

ISSN: 2602-3563



ACTA INFOLOGICA (ACIN)

JUNE, 2021
Volume: 5 | Issue: 1

ISTANBUL UNIVERSITY, INFORMATICS DEPARTMENT

dergipark.gov.tr/acin
acin.istanbul.edu.tr



İSTANBUL
UNIVERSITY
PRESS

Dizinler / Indexing and Abstracting
TÜBİTAK-ULAKBİM TR Dizin





Acta INFOLOGICA (ACIN) | ISSN: 2602-3563

Haziran (June) 2021
Cilt (Volume): 5 | Sayı (Issue): 1

Sahibi / Owner

Prof. Dr. Sevinç GÜLSEÇEN
İstanbul Üniversitesi, Enformatik Bölümü, İstanbul, Türkiye
Istanbul University, Informatics Department, Istanbul, Turkey

Sorumlu Yazı İşleri Müdürü / Responsible Manager

Prof. Dr. Sevinç GÜLSEÇEN
İstanbul Üniversitesi, Enformatik Bölümü, İstanbul, Türkiye
Istanbul University, Informatics Department, Istanbul, Turkey

Yazışma Adresi / Correspondence Address

İstanbul Üniversitesi Enformatik Bölümü
Kalenderhane Mahallesi, 16 Mart Şehitleri Caddesi, No: 8
Vezneciler, Fatih, İstanbul, Türkiye
Telefon / Phone: +90 212 440 00 00/10037
E-mail: acin@istanbul.edu.tr
<http://iupress.istanbul.edu.tr/tr/journal/acin/home>
<https://dergipark.org.tr/tr/pub/acin>

Yayıncı / Publisher

İstanbul Üniversitesi Yayınevi / Istanbul University Press
İstanbul Üniversitesi Merkez Kampüsü, 34452 Beyazıt,
Fatih, İstanbul, Türkiye
Telefon / Phone: +90 212 440 00 00

Dergide yer alan yazılardan ve aktarılan görüşlerden yazarlar sorumludur.
Authors bear responsibility for the content of their published articles.

Yayın dili Türkçe ve İngilizce'dir.
The publication languages of the journal are Turkish and English.

Haziran ve Aralık aylarında, yılda iki sayı olarak yayımlanan uluslararası,
hakemli, açık erişimli ve bilimsel bir dergidir.
*This is a scholarly, international, peer-reviewed and open-access journal
published biannually in June and December.*

Yayın Türü / Publication Type
Yaygın Süreli / Periodical

DERGİ YAZI KURULU / EDITORIAL MANAGEMENT

Baş Editör / Editor-in-Chief

Prof. Dr. Sevinç GÜLSEÇEN - İstanbul Üniversitesi, Enformatik Bölümü, İstanbul, Türkiye - gulsecen@istanbul.edu.tr

Baş Editör Yardımcısı / Co-Editor in Chief

Doç. Dr. Çiğdem EROL - İstanbul Üniversitesi, Enformatik Bölümü, İstanbul, Türkiye - cigdem@istanbul.edu.tr

Alan Editörleri / Section Editors

Dr. Serra ÇELİK - İstanbul Üniversitesi, Enformatik Bölümü, İstanbul, Türkiye - serra.celik@istanbul.edu.tr

Dr. Emre AKADAL - İstanbul Üniversitesi, Enformatik Bölümü, İstanbul, Türkiye - emre.akadal@istanbul.edu.tr

Dr. Fatma Önay KOÇOĞLU - Muğla Sıtkı Koçman Üniversitesi, Mühendislik Fakültesi, Yazılım Mühendisliği Bölümü, Muğla, Türkiye - fonaykocoglu@mu.edu.tr

Dil Editörleri / Language Editors

Elizabeth Mary EARL - İstanbul Üniversitesi, Yabancı Diller Yüksek Okulu, İstanbul, Türkiye
- elizabeth.earl@istanbul.edu.tr

Alan James NEWSON - İstanbul Üniversitesi, Yabancı Diller Yüksek Okulu, İstanbul, Türkiye
- alan.newson@istanbul.edu.tr

YAYIN KURULU / EDITORIAL BOARD

- Prof. Dr. Malgorzata PANKOWSKA - Katowice Ekonomi Üniversitesi, Enformatik Bölümü, Katowice, Polonya
- malgorzata.pankowska@ue.katowice.pl
- Prof. Dr. Mehpare TİMOR - İstanbul Üniversitesi, İşletme Fakültesi, İşletme Bölümü, İstanbul, Türkiye
- timorm@istanbul.edu.tr
- Prof. Dr. Meltem ÖZTURAN - Boğaziçi Üniversitesi, Uygulamalı Bilimler Yüksekokulu, Yönetim Bilişim Sistemleri Bölümü, İstanbul, Türkiye- meltem.ozturan@boun.edu.tr
- Prof. Dr. Orhan TORKUL - Sakarya Üniversitesi, Mühendislik Fakültesi, Endüstri Mühendisliği Bölümü, Sakarya, Türkiye - torkul@sakarya.edu.tr
- Prof. Dr. Selim YAZICI - İstanbul Üniversitesi, Siyasal Bilgiler Fakültesi, İşletme Bölümü, İstanbul, Türkiye
- selim@istanbul.edu.tr
- Prof. Dr. Sushil K. SHARMA - Ball State Üniversitesi, Miller İşletme Fakültesi, Indiana, ABD - ssharma@bsu.edu
- Prof. Dr. Türksel KAYA BENSGHİR - Hacı Bayram Veli Üniversitesi, İktisadi ve İdari Bilimler Fakültesi, İşletme Bölümü, Ankara, Türkiye - t.bensghir@hbv.edu.tr
- Prof. Dr. Üstün ÖZEN - Atatürk Üniversitesi, İktisadi ve İdari Bilimler Fakültesi, Yönetim Bilişim Sistemleri Bölümü, Erzurum, Türkiye - uozen@atauni.edu.tr
- Prof. Dr. Vesselina NEDEVA - Trakia Üniversitesi, Mühendislik ve Teknoloji Fakültesi, Elektrik Mühendisliği Bölümü, Bulgaristan - veselina.nedeva@trakia-uni.bg
- Prof. Dr. Yacine LAFİFİ - 8 Mayıs 1945 Guelma Üniversitesi, Matematik Fakültesi, Bilgisayar Bilimi ve Malzeme Bilimi, Bilgisayar Bilimi Bölümü, Cezayir - lafifi.yacine@univ-guelma.dz
- Prof. Dr. Elzbieta Magdalena WASIK - Mickiewicz Üniversitesi, Eski Cermen Dilleri Bölümü, Poznan, Polonya
- elawasik@amu.edu.pl
- Prof. Dr. Cem SÜTÇÜ - Marmara Üniversitesi, İletişim Fakültesi, Gazetecilik Bölümü, İstanbul, Türkiye
- csutcu@marmara.edu.tr
- Doç. Dr. Tuncay ÖZCAN - İstanbul Teknik Üniversitesi, İşletme Fakültesi, İşletme Mühendisliği Bölümü, İstanbul, Türkiye - tozcan@itu.edu.tr
- Dr. Öğr. Üyesi Enis KARAARSLAN - Muğla Sıtkı Koçman Üniversitesi, Mühendislik Fakültesi, Bilgisayar Mühendisliği Bölümü, Muğla, Türkiye - enis.karaarslan@mu.edu.tr
- Doç. Dr. Jan GUNCAGA - Bratislava Comenius Üniversitesi, Eğitim Fakültesi, İlköğretim Düzeyinde Doğa Bilimleri Didaktiği Bölümü, Slovakya - guncaga@fedu.uniba.sk
- Dr. Öğr. Üyesi Zerrin AYVAZ REİS - İstanbul Üniversitesi-Cerrahpaşa, Hasan Ali Yücel Eğitim Fakültesi, Bilgisayar ve Öğretim Teknolojileri Eğitimi Bölümü, İstanbul, Türkiye - ayvazzer@iuc.edu.tr
- Dr. Luis Miguel CARDOSO - Lizbon Üniversitesi, Portalegre Politeknik Enstitüsü, Karşılaştırmalı Araştırmalar Merkezi, Portekiz - lmcardoso@ippportalegre.pt
- Doç. Dr. Tetiana BONDARENKO - Ukrayna Mühendislik Pedagojisi Akademisi, Bilgi Bilgisayar Teknolojileri ve Matematik Bölümü, Kharkov, Ukrayna - bondarenko_tc@uipa.edu.ua
- Prof. Dr. Jagdish KHUBCHANDANI - New Mexico Eyalet Üniversitesi, Halk Sağlığı Departmanı, New Mexico, ABD - jagdish@nmsu.edu
- Doç. Dr. Natalija LEPKOVA - Vilnius Gediminas Teknik Üniversitesi, İnşaat Mühendisliği Fakültesi, İnşaat Yönetimi ve Emlak Bölümü, Litvanya - natalija.lepkova@vilniustech.lt

CONTENTS / İÇİNDEKİLER**Research Articles / Araştırma Makaleleri**

- A-Static-Periodic Solution Strategy for Dynamic Vehicle Routing Problem with Simultaneous Pickup and Delivery
Dinamik Eş Zamanlı Toplamalı ve Dağıtım Araç Rotalama Problemi için Statik Periyodik Çözüm Stratejisi
Mustafa Demirbilek..... 1
- Online Academic Resources With The Focus Of Eye Behaviors
Çevrimiçi Akademik Kaynaklarının Kullanımında Göz Davranışlarının Değerlendirmesi
İsmail Çetin, Emine Şendurur..... 13
- Adli Bilişim İncelemelerinde Şifre Kırma Yöntem ve Teknikleri
Password Cracking Methods and Techniques in Computer Forensic Investigation
İlker Kara..... 27
- Mobil Devlet Kullanımını Etkileyen Faktörlerin Belirlenmesi: Ankara İli Örneği
Determining Factors Affecting Mobile Government Usage: Ankara Province Case
Vildan Ateş..... 39
- Tedarikçilerin TOPSIS ile Seçilmesi ve Gelişim Durumlarının Sezgisel Optimizasyon ile Belirlenmesi
Selecting the Suppliers with TOPSIS and Determining Their Development Status by Using Heuristic Optimization
Ceylan Oğuz, İhsan Pençe, Melike Şişeci Çeşmeli, Özlem Çetinkaya Bozkurt..... 53
- Information Technology Governance Practices, Challenges and Effects on Enterprise Performance
Bilgi Teknolojileri Yönetişim Uygulamaları, Güçlükler ve Kurumsal Performansa Etkileri
Murat Paşa Uysal, Evrim Çetinkaya..... 65
- Two Level Kazakh Morphology
İki Düzeyli Kazak Morfolojisi
Züleyha Yiner, Atakan Kurt..... 79
- Audits in a Textile Company and Advantages
Bir Tekstil İşletmesinde Denetimler ve Avantajları
Mine Seçkin, Feden Koç, Ahmet Çağdaş Seçkin..... 99
- Süperpiksel Küme Bölgeleri Tabanlı Aktif Çevrit Ve GrabCut Sinerjisini Kullanarak İnsan Kulağının Otomatik Bölütlenmesi
Automatic Segmentation of the Human Ear Using Active Contour and GrabCut Synergy Based on the Superpixel Cluster Regions
Bahadır Karasulu..... 117
- Formel Yaklaşımlar ile Sürüm Değerlendirme
A Formal Methods Approach for Release Evaluation
Ebru Aydın Göl..... 129

Performance Comparison of Different Pre-Trained Deep Learning Models in Classifying Brain MRI Images <i>Beyin MR Görüntülerini Sınıflandırmada Farklı Önceden Eğitilmiş Derin Öğrenme Modellerinin Performans Karşılaştırması</i> Onur Sevli	141
Predicting Health Spending in Turkey Using the GPR, SVR, and DT Models <i>Türkiye Sağlık Harcamalarının GPR, SVR ve DT Modelleri ile Tahmini</i> Didem Güteryüz	155
Çocuklar İçin Otizm Spektrum Bozukluğunun Veri Madenciliği Yöntemleri ile Analizi* <i>Analysis of Autism Spectrum Disorder in Children with Data Mining Methods</i> Sümeyye Çelik, Melike Şişeci Çeşmeli	167
Keşifsel Veri Analizi Kullanarak Vakum Şarj Test Süresinin Azaltılması <i>Reducing Vacuum Charge Test Time Using Exploratory Data Analysis</i> Tuğba Demircan Kutluay, Süleyman Eken	187
Evaluation of Healthcare Economics of OECD Countries: Multi-Objective Statistical Optimization Model <i>OECD Ülkelerine ait Sağlık Ekonomilerinin Değerlendirilmesi: Çok Amaçlı İstatistiksel Optimizasyon Modeli</i> Abdulkadir Atalan, Cem Çağrı Dönmez	197

Review Articles / Derleme Makaleleri

Yapay Zekâ ve Tıbbi Görüntüleme Teknolojilerine Genel Bakış <i>An Overview of Artificial Intelligence and Medical Imaging Technologies</i> Furkan Atlan, İhsan Peñçe	207
Sağlık Hizmetleri ve Yönetiminde Yapay Zekâ <i>Artificial Intelligence in Health Services and Management</i> Betül Akalın, Ülkü Veranyurt	231

A-Static-Periodic Solution Strategy for Dynamic Vehicle Routing Problem with Simultaneous Pickup and Delivery

Dinamik Eş Zamanlı Toplamalı ve Dağıtım Araç Rotalama Problemi için Statik Periyodik Çözüm Stratejisi

Mustafa Demirbilek¹ 



ABSTRACT

Nowadays, logistic companies are dealing with the pressure of a highly competitive environment and trying to reduce operational costs. At the same time, customers are requiring not to wait so long to get their delivery and pickup demands satisfied. In this study, these two factors are considered with a-static-periodic solution strategy. The problem is modelled as Dynamic Vehicle Routing Problem with Simultaneous Pickup and Delivery (DVRPSPD) where delivery and pickup demands of customers continuously arrived in a planning horizon are satisfied at the same time and by the same vehicle. Furthermore, we proposed a solution methodology based on solving routing problems repeatedly for each time period by considering new and unvisited previous customers. A well-known heuristic, the Nearest Neighbourhood Algorithm (NNA), is used to solve VRPSPDs in each time period. Important real-life aspects such as demand fluctuations, different number of customers, and routing periods are integrated into the problem. According to results, the average waiting times per customer significantly increase whereas average travel times per visit notably decrease when the length of time period increases. If routes are constructed 8 times in a day, each customer waits an average 5 minutes compared to 66 minutes if it is made 4 times in the day. More routing plans cause more than 7-minute average travel times per visit. Different test settings are analysed and the results are explained to help decision makers find the best solution for both companies and customers.

Keywords: Vehicle routing problem with simultaneous pickup and delivery, The nearest neighbourhood algorithm, Dynamic customer arrivals, Decision making

ÖZ

Günümüzde lojistik firmaları yüksek rekabet ortamının getirdiği baskılarla mücadele etmekte ve operasyonel maliyetleri düşürmeye çalışmaktadırlar. Aynı zamanda müşteriler, dağıtım ve toplama taleplerinin karşılanması için çok beklemek istememektedirler. Bu çalışmada, bahsedilen iki faktör statik periyodik çözüm stratejisiyle beraber ele alınmıştır. Problem, bir planlama dönemi içinde sürekli olarak gelen müşterilerin, toplama ve dağıtım taleplerinin aynı zamanda ve araçla karşılandığı “Dinamik Eş Zamanlı Toplamalı ve Dağıtım Araç Rotalama Problemi (DEZTDARP)” olarak modellenmiştir. Bununla beraber yeni ve ziyaret edilmemiş eski müşterileri göz önüne alarak rotalama problemlerini her zaman periyodu için tekrar tekrar çözen bir çözüm metodu önerilmiştir. Literatürde çok kullanılan bir sezgisel olan En Yakın Komşu Algoritması, her periyotta EZTDARP’lerini çözmek için kullanılmıştır. Talep dalgalanması, farklı sayıda müşteri, farklı rotalama periyotları gibi gerçek hayatta karşılaşılabilecek önemli hususlar probleme entegre edilmiştir. Sonuçlara bakıldığında, rotalama periyotlarının süresi arttığı zaman müşteri başı ortalama bekleme sürelerinin ciddi şekilde arttığı, ziyaret başı ortalama seyahat sürelerinin ise azaldığı görülmüştür. Rotalama planları günde 8 sefer yapıldığında müşteri başı ortalama bekleme süresi 5 dakika iken, plan sayısı 4 olduğunda bekleme süresi 66 dakikaya yükselmektedir. Daha fazla rotalama ise ziyaret başı fazladan 7 dakika ortalama seyahat süresine neden olmaktadır. Karar vericilerin hem firmalar hem de müşteriler için en iyi çözümü bulmalarına yardımcı olmak amacıyla farklı testler analiz edilmiş ve sonuçları açıklanmıştır.

Anahtar kelimeler: Eş zamanlı toplamalı ve dağıtım araç rotalama problemi, En yakın komşu algoritması, Dinamik müşteri gelişleri, Karar verme

¹(Asst. Prof.), Gaziantep Islamic Science and Technology University, Faculty of Engineering And Natural Sciences, Department of Industrial Engineering, Gaziantep, Turkey

ORCID: M.D. 0000-0002-1520-2882

Corresponding author:

Mustafa DEMİRBILEK
Gaziantep Islamic Science and Technology University, Faculty of Engineering And Natural Sciences, Department of Industrial Engineering, Gaziantep, Turkey
E-mail address: mustafa.demirbilek@gibtu.edu.tr

Submitted: 26.11.2020

Revision Requested: 26.04.2021

Last Revision Received: 02.05.2021

Accepted: 31.05.2021

Published Online: 07.06.2021

Citation: Demirbilek, M. (2021). A-static-periodic solution strategy for dynamic vehicle routing problem with simultaneous pickup and delivery. *Acta Infologica*, 5(1), 1-12. <https://doi.org/10.26650/acin.831973>

1. INTRODUCTION

Vehicle Routing Problem (VRP) can be simply defined as finding the shortest route between a depot and geographically scattered points by satisfying a set of constraints. VRP has attracted many researchers' attention since the study of Dantzig and Ramster (1959), "The Truck Dispatching Problem." They aimed to satisfy demands of fuel stations when minimizing total travel times of a given truck fleet. A linear programming approach was used to solve the problem. Since then, many different types of VRP such as VRP with Time Windows, Stochastic VRP, Multi Depot VRP, Periodic VRP, Dynamic VRP, and different combinations of these have been studied. Not only problem types are evolved, but also a variety of solution methodologies have been developed. Exact solution methods, such as Branch-and-Bound Algorithms (Christofides & Eilon, 2006), Dynamic Programming (Columbus, 1972), Set Partitioning Formulations and Algorithms, and Commodity Flow Formulations and Algorithms, and heuristic/metaheuristic solution methods, such as The Savings Algorithm, Cluster-First-Route-Second Heuristics, Local Search, and Population Search algorithms, have been proposed over 50 years (Laporte, 2009).

VRP with simultaneous pickup and delivery (VRPSPD) is getting attention among other VRPs due to scarce resources, environmental sensitivities, and regulations (Demirbilek, 2020). In VRPSPD, vehicles deliver as well as collect goods at the same time. Each day in the planning horizon, a given number of vehicles start their travels from a depot, deliver and pick up goods to/from customers by visiting them only once, and return to the depot at the end of the day. As classical VRPs, the aim is to minimise total distances as well as to maximise total number of visited customers by considering the capacity of each vehicle (Goksal et al., 2013; Chen & Wu, 2006). The online order system of grocery stores is a good example for VRPSPD. Orders are distributed to customers while the same vehicles pick up undated goods or empty bottles. Because of regulations, some companies are responsible for products during their whole life cycles (as in the disposal of laser printer cartridges) (Montané & Galvão, 2006).

On the other hand, the dynamism issue that considers the arrival of customers over a planning time and assigns them to a schedule as soon as possible is becoming a more and more important issue since the hectic business life and the unwillingness of customers to wait for vehicles to get their packages delivered or picked up. Therefore, the number of studies about Dynamic VRP as well as VRPSPD has significantly increased recently.

In this study, a static-periodic-resolution-based solution methodology for a Dynamic VRPSPD (DVRPSPD) is proposed. The idea behind the solution methodology is to solve VRPSPD for a prespecified time period by covering new customers and previous customers that are not served until that time period. Therefore, customer arrivals during a planning horizon are taken into consideration in a dynamic fashion instead of optimizing routes with known customers before the planning horizon starts as in static solution approaches in the literature. A well-known heuristic algorithm, the Nearest Neighbourhood Algorithm (NNA), is used to solve VRPSPD in each period. The main contribution of the study is to examine relationships between average travel times for vehicles and the waiting times for customers under consideration of different routing period lengths and demand fluctuations.

In the next section, recent studies in the literature are reviewed. In Section 3, the mathematical formulation and NNA are explained. In Section 4, experimental setting, necessary data, and results are described. In the last section, we conclude our study and emphasize some assumptions and limitations.

2. LITERATURE REVIEW

Although many studies about DVRP and VRPSPD exist in the literature, it was observed that only the study of Aydogdu and Ozyoruk (2020) considered a similar problem. Therefore, VRPSPD studies are first reviewed, and differences between this study and the study of Aydogdu and Ozyoruk (2020) are discussed later in this section.

VRPSPD studies began with the study of Min (1989), "The Multiple Vehicle Routing Problem with Simultaneous Delivery and Pickup Points." He found optimum routes for vehicles that delivered library materials from a main library to 22 branches and picked up the materials from branches to carry to the main library. Hezer and Kara (2013) developed a metaheuristic

method, bacterial foraging optimization algorithm, for VRPSPD. The proposed algorithm ensured superior results for 24 out of 40 test sets compared to the heuristic solution method developed by Dethloff (2011). Nagy and Salhi (2015) developed a heuristic solution method to be able to solve both VRPSPD and VRP with Mixed Pickup and Delivery (VRPMPD). They also extended the solution method to solve the multi-depot VRPSPD. Dell'Amico et al. (2006) proposed a branch and price algorithm to solve VRPSPD for 40 requests. Wassan et al. (2007) developed a reactive tabu search metaheuristic that could control feasibility of predetermined moves quickly and acted on iterations to guide the search. Ai and Kachitvichyanukul (2009) presented a particle swarm optimization method. Subramanian et al. (2010) presented a parallel algorithm based on the sequential heuristics. They embedded the algorithm with a multi-start heuristic containing the VND integrated to an iterated local search framework. Avci and Topaloglu (2015) proposed an adaptive local search algorithm for solutions of both VRPSPD and VRPMPD. Li et al. (2015) considered Multi-Depot VRPSPD and developed an iterated local embedded adaptive neighbourhood selection approach. Results showed that the proposed algorithm was superior than a large neighbourhood search, particle swarm optimization, and an ant colony optimization approach. Zachariadis et al. (2016) considered two-dimensional loading constraints beside simultaneous linehauls and backhauls. Kalayci and Kaya (2016) developed an ant colony system empowered variable neighbourhood search algorithm, a hybrid algorithm that combines an ant colony system and a variable neighbourhood approach. Experiments showed that the proposed algorithm provided better results compared to the individual performance of two methods. Kececi et al. (2015) considered a heterogeneous vehicle fleet consisting of different type of vehicles with costs. They proposed a mathematical model and an insertion-based heuristic for the solution. Yazgan and Buyukyilmaz (2017) developed a greedy heuristic solution method and carried out a regression analysis to find the relationship between travel distances and the number of customers and the capacity of vehicles. Goksal et al. (2013) presented a discrete particle swarm optimization method and a variable neighbourhood descent algorithm for VRPSPD. The algorithm kept swarm diversity by improving randomly selected solutions in each iteration. Montane and Galvao (2006) modified a tabu search algorithm to become compatible with VRPSPD. They developed three types of movements to observe inter-route adjacent solutions: the relocation, interchange, and crossover movements. Table 1 shows publications in terms of problem types, years, and solution methods. The review study of Koc and Laporte (2018) in those VRPSPD studies were classified in terms of models, solution methods, variants, and industrial applications, and case studies are advised to interested readers.

Table 1

Publications in terms of problem types and solution method

Year	Author	Problem	Solution Methods
1989	Min	Book Delivery and Pickup	Clustered Based Heuristic Method
2001	Dethloff	VRPSPD	Mathematical Model and Heuristic Method
2005	Nagy and Salhi	VRPSPD, Multi-Depot VRPSPD	Mathematical Model and Insertion-based Heuristic
2006	Montane and Galvao	VRPSPD	Tabu Search Algorithm
2006	Dell'Amico et al.	VRPSPD	Branch and Price Algorithm
2007	Wassan et al.	VRPSPD	Reactive Tabu Search Algorithm
2009	Ai and Kachitvichyanukul	VRPSPD	Particle Swarm Optimization Algorithm
2010	Subramanian et al.	VRPSPD	Parallel Algorithm Based on the Sequential Heuristics
2013	Hezer and Kara	VRPSPD	Bacterial Foraging Optimization Algorithm
2013	Goksal et al.	VRPSPD	Discrete Particle Swarm Optimization
2015	Avci and Topaloglu	VRPSPD and VRPMPD	Adaptive Local Search Algorithm
2015	Li et al.	Multi-Depot VRPSPD	Adaptive Neighbourhood Selection
2015	Kececi et al.	VRPSPD with heterogeneous vehicle fleet	Mathematical Model and Insertion-based Heuristic
2015	Nagy and Salhi	VRPSPD	Mathematical Model and Variable Neighbourhood Search
2016	Zachariadis et al.	VRPSPD with two-dimensional loading constraints	Local Search Algorithm
2016	Kalayci and Kaya	VRPSPD	Ant Colony Optimization
2017	Yazgan and Buyukyilmaz	VRPSPD	Insertion-based Heuristic
2020	Demirbilek	VRPSPD	Mathematical Model
2020	Aydogdu and Ozyoruk	DVRPSPD	Random Iterative Local Search Variable Neighbourhood Descending

Aydogdu and Ozyoruk (2020) considered two time periods for VRPSPD. Furthermore, they covered dynamic demands, 10% to 50% of total daily demands. They developed the Random Iterative Local Search Variable Neighbourhood Descending algorithm as a solution methodology and compared results to results of the mathematical model. Their aim was to minimize total travel times. In our study, 2 to 8 time periods and 50% to 87.5% dynamism levels are considered. The relationships between travel and waiting times based on the different lengths of routing periods and demand fluctuations are also examined. On one hand, shorter planning horizons can provide shorter waiting times for customers. On the other hand, longer ones ensure optimal routes as well as shorter travel times for vehicles and companies. Therefore, understanding the relationship between waiting and travel times under different routing period lengths helps companies to decide the best period length that establishes a trade-off between customer satisfaction and profit.

3. MATERIALS AND METHODS

In this section, the mathematical model for VRPSPD is first introduced. Next, the NNA is explained in detail.

3.1. Mathematical Model for VRPSPD

There are n customers that must be visited by k homogeneous vehicles. Each customer has a certain amount of pickup and delivery. When a vehicle visits a customer, the amount of pickup and delivery demands of a customer cannot exceed capacity of the vehicle. Vehicles must start from and return to the depot. The mathematical formulation for VRPSPD is mostly taken from the study of Montane and Galvao (2006). However, the maximum distance constraint is removed, and the maximum daily working time for vehicles and service times for customers are added into the model. The following shows notations and decision variables:

Notations:

V : set of customers

V_0 : set of customers and the depot

K : set of vehicles

c_{ij} : distance between locations of customers i and j

p_j : pickup demand of customer j , $\forall j \in V$

d_j : delivery demand of customer j , $\forall j \in V$

Q : vehicle capacity

T : working time limit for vehicles

S : service time

Decision Variables:

$x_{ijk} = 1$, if vehicle k moves from customer i to j . 0, otherwise.

y_{ij} = Cumulative amount of pickup at node i and transported in arc (i, j)

z_{ij} = Cumulative amount of delivery at node i and transported in arc (i, j)

The corresponding mathematical formulation is given by

Minimize:

$$\sum_k^K \sum_i^{V_0} \sum_j^{V_0} c_{ij} x_{ijk} + \sum_k^K \sum_i^{V_0} \sum_j^{V_0} S x_{ijk} \quad i \neq j \quad (1)$$

Subject to:

$$\sum_k^K \sum_i^{V_0} x_{ijk} = 1, \quad \forall j \in V \quad (2)$$

$$\sum_i^{V_0} x_{ijk} - \sum_i^{V_0} x_{jik} = 0, \quad \forall j \in V_0, \forall k \in K, i \neq j \quad (3)$$

$$\sum_j^V x_{0jk} \leq 1, \quad \forall k \in K \quad (4)$$

$$\sum_i^{V_0} y_{ij} - \sum_i^{V_0} y_{ji} = p_j, \quad \forall j \in V \quad (5)$$

$$\sum_i^{V_0} z_{ij} - \sum_i^{V_0} z_{ji} = d_j, \quad \forall j \in V \quad (6)$$

$$y_{ij} + z_{ij} \leq Q \sum_k^K x_{ijk}, \quad \forall i, \forall j \in V \quad (7)$$

$$\sum_i^{V_0} \sum_j^{V_0} c_{ij} x_{ijk} + S \sum_i^{V_0} \sum_j^{V_0} x_{ijk} \leq T, \quad \forall k \in K, i \neq j \quad (8)$$

$$x_{ijk} \in \{0,1\}, \quad \forall i, \forall j \in V_0, \forall k \in K, i \neq j \quad (9)$$

$$y_{ij}, z_{ij} \geq 0, \quad \forall i, \forall j \in V_0, i \neq j \quad (10)$$

The objective function (1) is to minimize total travel and service times. Constraints (2) make sure that each customer is visited by exactly one vehicle. Constraints (3) ensure that the same vehicle arrives and departs from each customer. Constraints (4) show the maximum number of vehicles allowed. Constraints (5) and (6) are flow equations for pick-up and delivery, respectively; they guarantee that both are satisfied for each customer. Constraints (7) ensure that pickup and delivery demands do not exceed capacity of the vehicle moving from customer i to j . Constraints (8) show travelling and service times for a vehicle do not exceed the allowed working time. Constraints (9) and (10) demonstrate boundaries of decision variables.

3.2. The Nearest Neighbourhood Algorithm (NNA)

The NNA is one of the most popular heuristic solution methods to solve VRPs and variants in the literature. The main idea behind it is to find the nearest neighbour in terms of the distance to a reference point (customer or depot) and insert it to the schedule. This procedure is repeated until all customers are assigned or the schedule is no longer available. Since we deal with VRPSPD, the pickup and delivery demands of customers as well as capacity levels of vehicles as constraints must be considered. Therefore, when assigning customers in terms of seeking minimum distances, the demands of customers and

the capacities of vehicles are taken into consideration. If the delivery and pickup demands of a customer violates the capacity of a vehicle, the algorithm does not assign it to the schedule even though it has the minimum distance cost.

The other important factor is the waiting time. The NNA is originally designed to consider distances between customers when assignment decisions are made. However, waiting times for customers must be considered since the arrival of customers in different time periods causes significant waiting times. Let us assume that there are 3 vehicles, V_1 , V_2 , and V_3 , customers are visited by these three vehicles as illustrated in Fig. 1. In this example, two time periods exist, at morning and noon, and the rerouting activity must be made at noon. The last customers are visited at 11:00, 11:20, and 10:50 am by V_1 , V_2 , and V_3 , respectively. If the earliest time the first customer (red one) can be visited is 12:00, the algorithm checks waiting times, basically the differences between the finish times of the previous services and the earliest starting times of the next customers. It is assumed that waiting times, w_1 , w_2 , and w_3 , are 60, 40, and 70 minutes for V_1 , V_2 , and V_3 respectively and travel times between the last customers visited, and the first customer after rerouting, d_1 , d_2 , and d_3 , are 20, 35, and 10 minutes. The algorithm selects V_3 to visit the first customer if minimizing distances is considered as the objective function. However, the vehicle must wait 60 minutes to be able to start the service. If V_2 is selected to visit the customer, the vehicle spends 40 minutes, 35 minutes during traveling and 5 minutes during waiting. Note that we refer the cost of the travelling and waiting of vehicles as “average travel time per visit” in the remaining part of the study since fuel consumption of vehicles are ignored.

It is important to note that the NNA works for each time interval as a static fashion. In other words, the algorithm only considers new and unvisited customers when routes are constructed at the beginning of an associated period. Customers arriving after routes are constructed must wait till the next period. Therefore, each period in a day is actually a static VRPSDP. However, considering many periods and inserting customers arriving during the planning horizon to the daily routes adds a degree of dynamism to the problem.

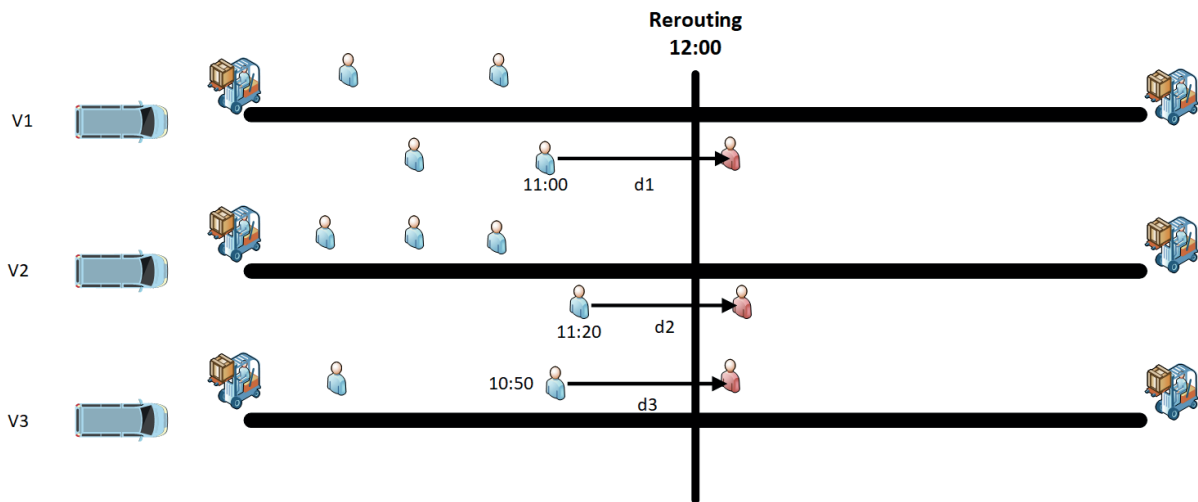


Figure 1. Illustration of the rerouting activity and the decision of customer visits

Fig. 2 shows the pseudocode for the NNA. First, the number of routing periods and vehicles are defined. In each period, a static time interval, new and unvisited customers are added to the “Customers” list on Line 4. Between Line 5 and 25, NNA works for the defined customers and vehicles. In each iteration, the cheapest insertion cost is calculated for each of the customers by considering all vehicles. After insertion costs are calculated for each customer as between Line 7 and 22, the customer with the cheapest insertion cost is assigned to the vehicle and removed from the list if the capacity and demand constraints are not violated. This procedure continues till all customers are scheduled.

4. COMPUTATIONAL EXPERIMENTS AND RESULTS

Three time periods, 60, 120, and 240 minutes, are considered to be able to understand the effect of period length on waiting and travel times for same customer sets. Since the daily working time of vehicles is defined as 480 minutes, there are 8, 4, and 2 routing periods that new customers can arrive and rerouting activities can be made. The expected number of customers

for a day is divided by the number of periods to be able to calculate the number of customers scheduled for the current time period. Let us assume that a 60-minute period length (8 periods) is considered and 48 customers for that day are expected. Thus, the number of customers that arrived in each period is 6. For a 120-minute time period (4 periods), the arrival of 12 customers will be expected. In real life, it is hard to observe that same number of customers arriving to the system in each time period, of course. By considering this possibility, period by period, demand fluctuations are also inspected. We randomly increase or decrease the number of customers by 30% from the expected values. Based on the above example, between 8 and 16 customers in each 120-minute routing period are expected.

Algorithm 1 The Nearest Neighbourhood Algorithm

```

1: Initiate Periods (P)
2: Set Vehicles (V)
3: for p= 1 To P do
4:   Set New and Unvisited Customers (C)
5:   while Customers list is not empty do
6:     MinCustomerCost ← ∞
7:     for c= 1 To C do
8:       MinVehicleCost ← ∞
9:       for v= 1 To V do
10:        Calculate insertioncost of Customers[c] for Vehicles[v]
11:        if insertionCost =< MinVehicleCost then
12:          if c satisfies capacity and time constraints for Vehi-
13:          cles[v] then
14:            MinVehicleCost←insertionCost
15:            Index←v
16:          end if
17:        end if
18:      end for
19:      if MinVehicleCost =< MinCustomerCost then
20:        MinCustomerCost←MinVehicleCost
21:        CustomerIndex←c
22:      end if
23:    end for
24:    Vehicles [Index]← Customers[CustomerIndex]
25:    Remove Customers[CustomerIndex] from Customers list
26:  end while
27:  Move Next Period
28: end for
29: Print Waiting and Travel Times

```

Figure 2. Pseudocode for NNA modified for DVRPSPD

There are some methods that can be used to measure the degree of dynamism (DD) in the literature. The method of Lund et al. (1996) is one of the most common methods. In this method, the DD is calculated by dividing the number of dynamically arriving customers by the total number of customers. In this study, dynamism levels change according to the length of periods since the algorithm starts with the number of customers arriving before the first period. Therefore, the first period in each day begins with known customers, and the remaining periods contain dynamic customers. In this situation, the DD level increases with the number of time periods accordingly. For example, if the 120-minute period length is considered, there are four routing periods in the 480-minute day length. The first period is composed of known customers. Thus, the remaining three contain dynamic customers, and the DD will be 0.75 (3/4). In this condition, there are three DD levels, 0.875, 0.75, and 0.5 for 60, 120, and 240-minute period lengths. Note that the case that all customers are known at the beginning is also considered for the comparison purpose.

As it is mentioned before, the day length is 480 minutes. The service time for each customer is 5 minutes. To test both models only for one customer set does not provide accurate results since results can significantly change for another customer set. Therefore, 30 sets that include different customers' data are taken into consideration. For each test, locations of customers and pickup and delivery demands are changed. It can be considered as operations of a company for 30 days. Three customer sets, 48, 96, and 144, are considered. Delivery demands of customers are assigned uniformly between 4 and 10 packages, and pickup demands are assigned uniformly between 3 and 7 packages. There is a different number of vehicles depending on the number of customers. Each vehicle capacity is 75 packages. A sufficient number of vehicles, as all customers can be visited by satisfying their pickup and delivery demands, is defined. The total area size is 3600 km² (60 km X 60 km) and represented, $X \in [0,60]$ and $Y \in [0,60]$, in the coordinate system. The depot is located at the middle of the service area. All related data are shown in Table 2. Finally, NNA is coded with Python Programming Language, and all tests are conducted on a PC with an Intel i5 7200U 2.5 GHz CPU and 8 GB Ram.

Table 2

Test settings

Daily working time (minute)	480
Service time (minute)	5
Service area	$X \in [0,60]$ and $Y \in [0,60]$
Number of customers	48,96,144
Demand fluctuation	$\pm 30\%$
Vehicle capacity (package)	75
Pickup demand (package)	Uniform (3,7)
Delivery demand (package)	Uniform (4,10)
Time period (minute)	60,120,240
Degree of dynamism	0.875,0.75,0.5

Table 3 shows travel times (minute) per visit and waiting times (minute) per customer under the stable demand structure. Confidence intervals with a 95% confidence level are also provided. For example, an average travel time per visit can be between 33.85 and 34.95 minutes (34.40 ± 0.55) with 95% confidence level if the 48 customers and 60-minute routing period test set are considered. First, average travel times per visit decrease when the length of periods increases as expected. A 480-minute period demonstrates the case that all customers are known at the beginning. Compared to the 60-minute time period (87.5% DD), average travel times per visit decrease by the least, more than 50%, if all customers are known before optimizations start. This is called "Value of Information" and shows that the difference between fully static and dynamic solutions for the same problem set. When the number of customers increases from 48 to 144, average travel times per visit slightly decrease. However, once we know all customers at the beginning, average travel times per visit decrease by 11% and 18% when increasing the number of customers to 96 and 144. In this study, waiting times for customers that represent differences (minute) between arrival times and starting times to services are also taken into consideration. Waiting times for a fully static case (480-minute day length) are not calculated since we should know when each customer arrives on the previous day for accurate calculations. This is a limitation for this study and discussed in the Conclusion Section. Although average travel times per visit change between 9% and 23% period by period, waiting time sensitivity based on different time periods is relatively enormous. If the period length increases 60 to 120, each customer must wait almost 62 minutes more to

be serviced whereas vehicles spend 7 minutes less per visit during travelling. When the number of customers increases, waiting times per customer increase whereas average travel times stay almost the same.

Table 3

*Average travel times (minute) per visit and waiting times (minute) per customer under the stable demand structure (*CI: Confidence Interval +/-)*

Customers	Time Periods	Travel Time (TT)	*CI for TT (\pm)	Waiting Time (WT)	*CI for WT (\pm)
48	60	34,40	0,55	4,63	0,53
	120	31,15	0,80	34,07	1,67
	240	27,84	0,60	66,39	2,60
	480	16,67	0,34
96	60	34,31	0,55	5,42	0,39
	120	28,89	0,35	41,05	1,03
	240	26,57	0,41	76,83	1,36
	480	14,81	0,21
144	60	33,47	0,44	6,79	0,32
	120	28,41	0,32	41,07	0,82
	240	25,59	0,31	78,73	1,05
	480	13,63	0,12

Table 4 shows average travel times (minute) per visit and waiting times (minute) per customer under demand fluctuations. As mentioned, the number of customers arriving in each period changes $\pm 30\%$ based on expected values. First, average travel times per visit increase compared to the stable demand scenario. On the other hand, average waiting times per customer are slightly less than average waiting times in the stable demand scenario. Similarly, average waiting times are very sensitive to time periods. When the length of time periods increases, waiting times per customer sharply escalate. Overall, results under demand fluctuations follow the same pattern with results under the stable demand condition. Although it sometimes seems there are only several minute differences per customer between two results, it has a significant effect when considering all customers. For example, the difference between the average travel time of a 60-minute period and a 240-minute period is around 8 minutes per visit. If we consider 144 customers/visits, it costs 1152 minutes daily. It means that vehicles spend more than 19 hours during daily travels if the 60-minute routing period is applied.

Table 4

*Average travel times (minute) per visit and waiting times (minute) per customer under demand fluctuations (*CI: Confidence Interval +/-)*

Customers	Time Periods	Travel Time (TT)	*CI for TT (\pm)	Waiting Time (WT)	*CI for WT (\pm)
48	60	38,99	0,76	3,67	0,37
	120	33,76	0,72	32,00	1,27
	240	29,60	0,95	68,40	2,55
	480	16,67	0,34
96	60	35,65	0,44	5,73	0,56
	120	31,27	0,54	35,41	1,24
	240	27,34	0,85	74,77	2,61
	480	14,81	0,21
144	60	35,65	0,50	6,23	0,46
	120	29,72	0,45	40,34	1,12
	240	27,23	0,74	82,07	2,29
	480	13,63	0,12

5. CONCLUSION

Vehicle Routing Problem with Simultaneous Pickup and Delivery is getting more attentions day by day since optimal or near optimal solutions significantly decrease operational costs of companies by delivering and collecting goods at the same time. Although many studies that consider static problem settings where all customer information is known at the beginning exist in the literature, the dynamism issue where customers arrive during a planning period and schedules are evolving dynamically is rarely taken into account. In this study, we consider the dynamism issue as static routing periods when optimizations are made. In other words, we find optimal routes for new customers and unvisited customers assigned from previous periods in each routing period. The Nearest Neighbourhood Algorithm is used to solve optimization problems quickly. Some real-life

aspects such as different numbers of customers, daily working times for vehicles, service times for customers, and demand fluctuations for routing periods are considered as parameters and constraints. Moreover, different degree of dynamism (DD) levels, 50% to 87.5%, are examined to observe how different DD levels affect results. Since results can significantly change based on divergent data sets, 30 different data sets that include a variety of customers' locations and pickup and delivery demands are evaluated, and confidence intervals are calculated for average travel times per visit and waiting times per customer to provide more accurate results for readers.

First, as expected, considering dynamic customer arrivals significantly increases average travel times per visit. On the other hand, when the length of periods decreases, the average waiting times also sharply decrease. If routes are constructed 8 times a day, each customer waits an average 5 minutes compared to 66 minutes if it is made 4 times a day. In this situation, each vehicle must spend seven more minutes visiting each customer. When the demand fluctuation case is examined, longer average travel times are observed. Furthermore, average travel times per visit decrease, and waiting times per customer increase slightly when the number of customers goes up. Note that several minute differences among results per customer/visit make a huge impact when considering the high number of customers. On one hand, companies try to avoid unnecessary travel times to reduce operational costs. On the other hand, customers demand to be served as soon as possible. In this study, we attempt to show the relationship between travel times and waiting times to help decision makers find the best solution for both companies and customers.

There are some limitations and assumptions for this study as well. First, when calculating waiting times, it is assumed that customers in the first period are ready to be serviced. However, these customers are supposed to arrive the previous day, so waiting times can be higher than those calculated here. Although this can be achieved by defining interarrival times for customers, we have a different number of customers visited scenario by scenario in this time, and it causes miscalculations and miscomparisons for travel and waiting times. Furthermore, pickup and delivery demands and the capacity of vehicles are considered the same, test by test. Testing a variety of demands and capacity levels can also be useful. However, the first aim of this study is to examine how the length of periods affect waiting and travel times per customer/visit.

In future research, each customer can be considered as a decision point. Instead of considering a specific time period, whenever a new customer arrives to system, the routing and scheduling activity is made to find the optimum visit time for the customer.

Peer-review: Externally peer-reviewed.

Conflict of Interest: The authors has no conflict of interest to declare.

Grant Support: The author declared that this study has received no financial support.

Hakem Değerlendirmesi: Dış bağımsız.

Çıkar Çatışması: Yazar çıkar çatışması bildirmemiştir.

Finansal Destek: Yazar bu çalışma için finansal destek almadığını beyan etmiştir.

References/Kaynaklar

- Ai, T. J., & Kachitvichyanukul, V. (2009). A particle swarm optimization for the vehicle routing problem with simultaneous pickup and delivery. *Computers & Operations Research*, 36(5), 1693-1702.
- Avcı, M., & Topaloglu, S. (2015). An adaptive local search algorithm for vehicle routing problem with simultaneous and mixed pickups and deliveries. *Computers & Industrial Engineering*, 83, 15-29.
- Aydoğdu, B., & Özyörük, B. (2020). Dinamik eş zamanlı toplama dağıtım araç rotalama probleminin çözümü için matematiksel model ve sezgisel yaklaşım: Rassal iteratif yerel arama değişken komşu iniş algoritması. *Journal of the Faculty of Engineering & Architecture of Gazi University*, 35(2).
- B. Columbus, "Book Reviews," no. September, pp. 558–559, 1972.
- Chen, J. F., & Wu, T. H. (2006). Vehicle routing problem with simultaneous deliveries and pickups. *Journal of the Operational Research Society*, 57(5), 579-587.
- Christofides, N., & Eilon, S. (2006). An algorithm for the vehicle-dispatching problem. *Journal of the Operational Research Society*, 20(3), 309-318.
- Dantzig, G. B., & Ramser, J. H. (1959). The truck dispatching problem. *Management science*, 6(1), 80-91.
- Dell'Amico, M., Righini, G., & Salani, M. (2006). A branch-and-price approach to the vehicle routing problem with simultaneous distribution and collection. *Transportation science*, 40(2), 235-247.

- Demirbilek, M. (2020). A Tactical/Strategic Level Cost Analysis Based On Visit Time Preferences For Vehicle Routing Problem With Simultaneous Pickup And Delivery. *European Journal of Technique (EJT)*, 10 (2), 301-312.
- Dethloff, J. (2001). Vehicle routing and reverse logistics: the vehicle routing problem with simultaneous delivery and pick-up. *OR-Spektrum*, 23(1), 79-96.
- Goksal, F. P., Karaoglan, I., & Altıparmak, F. (2013). A hybrid discrete particle swarm optimization for vehicle routing problem with simultaneous pickup and delivery. *Computers & Industrial Engineering*, 65(1), 39-53.
- Hezer, S. & Kara, Y. (2013). Eşzamanlı dağıtım ve toplama araç rotalama problemlerinin çözümü için bakteriyel besin arama optimizasyonu tabanlı bir algoritma. *Gazi Üniversitesi Mühendislik-Mimarlık Fakültesi Dergisi*, 28(2).
- Kalaycı, C. B., & Kaya, C. (2016). An ant colony system empowered variable neighbourhood search algorithm for the vehicle routing problem with simultaneous pickup and delivery. *Expert Systems with Applications*, 66, 163-175.
- Kececi, B., Altıparmak, F., & Kara, I. (2015). Heterogeneous vehicle routing problem with simultaneous pickup and delivery: Mathematical formulations and a heuristic algorithm. *Journal of the Faculty of Engineering and Architecture of Gazi University*, 30(2), 185-195.
- Koç, Ç., & Laporte, G. (2018). Vehicle routing with backhauls: Review and research perspectives. *Computers & Operations Research*, 91, 79-91.
- Laporte, G. (2009). Fifty years of vehicle routing. *Transportation science*, 43(4), 408-416.
- Li, J., Pardalos, P. M., Sun, H., Pei, J., & Zhang, Y. (2015). Iterated local search embedded adaptive neighbourhood selection approach for the multi-depot vehicle routing problem with simultaneous deliveries and pickups. *Expert Systems with Applications*, 42(7), 3551-3561.
- Lund, K., Madsen, O. B., & Rygaard, J. M. (1996). Vehicle routing problems with varying degrees of dynamism. IMM Institute of Mathematical Modelling.
- Min, H. (1989). The multiple vehicle routing problem with simultaneous delivery and pick-up points. *Transportation Research Part A: General*, 23(5), 377-386.
- Montané, F. A. T., & Galvao, R. D. (2006). A tabu search algorithm for the vehicle routing problem with simultaneous pick-up and delivery service. *Computers & Operations Research*, 33(3), 595-619.
- Nagy, G., & Salhi, S. (2005). Heuristic algorithms for single and multiple depot vehicle routing problems with pickups and deliveries. *European journal of operational research*, 162(1), 126-141.
- Subramanian, A., Drummond, L. M. D. A., Bentes, C., Ochi, L. S., & Farias, R. (2010). A parallel heuristic for the vehicle routing problem with simultaneous pickup and delivery. *Computers & Operations Research*, 37(11), 1899-1911.
- Wassan, N. A., & Nagy, G. (2014). Vehicle routing problem with deliveries and pickups: modelling issues and meta-heuristics solution approaches. *International Journal of Transportation*, 2(1), 95-110.
- Yazgan, H. R., & Büyükyılmaz, R. G. (2017). Eş zamanlı topla dağıt araç rotalama problemine sezgisel bir çözüm yaklaşımı. *Sakarya Üniversitesi Fen Bilimleri Enstitüsü Dergisi*, 22(2), 436-449.
- Zachariadis, E. E., Tarantilis, C. D., & Kiranoudis, C. T. (2016). The vehicle routing problem with simultaneous pick-ups and deliveries and two-dimensional loading constraints. *European Journal of Operational Research*, 251(2), 369-386.

Online Academic Resources With The Focus Of Eye Behaviors*

Çevrimiçi Akademik Kaynaklarının Kullanımında Göz Davranışlarının Değerlendirmesi

İsmail Çetin¹ , Emine Şendurur² 



*This submission appeared as part of İsmail Çetin's master thesis, titled "The Determine of Access Problems to Online Academic Resources and Usability Evaluation of Interfaces" with advisor Assoc. Prof. Dr. Emine Şendurur

¹(Res. Asst.), Ondokuz Mayıs University, Computer and Department of Instructional Technology Education, Samsun, Turkey

²(Assoc. Prof.), Ondokuz Mayıs University, Computer and Department of Instructional Technology Education, Samsun, Turkey

ORCID: İ.Ç. 0000-0002-7865-6080;
E.Ş. 0000-0002-0340-6378

Corresponding author:

İsmail ÇETİN
Ondokuz Mayıs Üniversitesi, Bilgisayar ve Öğretim Teknolojileri Eğitimi Bölümü,
Samsun, Türkiye
E-mail address: ismail.cetin@omu.edu.tr

Submitted: 08.09.2020

Revision Requested: 30.11.2020

Last Revision Received: 00.00.000

Accepted: 09.12.2020

Published Online: 12.03.2021

Citation: Çetin, I., & Sendurur, E. (2021). Online academic resources with the focus of eye behaviors. *Acta Infologica*, 5(1), 13-26.
<https://doi.org/10.26650/acin.791978>

ABSTRACT

The aim of this study is to investigate the eye movements of participants during academic searching tasks. Since the Web of Science, Google Scholar, and Council of Higher Education Thesis Center are among the popular academic search platforms in Turkey, we focused on the usability of them. During the preliminary analysis, we explored the most problematic tasks, and then analyzed the eye movements. The problematic tasks were detected with respect to the unexpected differences of completion duration; the failure of completion status; unexpected number of clicks; high number of fixations; high fixation duration; unrelated fixation areas; and think aloud data. Tobii Studio 3.4.2 was utilized for both the collection and analysis of eye-tracking data. 21 academicians participated to the study. The results of the analysis indicated that due to interface design, the eye movements of participants tended to focus on unrelated parts of the screen. As a result, certain parameters of eye tracking demonstrated how participants' eye movements tended to generate intense fixations, increased amount spent for unrelated areas, intense mouse clicks, and increased total amount spent for tasks. In addition to those parameters, users' satisfaction levels showed a decrease as a result of unexpected experiences.

Keywords: Usability, Eye Tracking, Eye Movements, Online Search, Online Academic Sources

ÖZ

Bu çalışmanın amacı, çevrimiçi akademik kaynak kullanımında göz hareketlerini incelemektir. Bu kapsamda ülkemizde en yaygın kullanıma sahip çevrimiçi akademik kaynaklardan olan Web of Science, Google Akademik (Google Scholar) ve YÖK Tez Merkezi web sitelerinin kullanılabilirlik testi esnasında en çok problem yaşanan görevler belirlenmiş ve kullanıcıların bu görevleri gerçekleştirirken sergilediği göz hareketleri analiz edilmiştir. En çok problem yaşanan görevler, görevlerin tamamlanması gereken süre ile tamamlanma süreleri arasındaki farkın yüksek olmasına, görevlerin tamamlanma durumlarına, fare tıklama sayılarına, sabit bakış sayılarına, sabit bakış sürelerine, sabit bakış yerlerine ve kullanıcı yorumlarına bakılarak belirlenmiştir. Bu işlemler için Tobii Studio 3.4.2 programı kullanılmıştır. Araştırma, çevrimiçi akademik kaynakları kullanma konusunda en az 2 yıl tecrübesi olan 21 akademisyen ile gerçekleştirilmiştir. Çalışmada geçerlilik için veri çeşitlenmesi sağlanarak nicel ve nitel veriler birlikte kullanılmıştır. Bulgular sonucunda ara yüz tasarımındaki sıkıntılardan dolayı kullanıcıların yanlış yerlerde ve daha yoğun sabit bakış sergilediği, bu durumun da sabit bakış süreleri, fare tıklama sayıları ve görev tamamlama sürelerini olumsuz etkilediği ve memnuniyetlerini düşürdüğü belirlenmiştir. Bu problemlerin sebepleri incelenmiş ve çözüm için önerilerde bulunulmuştur.

Anahtar kelimeler: Kullanılabilirlik, Göz İzleme, Göz Hareketleri, Çevrimiçi Arama, Çevrimiçi Akademik Kaynaklar

1. INTRODUCTION

The concept of information searching has been transformed into different formats due to rapid changes in electronic and digital sources. In the field of academia, the way to access any information has become easier in comparison to the past. Traditional printed resources have certain limitations. For example, borrowing it from a library may not be possible, but electronic or digital versions are accessible all the time. Searching for specific information within pages can be a challenging task for printed sources, but it is much easier in digital versions. Such different dimensions introduced by e-sources facilitate the effort spent by the researcher as well as the time spent (Kazan, 2017).

Rapidly increasing amounts of online information is expected to reach 35 zettabytes by the end of 2020 (Agrawal & Nyamful, 2016). Such an imbalanced increase in the electronic world is vulnerable to the quality of the information, i.e. not all information in that world is valid and reliable. Since users deal with different types of user interfaces on the web, experiencing various usability problems is very likely. These problems can result in frustrations for users. System performance such as response time, availability of reliable and valid content, broken or outdated links, and inadequate web search results are some of the things that affect user experience. (Brandt et al., 1996; Kovacs & Elkordy, 2000; Land, 2000). In addition, the design and usability of user interfaces may affect the way people access and use online information. In typical software, 48% of codes are related to interfaces (Myers & Rosson, 1992), which might be a reason why the design of appealing interfaces should be achieved. According to Norman (1988), reducing the possibility of errors is an important mission of any computer programs as well as making the recovery of errors possible. Similarly, in Nielsen's heuristics, the importance of interface designs was also emphasized. Since the online search is important for today's computer users, the usability of available tools is quite important.

Usability as a term is not very easy to define because the measurement of it is considerably hard and sophisticated, and therefore defining it in a scientific framework took time and was not always accepted by everyone (Hertzum et al., 2009). Nevertheless, in daily life, people generally use the semantic meaning of the term rather than the terminological one. For example, in ads, one can come across slogans such as "easy cleaning", "make your life easier", "easy to connect", "friendly screen". These are all examples of using the term in everyday meaning. From its semantic meaning, the term usability can also be expressed in the following terms: "learnability", "easy to use", "user friendly", and "transparent to users" (Cagiltay, 2011).

In the literature, the term usability can also be defined in various ways by different researchers. For example, according to Nielsen (2012), usability is a qualified indicator showing how easy to use a system is in addition to consisting of methods for designing usable systems. Shackel (2009) defines usability as the ability to complete the given tasks easily and effectively within the given scenario providing support. In the International Organization for Standardization (ISO) standards book, usability is located into two different parts. In the ISO 9241-11 (2018) part, the definition focuses on effectiveness, efficiency, and satisfaction within a context. In the ISO/IEC 2510 (2011) part that covers a wide range of computer related systems, the usability is associated with quality within effectiveness, efficiency, satisfaction, freedom from risk, and context coverage characteristics.

Although there are clearly defined standards, the usability testing is not very easy to pursue because there are lots of elements to be considered. Technical problems, calibration problems and ethical principles are some of these elements. The usability tests can be designed upon users, experts, models, and design principles. Each approach has its own advantages and disadvantages. Moreover, utilizing different test methods can generate varying results (Cagiltay, 2011). One of these approaches is Eye-tracking methodology. This technique is a system that determines eye movements and focus areas of the eye (Almedia et al. 2011). Eye-tracking methodology is a relatively objective technique used to collect data for user testing. As the user interacts with the interface, his/her fixations can be viewed through this technique in a detailed way (Nakatani & Pollatsek 2004; Renshaw & Webb 2006; Scheiter & Van Gog 2009). Such a technique enables the researchers and designers to diagnose the errors, misunderstandings, misguidance, and the most attractive locations on the screen (Pernice & Nielsen, 2009). Moreover, the researchers or designers can save time by locating the exact issues related to usability.

There are various studies investigating usability through the utilization of eye trackers. University web sites, e-libraries, governmental sites, handbooks, educational multimedia environments, and online academic resources have been the popular environments for usability studies. As a result of those, the researchers revealed many practical implications including misleading buttons, links, colors, and so on (Alemdag & Cagiltay 2018; Cetin & Sendurur, 2016; Lund, 2016; Yavuz et al. 2016; White, 2016). Eye tracking is a technique that has been used in many different areas of research other than usability. Factors affecting web visits, the effects of visuals on online reading performance, behavior of children using the Internet, web searching behavior and the quality of online ads are some of the other areas in which eye tracking is used (Beymer et al., 2007; Buscher et al., 2010; Byerly, 2007; Cutrell & Guan, 2007; Pan et al., 2004). In addition, the eye tracking technique has been used to examine the cognitive process of users on search engine results pages (Lewandowski et al., 2018).

Scholars and graduate students generally use online academic resources due to their convenience of accessibility and coverage. Instead of searching among hardcopies of journals, searching with the help of e-databases, search engines, etc. is more convenient for many researchers. Unlike general purpose online tools like Google search engine, the online academic tools have specific user profiles, which are scholars. Nevertheless, the design of interfaces can hinder the search process. In the literature, there are many usability studies, but they are mostly limited with library resources. In order to provide guidance with the designers as well as publishers, this study aims to fill the gap with a specific focus of comparing the usability of three popular online academic resources in a qualitative way: Web of Science, Google Scholar, and National Thesis Center (NTS). The main research question of the study is as follows:

RQ: How do users' eye behaviors change when the design of the interface has some usability problems?

2. METHOD

The research design of this study follows the simultaneous mixed methodology (Tashakkori and Teddlie 1998). During the user test sessions, both qualitative and quantitative data were collected, and then they were analyzed and interpreted in a complementary manner. Table 1 summarizes the design.

Table 1

Research design

Data Collection Method	Data Analysis Method
Eye Tracking Record	Quan – Descriptive Statistics (Tobii Studio 3.4.2) Qual – Content Analysis
Observation Form	Qual – Content Analysis
Satisfaction Form	Qual – Content Analysis

2.1. Participants

In this study, the convenient sampling method was used to select participants because they have to travel to the human-computer interaction laboratory, and spend time for the session. The participants of this study were selected according to their experience on using online academic resources because the interface should be familiar to the users (Dumash and Redish 1999). The subject area of researchers varied. According to Virzi (1992), 5 participants are enough for user tests. However, researchers like Lindgaard and Chattratichart (2007) were against this idea. Nielsen and Landauer (1993), on the other hand, made a classification depending on the scale of the research, and claimed that 7-15 could sometimes reveal the majority of the errors. In reference to Nielsen and Landauer (1993), we included 21 users in this study. 9 of them were male, and 12 of them were female. Detailed demographics were presented in Table 2.

Table 2
Demographics of Participants

	n	%
Gender		
Female	12	57.1
Male	9	42.9
Age		
23-27	9	42.9
28-32	7	33.3
33-38	5	23.8
Experience of Online Academic Resources		
2-3 years	10	47.7
5-6	9	42.9
7-10	2	9.4
Department		
Civil Engineering	5	23.9
Computer Education & Instructional Technologies	4	19
Industrial Engineering	3	14.3
Electronic & Electrical Eng.	2	9.5
Chemistry	2	9.5
Computer Engineering	2	9.5
Geography	2	9.5
Informatics	1	4.9

2.2. Instruments & Apparatus

Before deciding on the tasks to be included in the experiment, the researchers talked to and observed the scholars around them. As a result of those unstructured and informal interviews, a group of frequently performed tasks emerged and formed the boundaries of our user tests. Finding a specific publication of a scholar, ordering the publications with regards to given criteria, and accessing dissertations published on a given date were the initial tasks that we aimed to include. In order to decide on the online academic resources, we delivered a questionnaire (Author(s), 2016). In the survey study, users were asked about the online academic resources they use most frequently, the reasons for choosing this resource, and the most common problems they encounter while accessing these resources. Since the detailed results are not the main concern of this study, the detailed qualitative analysis of user experience of hard tasks is the emphasis of the current study. In a preliminary study, Author(s) (2016) explored that the following two tasks were the most complicated ones according to users, and the analysis of eye movement approved that claim which is why this study has a deep focus on the following two tasks:

Task 1: List the studies of which title includes “instructional design” keywords.

Task 2: Download the dissertation written by “* (removed for blind review)”.

The tasks were the main instruments, but in order to observe and record data, Tobii Studio 3.4.2 was utilized as another instrument. Tobii Studio is an eye-tracking platform for the recording and analysis of eye gaze data. In addition to screen recordings, participants’ voice recordings were also important for gaining insights into what they were thinking as they moved on. There was a Tobii T120 eye tracker with 0.5 degree of sensitivity integrated into the desktop with a 17-inch monitor. The screen resolution was 1024x768 during data collection. The room was surrounded by cameras to record gestures and mimics. The participants were given wearable microphones for voice recording. The chair was adjustable and thus the height variety of participants was not a problem. Although certain head movements were tolerated by the eye tracker, we advised participants to keep an approximate 60 cm distance from the tracker (Tobii Studio 2008).

2.3. Data collection & analysis

Before starting the test, the users were informed about the procedure and ethical issues. They were free to stop, leave or have a break whenever they wanted. Then, the tasks were given in a written format and stayed in front of them during the test. The calibration was run for each participant to increase the reliability of recordings. 9 point calibration was conducted. The author(s) did not intervene in the sessions, i.e. users were encouraged to perform as if they were at home/library/office. They were also encouraged to think aloud. An average session took about 45 minutes. The users were expected to complete each task separately. No stimulus was presented at the beginning. In the beginning of the experiment the web browser was ready to be used. At the end of each session, users were asked to complete the satisfaction questions. The experiment was implemented in a public university's human-computer interaction laboratory.

The usability criteria consist of three facets: effectiveness, efficiency, and satisfaction. In order to evaluate the effectiveness of the chosen academic resources, we focused on the completion status of tasks and tried to investigate the reasons for failure. Efficiency in the current user test was decided according to the following metrics: time to first fixation, first fixation duration, total fixation duration, fixation count, time to first mouse click, time from first fixation to next mouse click, and mouse click count. For the third dimension, satisfaction, we assigned each user a satisfaction questionnaire after the test. All data were analyzed with the combination of both qualitative and quantitative analysis methods. Tobii Studio 3.4.2 program mostly benefited from quantitative analysis. Think aloud data, observation notes, and satisfaction questionnaires were analyzed in a qualitative way utilizing content analysis techniques. At the end of the analysis, the researchers diagnosed the two tasks that were perceived as the most difficult ones. As a result, the eye movements of users were closely analyzed in this part of the study.

3. RESULTS

When deciding on the most problematic tasks to be included in this study, we focused on the gap between expected and observed task completion duration, the task completion status, number of mouse clicks and fixations, fixation durations and locations as well as user comments. We examined gaze plots of users for detailed eye movements. As a result of that preliminary analysis, we decided to include two tasks, which seemed problematic.

The critical screen parts, i.e. the related areas leading to the completion of tasks, consisted of areas of interest (AOI) in this study. In the first task, locating and clicking the detailed search menu was a critical action for the task completion. In the second task, locating where to download the dissertation was an important phase. That's why the menu area was signed as an AOI, and the eye movements to the first fixation to that AOI was analyzed in detail. Table 3 summarized the metrics of those two tasks. Table 4 includes metrics of eye movements: time to first fixation to the target AOI, number of fixations on the AOI, and time to first mouse click on the AOI.

Table 3

Expected vs. Found Results

Tasks	Average TCD*	Shortest TCD*	Longest TCD*	Expected TCD*	Average MC#	Min MC#	Max MC#	Expected MC#
Task 1	60.88	25	154	16	8	5	11	4
Task 2	52.9	13	126	15	8.5	7	15	6

*TCD: Task Completion Duration; #Mouse Clicks

Table 4

Average Fixations and Mouse Clicks

Tasks	Time to First Fixation	Number of Fixations	Time to First Mouse Click
Task 1	9.81	11.06	8.58
Task 2	4.03	81.10	5.39

The expected ones in Table 3 represent the optimum possible way to follow. For instance, one can complete the Task1 with 4 mouse clicks, but there is a considerable difference between minimum and maximum number of clicks, which may be a sign of the design problem. Similarly, task completion duration for this task takes about 16 seconds, however, the gap between participants' performance was striking. For example, one user completed it in 25 seconds, which is close to the expected one, whereas another user finished in 154 seconds. In addition, the average task completion duration is approximately four times longer than the expected duration. Therefore, this task is a good candidate for deep focus with regards to eye movements. Most of the users explained how they had difficulties finding the detailed search option. Most of the participants failed to complete the task, and therefore gave up. As shown in Table 4, users fixated on the related AOI, but it took about 8.58 seconds to associate it with the target task and to click on it. The following excerpts exemplifies the satisfaction status of users:

"...detailed search cannot be found. I thought that keywords of previous searches would appear when I click the arrow sign." (K4)

"...I think finding detailed search is hard for beginners." (K5)

"It would be better if there is a hint for a detailed search option." (K13)

"I really had difficulty finding the detailed search. The location of it does not make any sense. I wouldn't think that it would be there." (K16)

"...there are no visuals of detailed search, I couldn't have completed the task." (K20)

In the second task, users were expected to search and download the given dissertation. All users found it, but there were gaps between expected and performed values of metrics. The average task completion duration was 52.9 seconds, which is above the expected duration. Moreover, the shortest duration was almost ten times shorter than the longest task completion duration. Besides the large gap among the performances of the users, the number of mouse clicks was above expectations. There were too many fixations (N=81) on the target AOI, whereas it took 5.39 seconds to detect and decide to click. According to users, the link did not seem to be clickable, and therefore they hesitated, which brought about a longer task completion period. The following excerpt are some examples of users' hesitations:

"Downloading the dissertation was very difficult. I found it incidental , I clicked the first clickable place and downloaded." (K1)

"...there was neither a hand sign nor any other sign to guide the downloadable area. I worked hard to find the location of the dissertation to download." (K6)

"I found the dissertation, but the button was hardly visible. One needs to click the dissertation number to download. Incidentally, I clicked hoping that would be the right link." (K19)

The Fastest vs. Slowest Users' Eye Movements for Task 1: Examining the task completion processes of all users showed difficulties in finding the detailed search option. An average user can complete the task in about 16 seconds. In this study, the fastest duration was recorded as 25 seconds (K12). The gaze plot of user K12 was demonstrated on Figure 1. It can be observed that the user detected the related field of the screen after a few fixations over the top left corner of the screen. On the other hand, the longest duration was recorded as 154 seconds (K1). Since user K1 did not detect the menu part for the detailed search, s/he first did a regular search, and then tried to access the detailed search menu on the results page, which brought about a longer task duration. Figure 2 demonstrates the heat-map of K1.

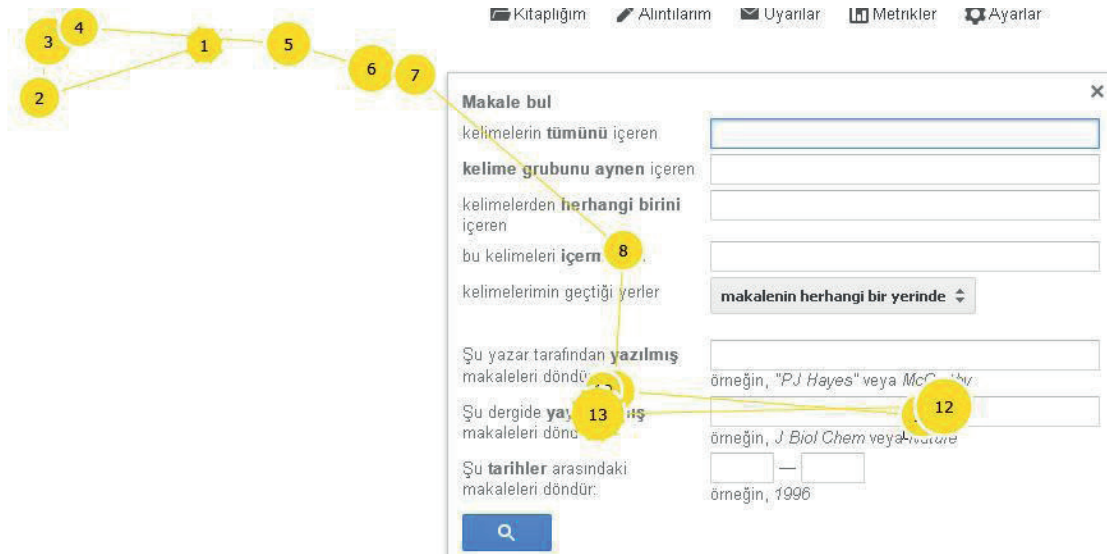


Figure 1. The gaze plot of K12 during Task 1

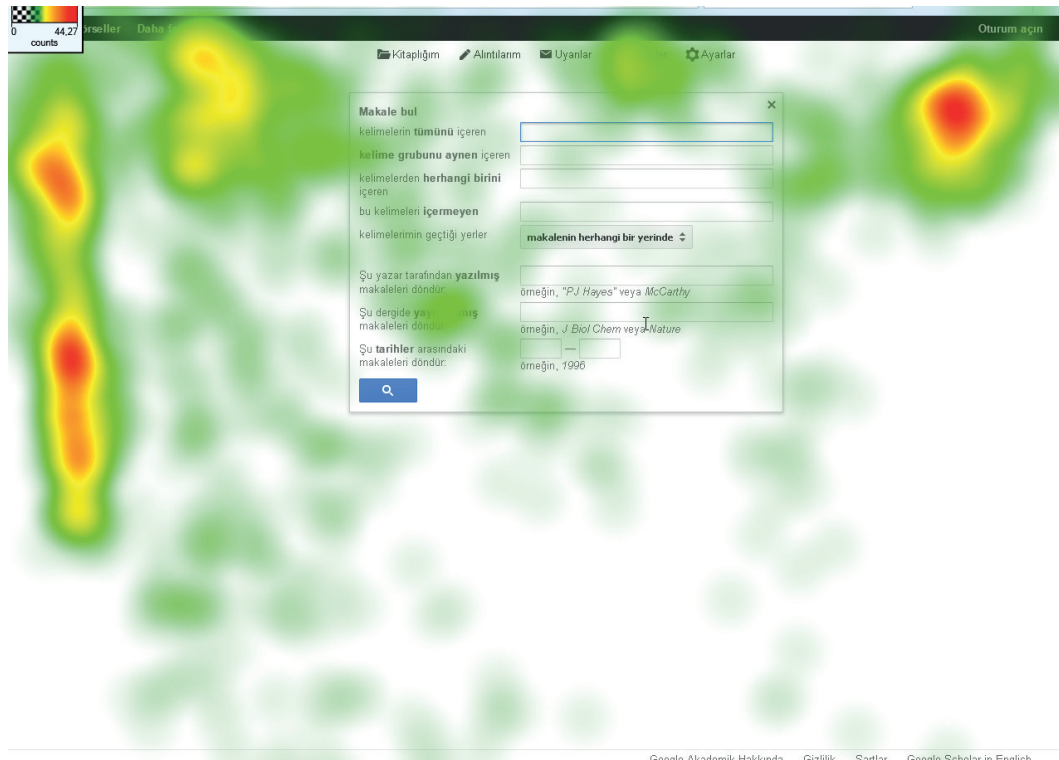


Figure 2. Gaze plot of K1 during Task 1

User K1 spent a considerable amount of time on the left side of the page, then skimmed all pages around, and then due to the lack of hints, s/he focused on the article links on the results page for a detailed search tool. Finally, despite fixations on the target area, s/he did not locate/click the item. After spending too much time, s/he went back to the main search page and found the detailed search tab. In short, realizing the existence of the option took about 103 seconds, which is quite long. As pop-up appeared, user K1's examining the window, writing keywords, and searching processes took 51 seconds, which is still longer than the shortest task completion duration. Figure 3 shows the first 7 seconds of user K1, after pop-up windows.

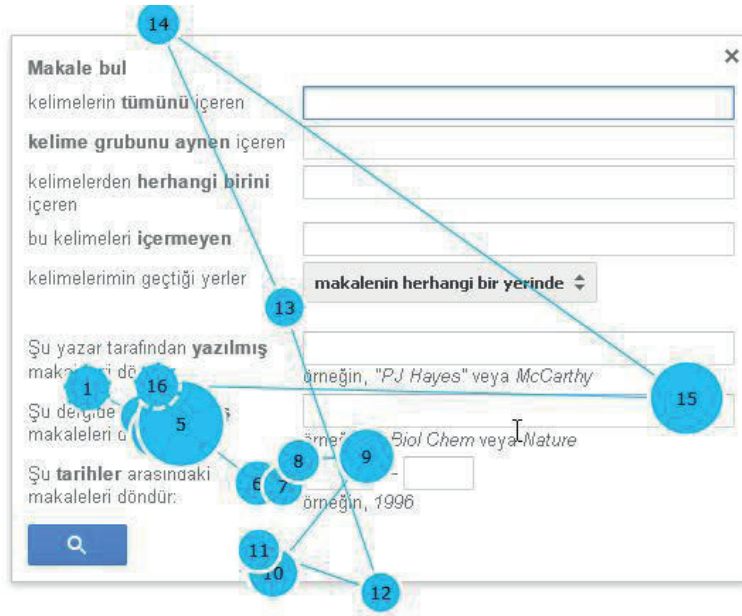


Figure 3. The first 7 seconds of K1 on Task 1

Failures of Task 1: Users K13, K16, and K20 were not able to finish Task 1. The eye movements of those users show intense fixations on the menus on both sides of the page. User K13 spent too much time focusing on the interval option of the menu, but as soon as realizing the date specific facet of that part, s/he shifted the focus. Like user K13, user K16 tried to find the detailed search on the both sides of the page . Despite shorter fixations, K16 spent too much time locating the correct menu item. Although s/he saw the item on the settings button at the upper right corner, s/he did not associate it with the task, and thus failed to complete. The longest fixations of that user were also on the target area. User K20 showed short fixations all around the page, but could not have realized the target, and as a result of that s/he gave up the task. Figure 4 consists of the eye movements of users who failed during Task 1.

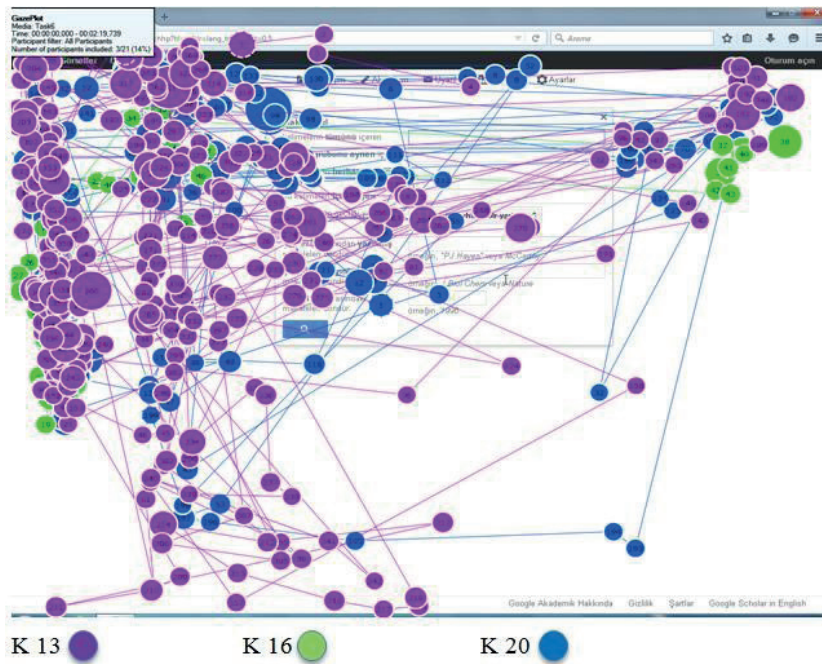


Figure 4. Eye movements of users who gave up the Task 1

The Fastest vs. Slowest Users' Eye Movements for Task 2: An average user can complete the task at about 15 seconds, but one of the participants (K9) was faster with 13 seconds. The slower performers generally got stuck on finding the download link. User K9 first entered the keywords on the main page with necessary options. Once realizing the dissertation number, s/he clicked and the downloading process started. That user had slight fixations on general information about the thesis, and never got distracted by other items on the page. Figure 5 shows smooth eye movements of the user.

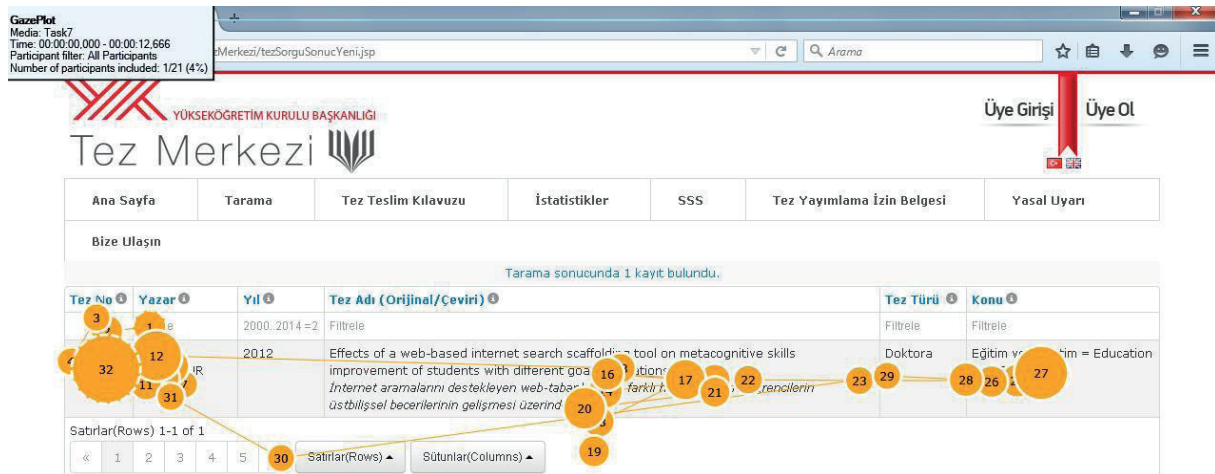


Figure 5. Gaze plot of user K9 during Task 2

User K11 completed Task 2 within 109 seconds. As shown on Figure 6, s/he gazed around all over the page, fixated and clicked each item one by one. Despite frequent fixations on that area, s/he did not click on the dissertation number, which was the target area for downloading. It can be observed that the duration of fixations on that area increased towards the end of the task. As the user was sure of the functions of other items, s/he decided to click on the correct link. Figure 7 shows the first 6 seconds of the user. According to eye movements on that figure, s/he did not fixate on the correct place for 5 seconds. After 6 seconds, s/he fixated on the right item, but did not click.

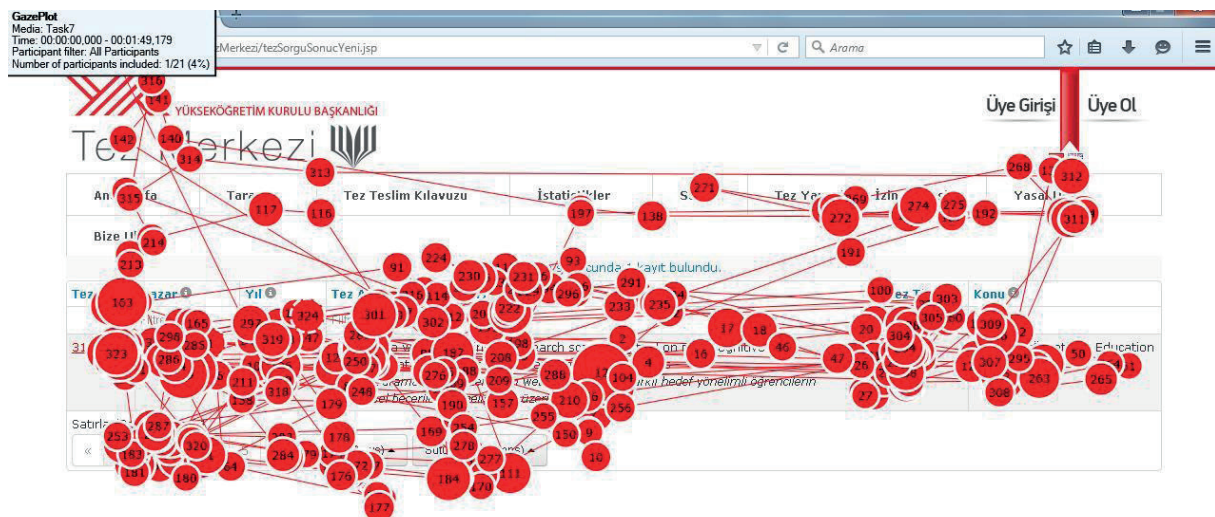


Figure 6. Gaze plot of user K11 during Task 2



Figure 7. The first 6 seconds of K11 on Task 2

Failure of Task 2: There was only one user who failed to complete the task (K1). The search patterns of K1 were very similar to K11, the slowest user in Task 2. User K1 gazed all around the page with similar fixation durations, and fixated on the dissertation number, but did not click due to lack of associations with the download link. As a result, s/he gave up the task. Figure 8 demonstrates the eye movements of the user during Task 2.

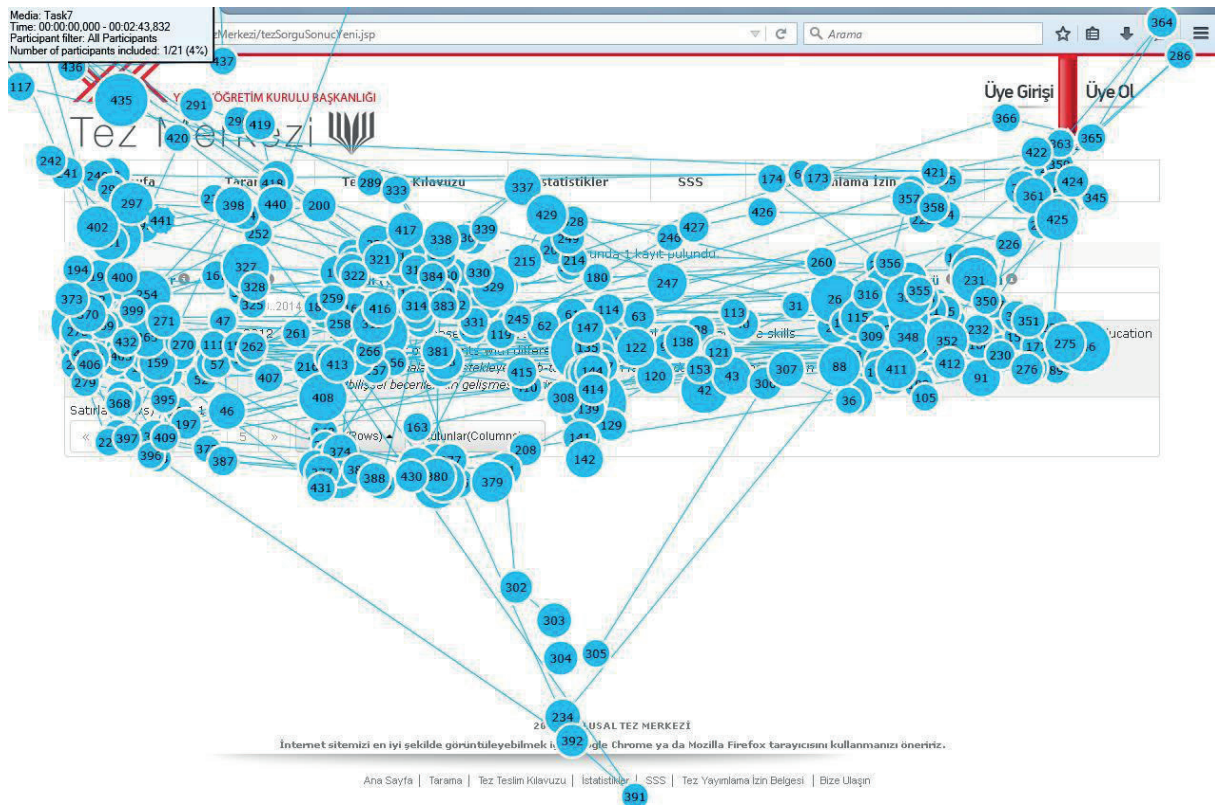


Figure 8. Gaze plot of K1 during Task 2

3. DISCUSSION AND CONCLUSION

It was observed that users' gaze plots, efforts to complete the tasks, and task completion durations/status were very different from each other during the two tasks, which caused problematic user experience. In both tasks, users generally focused on both sides of the pages, where menu items are located. Although the majority of the users showed fixations at the correct places, they kept searching for the target at different parts of the screen due to lack of associations. This situation led to

increases in fixation counts, fixation durations, mouse clicks, and task completion durations. According to Rayner (1998), long fixation durations are associated with intense cognitive effort along with longer thinking durations. This assumption can explain the relation between long fixation durations and task completion duration. In addition, the failure of completion may also be related to cognitive overload as a result of longer fixations everywhere accompanied by mouse clicks. Situations having too many options can cause confusion, which can be a sign of cognitive overload (Murray 2001). For example, the study conducted by Kilic and Karadeniz (2004) indicated that students, whose cognitive load is high, showed lower performance and longer task completion durations in hypermedia. The current study showed that when the user fails to associate the target design element, the number of fixations and mouse clicks increase, and thus the gaze plots exemplify the cognitive overload resulting in either failure or longer task completion durations.

In the first task, users were assigned the task of making a detailed search on Scholar Google. Users generally tended to focus on menu items, but failed to locate the targeted one. Providing users with flexible interfaces including freedom of choices is an important part for usable designs (Nielsen 1994). Although the experienced users accessed the detailed search quickly, the users who rarely used the details option had difficulties learning where to click. In the second task, the main problem was the lack of cues about the existence of a download link, which can be overcome easily by defining the cursor hover style as a pointer. According to Nielsen's (1994) heuristics, consistency and standards should be taken into account for universal designs. In this way, the designer can prevent misconceptions in the interface. In conclusion, the problems in the design of two frequently used online resources resulted in negative user experiences. The users spent extra time and effort due to lack of salient hints. The signals and cues are valuable in increasing the efficiency in learning (Clark et al. 2006).

Finally, the study has limitations in terms of certain issues, and therefore one should be careful when interpreting the results of this study. First, the experiment is limited to assigned tasks. Because of the location of the human-computer interaction lab, we only included the volunteers who were available around us in the study. In the beginning protocol, we encouraged users to behave as if they were at home/office, but being in a controlled environment may have affected the overall performance of the users.

The research implications can be summarized as follows:

- Since we utilized convenient sampling, the further studies can enlarge the number and variety of participants.
- One can provide alternative designs and compare them with the old ones.
- One can compare the resources based on the participants' age or gender.
- The practical implications can be summarized as follows:
 - Before eye tracking experiments, the researcher should make sure the physical and technical settings are all in good condition.
 - Having a plan-B is needed when planning data collection period, because the logistic issues may sometimes restrict the number of accessible users. Scheduling according to users, and having extra users in advance may prevent possible data collection problems.
 - Large data size can bring about freezing screens or slow performance. In order to have optimum performance, the software enabling the analysis of eye tracker data should be installed on a high performance computer.
 - In user tests, data collection with eye trackers is a considerably lengthy process, that's why time management and planning is necessary to increase efficiency.

Hakem Değerlendirmesi: Dış bağımsız.

Çıkar Çatışması: Yazarlar çıkar çatışması bildirmemiştir.

Finansal Destek: Bu çalışma, Ondokuz Mayıs Üniversitesi Bilimsel Araştırmalar Proje Birimi (BAP) tarafından desteklenmiştir. (Proje No: PYO.EGF.1904.15.005)

Peer-review: Externally peer-reviewed.

Conflict of Interest: The authors have no conflict of interest to declare.

Grant Support: This research was supported by Ondokuz Mayıs University BAP. (Project number: PYO.EGF.1904.15.005)

Kaynaklar/References

- Agrawal, R., & Nyamful, C. (2016). Challenges of big data storage and management. *Global Journal of Information Technology*, 6(1), 1-10. <https://doi.org/10.18844/gjit.v6i1.383>
- Alemdag, E., & Çağiltay, K. (2018). A systematic review of eye tracking research on multimedia learning. *Computers & Education*, 125, 413-428. <http://doi.org/10.1016/j.compedu.2018.06.023>
- Almeida, S., Veloso, A., Roque, L., & Mealha, O. (2011, January). The Eyes And Games: A Survey Of Visual Attention. In *The Sbgames Brazilian Symposium On Computer Games And Digital Entertainment*. Salvador, Brazil, .
- Bahsisoglu, H. K. (2006). E-kaynaklar kullanımının akademik çalışmalara sağladığı kolaylıklar. Akademik Bilişim Bilgi Tek IV 9-11. <http://ab.org.tr/ab06/bildiri/133.doc>. Accessed 05 May 2019
- Beymer, D., Orton, P. Z., & Russell, D. M. (2007, September). An eye tracking study of how pictures influence online reading. In *IFIP Conference on Human Computer Interaction Springer Berlin Heidelberg 2007* (pp. 456-460). Springer, Berlin, Heidelberg.
- Brandt, D. S. (1996). Evaluating information on the internet. *Computers in Libraries*, 16(5), 44-46.
- Buscher, G., Dumais, S. T., & Cutrell, E. (2010, July). The good, the bad, and the random: an eye-tracking study of ad quality in web search. In *33rd international ACM SIGIR conference on Research and development in information retrieval 2010* (pp. 42-49). <https://doi.org/10.1145/1835449.1835459>
- Byerly, G. (2007). Look in their eyes-eye tracking, usability, and children. *School Library Media Activities Monthly*, 23(8), 30-32.
- Clark, R. C., Nguyen, F., Sweller, J., & Baddeley, M. (2006). Efficiency in learning: Evidence-based guidelines to manage cognitive load. *Performance Improvement*, 45(9), 46-47. <http://doi.org/10.1002/pfi.4930450920>
- Cutrell, E., & Guan, Z. (2007) What are you looking for?: an eye-tracking study of information usage in web search. In: *SIGCHI conference on Human factors in computing systems* (pp. 407-416). <https://doi.org/10.1145/1240624.1240690>
- Çağiltay, K. (2011). *İnsan bilgisayar etkileşimi ve kullanılabilirlik mühendisliği: Teoriden pratiğe*. ODTÜ Yayıncılık, Ankara
- Cetin, I., & Sendurur, E. (2016). Çevrimiçi akademik kaynakların kullanılabilirlik değerlendirmesi. *Mehmet Akif Ersoy Üniversitesi Eğitim Fakültesi Dergisi*, 1(40), 273-299. <http://doi.org/10.21764/efd.04062>
- Dumas, J. S., & Redish, J. (1999). *A practical guide to usability testing*. Intellect Books.
- Hertzum, M., Hansen, K. D., & Andersen, H. H. (2009). Scrutinising usability evaluation: does thinking aloud affect behaviour and mental workload?. *Behaviour & Information Technology*, 28(2), 165-181. <https://doi.org/10.1080/01449290701773842>
- ISO 9241-11 (2018). Ergonomic requirements for office work with visual display terminals (vdt) part 11: Guidance on usability. International Organization for Standardization, Geneva. <https://www.iso.org/obp/ui/#iso:std:iso:9241:-11:ed-2:v1:en>. Accessed 15 March 2019
- ISO/IEC 25010 (2011). Systems and software engineering – systems and software quality requirements and evaluation – system and software quality models. International Organization for Standardization, Geneva. <https://www.iso.org/obp/ui/#iso:std:iso-iec:25010:ed-1:v1:en>. Accessed 15 March 2019
- Kazan, H. (2017). Avantajları ve dezavantajları ile dijital dergicilik. *The Turkish Online Journal of Design, Art and Communication*, 7(1), 17-24.
- Kilic, E., & Karadeniz, S. (2004). Hiper ortamlarda öğrencilerin bilişsel yüklenme ve kaybolma düzeylerinin belirlenmesi. *Kuram ve Uygulamada Eğitim Yönetimi Dergisi*, 10(4), 562-579.
- Kovacs, D. K., & Elkordy, A. (2000). Collection development in cyberspace: building an electronic library collection. *Library Hi Tech*, 18(4), 335-361. <https://doi.org/10.1108/07378830010360446>
- Land, S. M. (2000). Cognitive requirements for learning with open-ended learning environments. *Educational Technology Research and Development*, 48(3), 61-78. <https://doi.org/10.1007/BF02319858>
- Lewandowski, D., Kerkmann, F., Rümmele, S., & Sünkler, S. (2018). An empirical investigation on search engine ad disclosure. *Journal of the Association for Information Science and Technology*, 69(3), 420-437. <http://doi.org/10.1002/asi.23963>
- Lindgaard, G., & Chattratichart, J. (2007). Usability testing: what have we overlooked?. In *Proceedings of the SIGCHI Conference on Human factors in Computing Systems* (pp. 1415-1424). ACM. <https://doi.org/10.1145/1240624.1240839>
- Lund, H. (2016). Eye tracking in library and information science: A literature review. *Library Hi Tech*, 34(4), 585–614. <http://doi.org/10.1108/LHT-07-2016-0085>
- Murray, T. (2001, October). Characteristics and Affordances of Adaptive Hyperbooks. In *WebNet* (pp. 899-904).
- Myers, B. A., & Rosson, M. B. (1992, May). Survey on user interface programming. In *Proceedings of the SIGCHI conference on Human factors in computing systems* (pp. 195-202). ACM.

- Nakatani, C., & Pollatsek, A. (2004). An eye movement analysis of “mental rotation” of simple scenes. *Perception & psychophysics*, 66(7), 1227-1245. <https://doi.org/10.3758/BF03196848>
- Nielsen, J. (1994). Heuristic evaluation. In Nielsen, J., and Mack, R.L. (Eds.), *Usability Inspection Methods*, (413-414). John Wiley & Sons, New York, NY.
- Nielsen, J. (2012). Usability 101: Introduction to usability. <https://www.nngroup.com/articles/usability-101-introduction-to-usability/>. Accessed 21 January 2019
- Nielsen, J., & Landauer, T. K. (1993). A mathematical model of the finding of usability problems. In *Proceedings of the INTERACT'93 and CHI'93 conference on Human factors in computing systems* (pp. 206-213). ACM.
- Norman, D. A. (1988). *Designing everyday things*. Currency-Doubleday, New York.
- Pan, B., Hembrooke, H. A., Gay, G. K., Granka, L. A., Feusner, M. K., & Newman, J. K. (2004, March). The determinants of web page viewing behavior: an eye tracking study. In *Proceedings of the 2004 symposium on Eye tracking research & applications* (pp. 147-154). ACM.
- Pernice, K., & Nielsen, J. (2009). *Eyetracking methodology: How to conduct and evaluate usability studies using eyetracking*. Berkeley: New Riders Press.
- Rayner, K. (1998). Eye movements in reading and information processing: 20 years of research. *Psychological bulletin*, 124(3), 371-422. <https://doi.org/10.1037/0033-2909.124.3.372>
- Renshaw, J. A., & Webb, N. (2007). Eye tracking in practice. In *21st British HCI Group Annual Conference University of Lancaster, UK* (pp. 239-241). <http://doi.org/10.14236/ewic/HCI2007.41>
- Scheiter, K., & Van Gog, T. (2009). Using eye tracking in applied research to study and stimulate the processing of information from multi-representational sources. *Applied Cognitive Psychology*, 23(9), 1209-1214. <http://doi.org/10.1002/acp.1524>
- Shackel, B. (2009). Usability-Context, framework, definition, design and evaluation. *Interacting with Computers*, 21(5-6), 339-346.
- Tashakkori, A., & Teddlie, C. (1998). *Mixed methodology: Combining qualitative and quantitative approaches*. Sage.
- Tobii Studio (2008) Tobii Studio 1.2 User Manual. <https://manualzz.com/doc/4199338/tobii-studio-1.2-user-manual>. Accessed 16 September 2018
- Virzi, R. A. (1992). Refining the test phase of usability evaluation: How many subjects is enough?. *Human Factors: The Journal of the Human Factors and Ergonomics Society*, 34(4), 457-468. <https://doi.org/10.1177/001872089203400407>
- White, R. W. (2016). *Interactions with Search Systems*. New York: Cambridge University Press.
- Yavuz, I., Cinar, N. O., & Cagiltay, K. (2016). Kamu internet sitelerinde yer alan arama alanlarının kullanılabilirliği ve buna yönelik kullanıcı davranışlarının belirlenmesi. *International Journal of Informatics Technologies*, 9(1), 41. <http://doi.org/10.17671/btd.81715>

Adli Bilişim İncelemelerinde Şifre Kırma Yöntem ve Teknikleri

Password Cracking Methods and Techniques in Computer Forensic Investigation

İlker Kara¹ 



ÖZ

Bilgi sistemleri ve veri kullanımındaki sonsuz artış, bilgi güvenliğinde tehlikenin doğuşunu tetikledi. Son yayınlanan raporlara göre askeri kuvvetler ve e-ticaret web siteleri dışında sıradan kullanıcılarda sistemleri ve belgelerini korumak için şifreleme teknikleri kullanmaya başlanmışlardır. Alınana tedbirlere rağmen çeşitli gizleme tekniklerini kullanarak hazırlanan akıllı saldırılar mevcut korunma yöntemlerini atlatarak hedef sistemdeki parola ve kullanıcı adlarını ele geçirebilmektedir. Kurumsal firmalar ve sıradan kullanıcılar verilerini gizlemek için yeni nesil şifreleme yöntemlerini yaygın olarak kullanmaktadır. Bu durum özellikle adli olaylara konu olan bilgi sistemleri ve dosyaların incelenmesinde büyük engeller oluşturmaktadır. Eğer şüpheli kişi kullanmış olduğu bilgi sistemi veya dosyalarını şifrelenmiş ise delil elde etmek için önce bu şifrelerin önceden bilinmesi ya da şifrenin kırılması gereklidir. Bu adımda şüpheli kendi rızasıyla parola kolluk kuvvetlerine vermemesi durumunda adli uzmanlar çeşitli yöntemlerle şifreleri kırmaya çalışmakta bu süreç genellikle zor olmakta ve bazı durumlarda şüpheli sistemdeki şifreli verilere ulaşılamamaktadır. Bu çalışma iki katkı sunmaktadır. İlk olarak en çok kullanılan şifre kırma yöntemleri detaylı olarak incelenmiştir. İkincisi, “BitLocker” veri şifreleme yöntemiyle şifrelenmiş örnek bir adli vaka incelenerek şifreli verileri kırılma adımları incelenmiştir. Sonuçlardan şifrelenmiş verilerin erişmek için kullanılan yöntemin etkili olduğunu ve şifrelerin kırıldığı göstermektedir.

Anahtar kelimeler: Şifre Kırma, Veri Şifreleme, Güvenlik Saldırıları, Analiz Yöntemleri

ABSTRACT

The unending increase in information systems and data use has triggered the birth of danger to information security. According to recently published reports, apart from military forces and e-commerce websites, ordinary users have begun to use encryption techniques to protect systems and documents. In spite of precautions, smart attacks prepared using a variety of concealing techniques overcome available protection methods and can obtain the passwords and user names of on the target system. Corporate firms and ordinary users commonly use new-generation encryption methods to hide their data. This situation creates large obstacles for the investigation of computer systems and files which are the subject of forensic events, especially. If a suspect uses a computer system with encrypted files, to obtain evidence, firstly, it is necessary to know these encryptions or to crack them. In this step, if the suspect does not give law enforcement the encryptions willingly, forensic experts attempt to break the encryption with a variety of methods. This process is generally difficult, and in some situations, the encrypted data on the suspect's system cannot be reached. This study provides two contributions. The first is that a detailed investigation of the most commonly used encryption cracking methods are investigated in detail. Secondly, an example forensic case encrypted with the “BitLocker” data encryption method is investigated and the steps to break the encrypted data are investigated. The results show that the methods used to access the encrypted data is effective and that the encryption was cracked.

Keywords: Password Cracking, Data Encryption, Security Attacks, Analysis Methods

¹(Dr. Öğr. Üyesi), Çankırı Karatekin Üniversitesi
Eldivan Sağlık Hizmetleri Meslek Yüksek Okulu,
Çankırı, Türkiye

ORCID: İ.K. 0000-0003-3700-4825

Corresponding author:

İlker KARA
Çankırı Karatekin Üniversitesi Eldivan Sağlık
Hizmetleri Meslek Yüksek Okulu, Çankırı,
Türkiye
E-mail address: karaikab@gmail.com

Submitted: 02.10.2021

Revision Requested: 19.02.2021

Last Revision Received: 22.02.2021

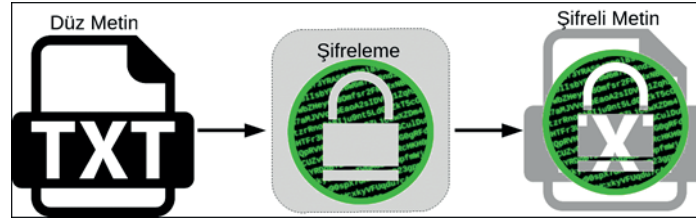
Accepted: 22.02.2021

Published Online: 31.05.2021

Citation: Kara, İ. (2021). Adli bilişim
incelemelerinde şifre kırma yöntem ve teknikleri.
Acta Infologica, 5(1), 27-38.
<https://doi.org/10.26650/acin.804201>

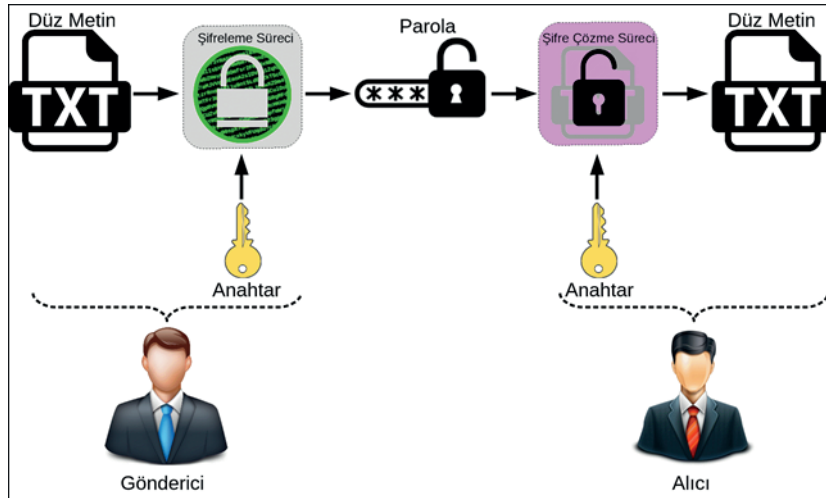
1. GİRİŞ

İlk ortaya çıktığından beri, veri şifreleme yöntemleri bilgi sistemlerini ve verileri en iyi koruma tekniklerinden biri olarak bilinmektedir (Guddeti vd. 2020). Bu güvenlik yaklaşımı, kullanıcının korunan sistemlerin ve verilerin içeriğini erişimi için bir şifre çözme anahtarı kullanılması olarak tanımlanır. Bu şekilde, yalnızca yetkili kişiler (anahtara erişimi olanlar) korunan sisteme veya verilere erişebilmektedir (Bhanot vd. 2015). Bir açık metnin bir şifreleme algoritması vasıtasıyla anlaşılamaz bir metin haline getirilmesi işlemine şifreleme (Encryption) denir (Şekil 1) (Dass vd. 2020). Şifre Çözme (Decryption) işlemi ise bir şifreleme algoritması şifrelenmiş veriyi çözerek ilk haline getirme işlemidir (Şekil 2) (Hur vd. 2019).



Şekil 1. Şifreleme adımları.

Veri şifreleme yöntemi, ticari işletmeler, resmi kurumlar, silahlı kuvvetler, kişisel kullanıcılar ve e-ticaret web sitelerinde ödeme ayrıntılarını korumak da dâhil olmak üzere bir dizi farklı alanda kullanılmaktadır (Saraçević vd. 2020). Günümüz teknoloji uygulamalarda veri şifrelemenin doğru bir şekilde yapılmasıyla, verilerin hemen hemen her tehdiye karşı koruyabilecek güce sahiptir. Bununla birlikte, birçok kurum ve kullanıcı (son derece hassas verilere sahip olanlar bile) henüz bu yöntemi uygulamamaktadır (Kara 2019). Bu verilerin varlığı şüphesiz siber suçlar için ilgi çekici bir hedeftir ve korunmasız olanlarda ilk saldırıya uğrayanlar arasındadır.



Şekil 2. Şifre oluşturma-çözme süreci.

Kurumlar ve kullanıcılar veri şifreleme yönteminin sunabileceği güvenlik avantajlarının farkında olsa da, şifrelenmiş veriler adli incelemelerde birtakım zorluklarla karşılaşılmasına neden olmaktadır. Adli bilişim uzmanı, bilgi sisteminde ki veriye ulaşabilmesi için bu şifreleme mekanizmalarını aşması gereklidir. Çoğu zaman yapılan adli incelemelerdeki şüpheli bilgi sistemlerinin şifrelemeyi aşmaları çok zor hatta imkânsız olabilmektedir. Günümüzde suç işleyenler veya şüpheli durumda olan insanlar kendilerini veya kişisel verilerini gizlemek amacıyla veri şifreleme yöntemlerini kullanarak işlerini kolaylaştırırken adli uzmanların işlerini zorlaştırmaktadır.

Öte yandan gelişen teknolojinin getirdiği avantajlardan faydalanan şifreleme yöntemleri de gün geçtikçe gelişmektedir. Apple iOS 11'de, Firefox arama motorunda, Whatsapp mesajlaşma uygulaması gibi birçok uygulaması güvenlik önlemlerini artırdı ve "end-to-end" olarak bilinen şifreleme kullanılmaya başlamışlardır. Microsoft Windows 8 sürümünde standart olarak

sunulan hard disk şifrelemesi kolaylıkla kullanıcılar tarafından uygulanabilmektedir. Sıradan kullanıcılar dahi iletişimlerini ve kişisel verilerini gizlemek için PGP, TrueCrypt, Microsoft BitLocker ve Tor gibi şifreleme uygulamalarını kullanmaktadır. Hukuka uygun olarak kişisel verilerin korunması ve gizlilik politikalar kapsamında kişisel veri sahibinin açık rızasının bulunmadan ikinci kişilerin erişmemesi kanunlar ile koruma altına alınmış olmakla birlikte özgürlüklerin ve kamu güvenlik arasındaki bu dengeyi kanunlar ile kullanıcılar arasında kurmak gereklidir.

Bu çalışmada veri şifreleme yöntemleri ve şifre kırma analizlerindeki yeni gelişmeleri anlatılmaktadır. Çalışmanın odak noktası da “BitLocker” veri şifreleme yöntemiyle şifrelenmiş gerçek bir şifreleme vakasının incelenmesi olacaktır. Bu amaçla “BitLocker” veri şifreleme yöntemiyle şifrelenmiş bilgisayardaki verilerin nasıl kurtulabileceğini gösteren uygulanabilir bir yaklaşım önerildi. Önerilen yaklaşımda şifre kırma analizlerinde Windows tabanlı bilgisayarlarda brute force şifre kırıcısı olan “Passware Kit Forensic Version 2019.2.0 (Ücretsiz sürümü)” yazılımı kullanıldı. Bu yaklaşımla benzer yöntemlerle şifrelenmiş bilgi sistemleri ve veriler için uygulanabilir olmasını hedeflenmiştir.

2. ŞİFRE KIRMA (PASSWORD CRACKING) YÖNTEMLERİ

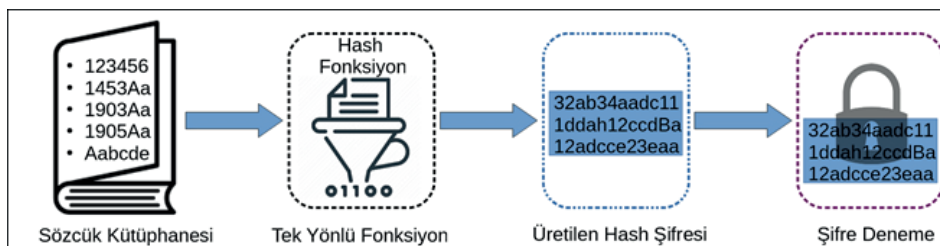
Şifre ve parolalar dijital dünyada siber saldırganların odağında olmakta ve kırılan tek bir şifre ile tüm güvenlik sistemleri atlatılarak tüm sistem veya veriler ele geçirilebilmektedir. Şifreli verilere veya bilgi sistemlerine ulaşmak saldırganlar kadar sıradan insanların ilgisini çekmekte bu durum şifre kırmak için her geçen gün yeni yöntemler geliştirmeye çalışılmaktadır. Bu nedenle bu çalışmada şifre kırmak için kullanılan en popüler yöntemlere değinilmiştir. Şifre kırma yöntemlerini “aktif” ve “pasif” yöntemler olarak iki ana gruba ayrılabilir. Bunlar (Kaya vd. 2017):

- Aktif şifre kırma yöntemlerinde; hedef bilgi sistemine veya veri dosyalarına çevrimiçi (anlık) olarak şifre kırma saldırıları gerçekleştirilmektedir.
- Pasif şifre kırma yöntemlerinde; ise çevrimdışı ortamlarda önceden hazırlanan tahmini şifreler hazırlanarak, incelemeler sonucunda ele geçirilen şifreler denenerek veya olası şifre tahminleriyle yapılan şifre kırma saldırılarıyla gerçekleştirilmektedir.

Aktif şifre kırma yöntemi kullanılabilmesi için bir şekilde hedef veri tabanı yâda kullanıcı bilgilerinin (olası hesap kullanıcı adı veya parolalar gibi) bilinmesi gereklidir. Bu durum dışında kalan senaryolarda pasif şifre kırma yöntemleri kullanılmaktadır.

2.1. Sözlük Saldırısı (Dictionary Attack)

Sözcük saldırısı “sözlük dosyasında” oluşturulan olası her kelimeyi otomatik şifre kombinasyonları halinde şifrelenmiş sistem veya veri dosyasına karşı deneyerek şifreyi kırma denemelerine denilmektedir (Bhanot vd. 2015). Sözcük dosyası içerisinde kullanıcıların daha önce en sık kullandıkları şifreleri içeren bir dosyadan oluşmaktadır. Sözcük saldırısı üretilen tek yönlü fonksiyon (Hash fonksiyonu) elde edilmiş şifreler hedef sistemdeki parolayı veri tabanında karşılık gelen olasılıkları tahmin etmesiyle şifre kırılmaktadır (Beşkirli vd. 2019) (Şekil 3).

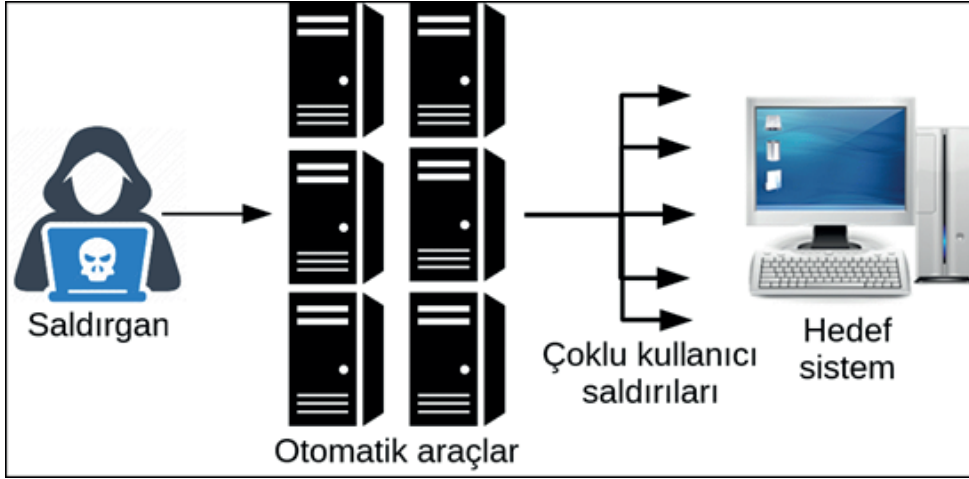


Şekil 3. Sözcük saldırı algoritması.

Sözcük saldırısında üretilen olası hash şifreler ile hedef sistemdeki şifreyi kırmak mevcut parolanın büyüklüğüne ve şifreleme tekniğine (Gelişmiş şifreleme algoritması-AES, Asimetrik şifreleme algoritması-RSA gibi) göre işlem süresi değişiklik göstermektedir (Beşkirli vd. 2019). Bazen hedef sistemdeki algoritmanın çözülme olasılık süresi yıllar alabilmektedir (Bhanot vd. 2015).

2.2. Kaba Kuvvet Saldırısı (Brute Force Attack)

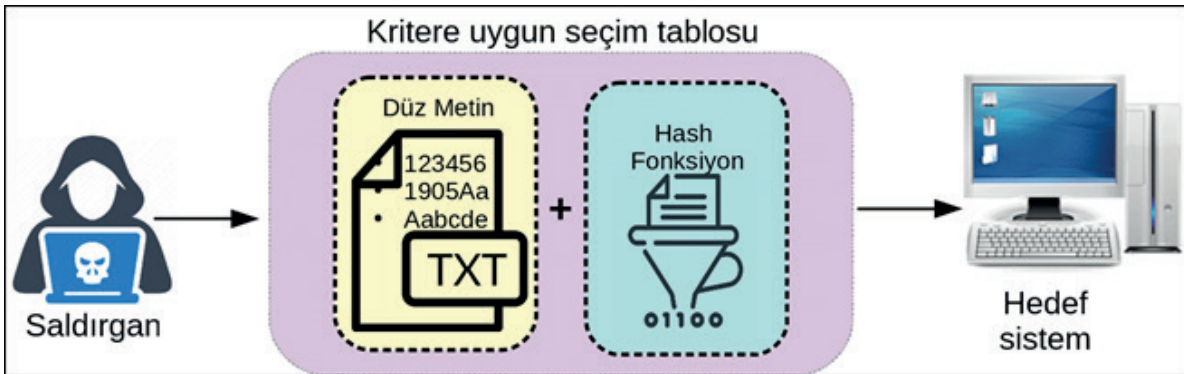
Kaba Kuvvet saldırıları yöntemi hedef sistemdeki parolayı deneme saldırılarıyla kıırma mantığıyla çalışmaktadır. Bu amaçla oluşturulan harfler ve özel karakterler içeren bir liste (Pass List) hazırlanır (Kara 2019). Kaba Kuvvet saldırısının başarı süresi hedef sistemdeki parolanın karmaşıklığına ve saldırıda kullanılan sistemimin donanım gücüne göre değişiklik göstermektedir (Şekil 4).



Şekil 4. Kaba kuvvet saldırı algoritması.

2.3. Gökkuşığı Tablosu Saldırısı (Rainbow Table Attack)

Gökkuşığı tablosu saldırısı yönteminde hedef sistemdeki parolaları kıırma için belirli ölçüte göre seçilen içerisinde düz ve hash fonksiyonlarını içeren bir şifre tablosu hedef sistemdeki parolayı kıırma mantığıyla çalışmaktadır (Bhanot vd. 2015; Beşkirli vd. 2019). Kaba kuvvet saldırılarında olduğu gibi hazırlanan şifre tablosu tek tek hedef sistemdeki parolayı kıırma için kullanılmaktadır. Kaba kuvvet saldırısından ayıran en önemli farkı ise tüm parolaları denemeye dayalı olmasıdır. Dahası denenilen parolaları belirli bir düzene dayalı olarak seçmektedir. Bu düzen ise düz karşılığı olan hash edilmiş parolaları bilgisayar sisteminde sürekli olarak denemekten geçmektedir (Beşkirli vd. 2019).



Şekil 5. Gökkuşığı tablosu saldırı algoritması.

Gökkuşığı tablosu saldırısı yönteminde tablo hazırlanırken açık kaynak araması ile kullanıcıların en çok tercih ettiği parola örnekleri ve özel servilerin (rockyou wordlist gibi) hazırladığı listeler kullanılmaktadır (Bhanot vd. 2015).

2.4. Otalama (Phishing) Saldırı Yöntem

Otalama saldırıları, kurbanı ait şifrelerini ele geçirmek üzere yapılan saldırılardır. Otalama saldırıları, sahte bir e-posta (İçeriğinde; indirim, bedava ürün ya da hediye çekışı kazandığını belirten) ile kurbanı ulaşarak ilk bakışta zararsız gibi

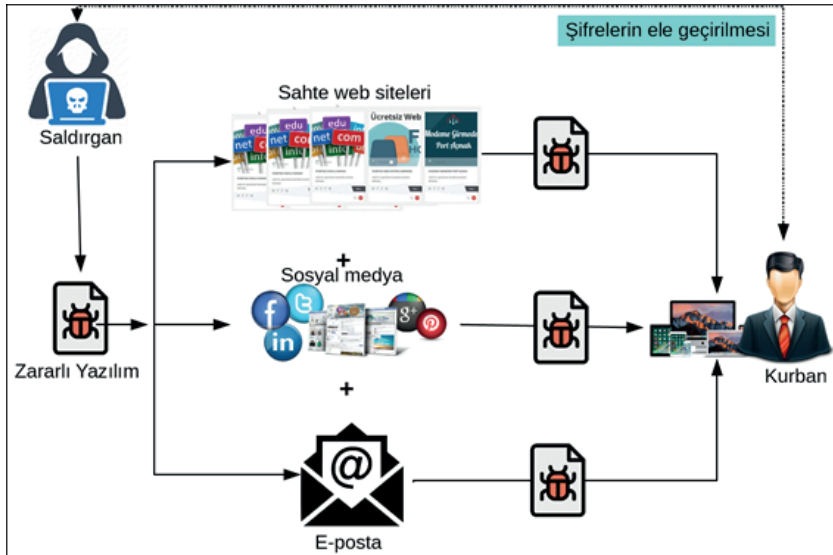
görülen ekinde hedef kişiyi sahta web sitelerine yönlendiren saldırgan kurban kişiden kullanıcı adı ve şifresini sisteme girmesi üzerine kurulmuştur (Şekil 6) (Kara 2019).



Şekil 6. Otalama saldırı algoritması.

2.5. Kötücül Yazılım (Malware) Saldırıları

Saldırganlar, şifre kırma yöntemlerinde en çok kötücül yazılım saldırıları tercih etmektedirler. Hedef bilgi sistemindeki parolayı kırmak ya da ele geçirmek için önceden hazırlanmış kötücül yazılımları (Virüs, tuş kaydedici (keylogger), truva atı, fidye yazılım gibi) kullanmaktadır (Kara 2019). Saldırgan bu tür kötücül yazılımları sahte web siteleri, önceden tahrip edilmiş sosyal medya platformları, haber siteleri, ilan siteleri, sahte e-postalar ile hedef sisteme ulaşmaktadır. Hedef sisteme sızan kötücül yazılım sistemde kayıtlı olan tüm kullanıcı adı ve parolayı ele geçirerek saldırganına bir e-posta ile göndermektedir. Böylece saldırgan amacına ulaşmış olup kurbandan ait kredi kartı bilgileri, sosyal medya hesaplarını, kimlik verileri, e-posta hesap şifrelerini kolayca ele geçirmektedir.



Şekil 7. Kötücül yazılım algoritması.

3. ŞİFRE KIRMA PROGRAMLARI

Günümüzde farklı stratejilerle çalışan birçok şifre kırma yazılımları vardır. Bunlardan en popüler olanları; Passware Kit Forensic Version, Aircrack-ng, Hydra, Elcomsoft Distributed Password Recovery, Hashcat Password Recovery, John the Ripper, fgdump, Medusa, Cain and Abel ve RainbowCrack'dır. Bu bölümde kısaca bu yazılımlar hakkında bilgi verilecektir.

- a) Passware Kit Forensic (PKF): Özellikle bilgisayarlardaki (Windows, Linux ve Amazon EC2(Amazon Web Services)) şifrelerin kırma için ve adli uzmanlarında adli analizlerde tercih ettikleri etkili bir şifre kırma yazılımıdır. PKF bilgisayardaki tüm alanları veya seçilen özel dosyalar üzerindeki tüm şifre korumalı öğeleri kaba kuvvet saldırı tekniğiyle eksiksiz çözebilme yeteneğine sahiptir. Yazılım 280'den fazla dosya türünü tanır ve şifreleri kırma için toplu modda (aynı anda birkaç şifreli alan veya dosya üzerinde) çalışabilmektedir.
- b) Aircrack-ng: Aircrack-ng yazılımı kablosuz iletişim WI-FI sahip olduğu WEP ve WPA-PSK (kablosuz ağları için bir güvenlik algoritmaları) güvenlik algoritmalarını kaba kuvvet saldırı tekniğiyle şifrelerini kıran bu alandaki en popüler yazılım araçlardan biridir.
- c) Hydra: Uzak erişim protokolündeki şifreleri kaba kuvvet saldırı tekniği kullanarak kırmaya yarayan ayrıntılı şifre kırma programlarıdır. Hydra programı HTTP, VNC ve TELNET ağ protokollerini desteklemektedir. Linux işletim sisteminde ücretsiz olarak sunulan programın Windows işletim sistemi sürümü de bulunmaktadır.
- d) Elcomsoft Distributed Password Recovery: Farklı dosya türlerinin (FAT, NTFS gibi) şifrelerini kırmaya yarayan Elcomsoft şifre kırma programı resmi kurumlarda ve kişisel kullanıcılar tarafından sıklıkla tercih edilmektedir. Windows işletim sistemlerinde yaygın olarak kullanılmaktadır. Kaba kuvvet saldırı tekniğiyle çalışan bu program TrueCrypt, VeraCrypt, Bitlocker, PGP Disk ve LUKS şifreleme programları üzerinde başarılı sonuçlar vermektedir.
- e) Hashcat Password Recovery: Hashcat, kolayca hash algoritma şifrelerini kırmaya yarayan bir şifre kırma programıdır. Çalışma mantığında kaba kuvvet saldırı tekniğini kullanan Hashcat'in Linux, OS X ve Windows işletim sistemleri için sürümler bulunmaktadır.
- f) John the Ripper: John the Ripper özellikle hash algoritma şifrelerini kırma programıdır. Linux işletim sisteminde yaygın olarak kullanılan bu programın Windows işletim sistemi versiyonu da bulunmaktadır. Çalışma mantığı Hashcat programına benzeyen John the Ripper, kaba kuvvet saldırı tekniğiyle şifre kırma saldırıları düzenlemektedir.
- g) fgdump: fgdump şifre kırma programı lokal veya diğer kullanıcı hesap şifrelerini kırma için tasarlanmıştır. Programın ilk sürümü olan "pwdump" 2019 yılında güncellenmesiyle programın adı "fgdump" olarak değiştirilmiştir. fgdump şifre kırma programının son sürümünde şifre geçmişlerini görüntüleme özelliğinde eklenmiştir.
- h) Medusa: Uzak erişim sistemlerin şifrelerini kırma üzerine geliştirilmiş bir şifre kırma programıdır. Medusa hedef sistemdeki şifreleri kırma için kaba kuvvet saldırı tekniklerini kullanmaktadır. Nisan 2018 yılında güncellenen son sürümüyle güncellenerek HTTP, SQL, POP3, Telnet ve VNC gibi birçok ağ protokolleri desteklemektedir.
- j) Cain and Abel: Microsoft Windows için tasarlanmış bir şifre kırma programıdır. Kaba kuvvet ve kripto analiz teknikleri kullanan Cain and Abel programı, IOS, RDP ve PIX gibi ağ protokollerini desteklemektedir. Ayrıca son sürümünde getirilen güncellemeler ile IP üzerinden ağ parolası dinleme ve kayıt konuşma özelliğide eklenmiştir.
- k) RainbowCrack: RainbowCrack şifre kırma programı kaba kuvvet saldırı tekniği yerine gökkuşağı tabloları üretme tekniğiyle hedef sistemdeki şifreleri kırma üzerine tasarlanmıştır. Gökkuşağı tabloları olarak isimlendirilen bu tablolar önceden planlanmış olarak hazırlanmasıyla geleneksel yöntemlere göre daha kısa sürede şifreleri kırabilmektedir. Ağustos 2020 yılında güncellenen RainbowCrack şifre kırma hızı ve yüksek başarı oranıyla dikkat çekici olmakla birlikte tabloların hazırlanmasındaki güçlükler ve şifreleme algoritmaların karmaşıklığının artması yöntemin uygulanabilirliğini olumsuz yönde etkilemektedir.

4. LİTERATÜR TARAMASI

Şifre kırma saldırılarına bakıldığında kayıtlara geçmiş ilk olayların 90'lı yıllarında kaydedilmiştir. 16 Temmuz 1998 yılında ABD'de, topluluk acil müdahale ekibi (Community Emergency Response Team (CERT)) 186.126 şifrelenmiş verileri ait bir parolaları bulunduğunu bildirmiştir (Madsen 1998). 2009 yılında o güne kadar yaşanmış en büyük şifre kırma saldırısı (SQL Injection) olan "Rockyou.com" web hizmeti 32 milyona yakın şifrenin kırılarak piyasaya sürülmüştür (Noorunnisa vd. 2016).

2011 Haziran ayında NATO (Kuzey Atlantik Antlaşması Örgütü) üyelerinin 11.000 kayıtlı kullanıcılarının meta-data bilgileri ele geçirildi. Temmuz ayına gelindiğinde Amerika Birleşik Devletlerinde görev yapan özellikle Deniz Piyadeleri, Hava Kuvvetleri personeli ve İç işleri çalışanlarına ait bilgiler ele geçirilerek terör örgütlerine dağıtıldığı iddia edildi. Bu saldırılardan sonra birçok resmi kurum ve e-ticaret siteleri zayıf şifre kullanımını (1234, abcd gibi) yasakladı ve güçlü şifreleme algoritması (Büyük harf simge içeren en az sekiz karakterden oluşan gibi) standarttı kullanımını zorunlu hale getirmiştir (Thakur vd. 2019).

Günümüze kadar yapılan şifre kırma çalışmalarına bakıldığında halen en güvenli yöntemlerinin sözlük saldırısına veya kaba kuvvet yöntemlerine dayandığını görülmektedir (Oechslin vd 2003, Thing vd. 2009). Bununla birlikte, bu yöntemlere ek olarak Gökkuşluğu tabloları oluşturmak için çok miktarda hesaplama süresi ve çaba gerektirdiğinden (Billet vd. 2006) çalışmalarda çok tercih edilmediği de görülmektedir.

2009 yılında Weir ve ark. (Weir vd. 2009, Weir vd. 2010) şifreli verileri metin yapısını kullanarak şifrelerin kırılmasını sağlayan yeni bir şifre kırma tekniği sunmuşlardır. Bu çalışma sonuçlarının John the Ripper aracından daha etkili olduğunu göstermişlerdir (Houshmand vd. 2017). Başka bir çalışmada, Zhang ve ark. (Zhang vd. 2010). Weir'in önerdiği algoritmasının mevcut şifre kırma teknikleri arasında en etkili olduğunu iddia etmişlerdir. 2011 yılında Kelley ve ark. (Kelley vd. 2012) simüle edilmiş şifreler oluşturmak ve güçlerini değerlendirmek için Weir'in algoritmasını ve Markov modelinin bir varyasyonunu uyguladı. 2012 yılında Castelluccia (Castelluccia vd. 2012) ve Narayanan (Narayanan vd. 2005), şifrelenmiş dosyaların parolasını bağlamsal sıklığına dayandığı karakter tahminlerinden oluşan bir model önermiştir. 2017 yılına gelindiğinde Houshmand (Houshmand vd. 2017) yaptığı çalışmada şifre kırmak için şifreyi oluşturan kullanıcı hakkında bilgileri kullanan bir yöntem önermiştir.

Adli bilişim şifre kırma incelemeleri, artan siber suç oranına bağlı olarak incelenecek delil miktarı ve boyutunun artması, şifre kırma sürecinin karmaşıklığına bağlı olarak işlem süreçlerinin uzaması ve kullanılan donanımın maddi boyutu gibi birçok zorluklar içermektedir. Bu süreç içerisinde şifre kırma yöntemlerinin tanımlanması, delilerin kategorize etmek ve uygun tekniklerin belirlenmesi yönelik çalışmalar yürütülmüştür. 2013 yılında Al Fahdi ve arkadaşları yaptıkları çalışmada adli incelemelerdeki dijital delilerdeki artış miktarına bağlı olarak şifre kırma işlemlerinde belirgin ölçüde artacağını öngören bir anket çalışması adli tıp uzmanları üzerinde uygulamıştır (Al Fahdi vd. 2013). 2016 yılına Harichandran ve arkadaşları benzer bir çalışmada adli tıp uzmanlarının şifre kırma işlemlerinde yasal zorluklar, teknik eğitim gereksinimleri ve yeni nesil karmaşık şifre algoritmalarının zorluklarının ön gören bir anket yapılmıştır (Harichandran ve ark., 2016). Diğer bir çalışmada Lillis ve arkadaşları adli incelemelerdeki zorlukları üç ana kategoriye ayırmıştır. Bunlar; teknik zorluklar, yasa zorluklar ve kullanılan kaynak temininde yaşanan zorluklar olarak belirlemişlerdir (Lillis vd. 2016).

Adli bilişim incelemeleri dışında bilgisayar sistemleri üzerinde penetrasyon testi veya hesap kurtarma işlemlerinde şifre kırma işlemlerine ihtiyaç duyulmaktadır (Lehto ve Neittaanmäki. 2018). Özellikle adli bilişim incelemeleri gerçekleştirebilmek için özel şifre kırma programları kullanılmaktadır (Hassan, 2019). Adli incelemelerde şifre kırmak için en basit yöntem kapsamlı bir deneme veya kaba kuvvet saldırısı yöntemi kullanmaktır (Raza vd. 2012). Bu çalışmada seçilen adli vaka örneğinde şifre kırma işlemi kaba kuvvet saldırı yöntemi kullanan özel bir program kullanılmıştır.

Adli incelemelerde şifre kırmak için kullanılan başka bir yöntem ise potansiyel şifrelerin oluşan bir tablo (Hellman tabloları veya Gökkuşluğu tabloları gibi) kullanmaktır (Hellman, 1980; Wang vd.2013). İlk uygulamalarda başarı oranı yüksek olsada günümüzde kullanılan şifrelerin boyutu ve karmaşıklığının artması önceden hesaplanmış bir tablo ile şifre kırma işleminin başarısını son derece düşürmekte neredeyse imkansız hale getirmiştir. Diğer bir yöntem ise sözlük saldırılarıyla şifreleri kırmak üzerinedir. Bu yöntemde de kullanıcıların eğilimlerini (ad, soyad, memleket, futbol takımı vb.) test etmekten oluşmaktadır. Bir önceli yöntem olan tablo hazırlamasındaki zorluklar sözcük saldırısı yönteminde de olması nedeniyle adli bilişim incelemelerinde tercih edilmemektedir (Aggarwal vd. 2018). Modern şifre kırma yöntemleri özellikle makine öğrenmesi yönteminden faydalanarak hazırlanmış otomatikleştirilmiş programlar vasıtasıyla gerçekleştirilmektedir (Dürmuth vd. 2015). Bu yöntemin temelinde şifre kırma programlarının azalan olasılıklarla kaba kuvvet saldırısı mantığının dayanmaktadır (Hitaj vd. 2019). Günümüzde kaba kuvvet saldırısı yöntemini kullanan programlar adli bilişim incelemelerinde adli bilişim

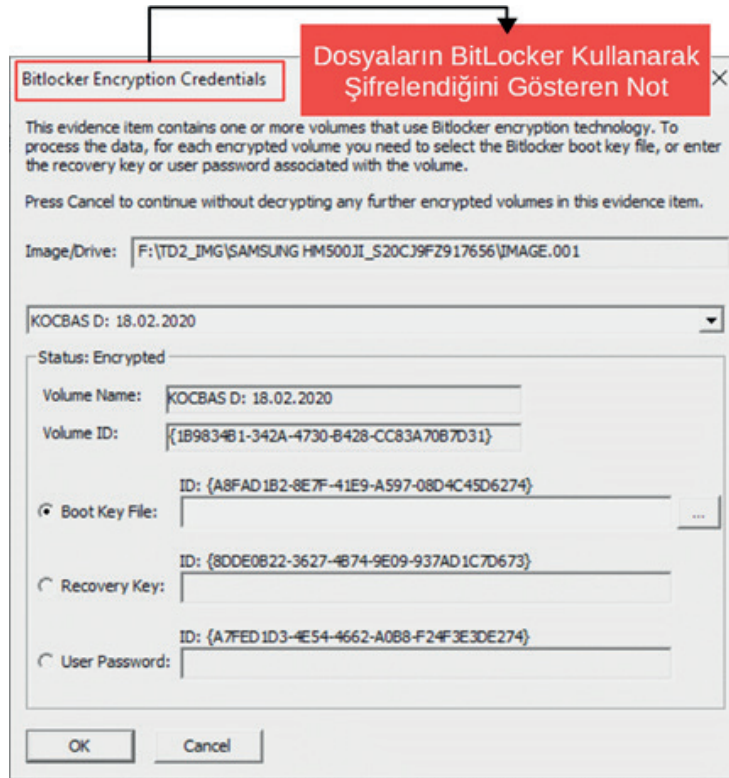
uzmanları tarafından şifre kırma incelemelerinde yaygın olarak tercih edilmektedir (Kanta vd. 2020). Benzer olarak bu çalışmada seçilen vaka örneği incelemelerinde kaba kuvvet saldırısı yöntemini kullanan PKF programı aracılığıyla gerçekleştirilmiştir.

5. ÖRNEK OLAY İNCELEMESİ

Bu çalışmada tüm analizler, Windows 10 Pro yazılıma sahip 2xXeon Gold 5118 /32GB / 256GB M.2 SSD sahip Dell Precision T7920 marka iş bilgisayarını ile yapıldı. Şifre kırma analizlerde “Passware Kit Forensic (PKF) V. 2019.2.0 (Ücretsiz sürümü)” programlarıyla yapılmıştır. Seçilen örnek adli bir olay olması (gerçek bir saldırı örneği) nedeniyle incelenecek şüpheli bilgisayarda bulunan verilere zarar vermemek ve delil bütünlüğünü koruma için öncelikle şüpheli bilgisayarın imal kopyası alınmıştır. İmaj kopyası alam işleminde uluslararası adli inceleme standartlarında kullanılan Forensic Toolkit (FTK) Imager (Ücretsiz sürümü) programı tercih edilmiştir.

İmaj kopyası alındıktan sonra analiz aşamasına geçilmiştir. Analiz aşamasında ilk adım olarak iş bilgisayarında analiz ortamının oluşturulması gereklidir. Analiz ortamı; incelenen delillerin olası saldırılarından iş bilgisayarının etkilenmemesi için gereklidir. Adli incelemelerde analiz ortamı sanal makine tercih edilmektedir. Sanal makine modunda çalıştırılan iş bilgisayarını incelemeler tamamlandıktan sonra ilk haline geri getirilebilmektedir. Bu çalışmada analiz ortamı Virtual Machine (VM) tercih edilmiştir.

Analiz ortamı kurulduktan sonra FTK Imager ile E02 formatında alınan imaj kopyası canlandırıldığında tüm bölümlerin şifreli olduğu görülmüştür (Şekil 8).

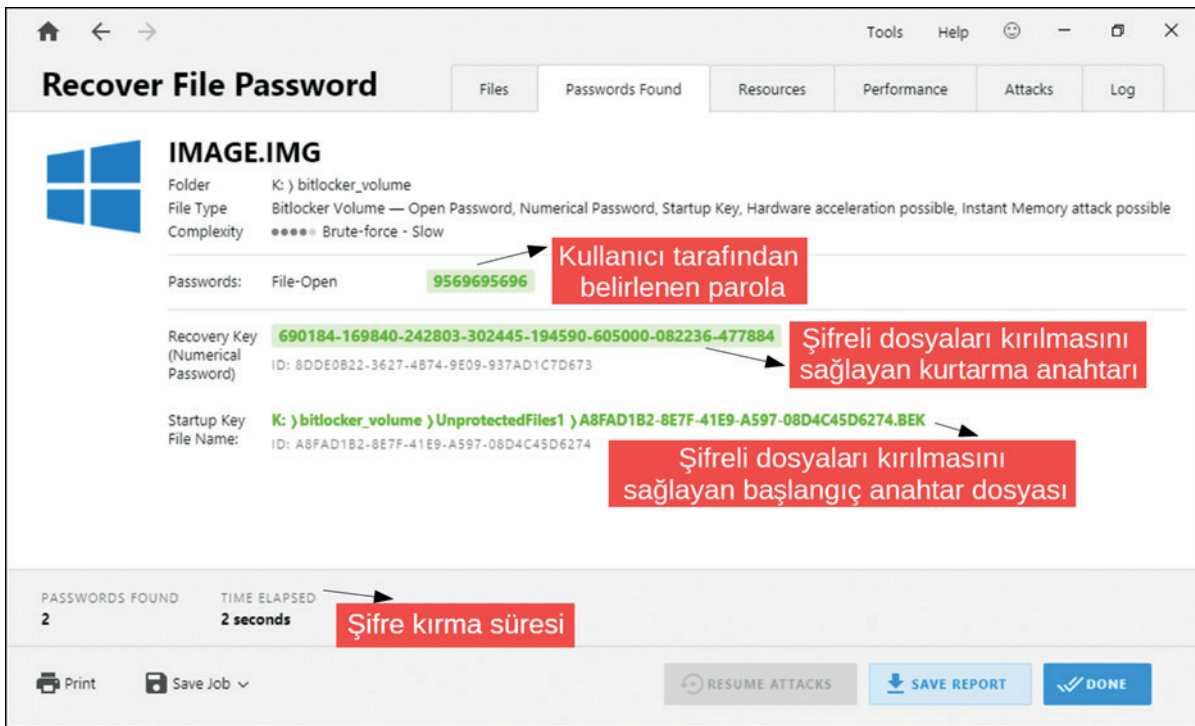


Şekil 8. Dosyaların “BitLocker” kullanarak şifrelendiğini gösteren ekran görüntüsü.

BitLocker, Windows işletim sistemlerinde kullanılan sistemde bulunan verileri ve dosyaları güvenli şekilde şifreleme yöntemidir. BitLocker, diğer dosya şifreleme yöntemlerinden farklı olarak, sistemdeki sürücünün tamamını şifrelemektedir. Kullanıcı sürücü içerisine bir veri ya da dosya eklendiğinde BitLocker otomatik şifreleme işlemini gerçekleştirmektedir. Böylece kullanıcı ağ üzerinden üçüncü kişilere veriler ya da dosyalarını paylaşsa da şifreli şekilde iletilir ve şifreler ancak yetkili kişi tarafından parolası girilerek açılmaktadır.

Şüpheli bilgisayardaki dosyaların şifreli olduğu ve şifreleme türü tespit edildikten sonra şifre kırma aşamasına geçilmiştir. Şifre kırmak için şifreleme türü olan “BitLocker” hangi programlar ile kırılacağı yönünde açık kaynak araştırması yapılmıştır. BitLocker, SHA-256 ve AES gibi oldukça karmaşık şifreleme algoritmasına sahiptir (Milo vd. 2011; Beşirli vd. 2019). 2019 yılında Agostini ve Bernaschi, BitLocker yöntemiyle şifrelenmiş dosyaların şifrenin kırılması ve erişim sağlanması üzerine vaka analizi çalışması yapmışlardır (Agostini, Bernaschi. 2019). Çalışmada Windows işletim sistemine sahip bilgisayarda BitLocker yöntemiyle şifrelenmiş dosyaların şifrelerini kırmak için geliştirdikleri “bitcracker” isimli bir program önermişlerdir. Çalışma sonuçları başarılı olsa da önerilen kullanılan bitcracker şifre kırma programının yeterliliği konusunun tartışmalı olduğu ve geliştirilmesi yönünde daha fazla çalışmalar yapılması gerekliliğini belirtmişlerdir. 2015 yılında Kumar yapmış olduğu çalışmada adli incelemelerde sıklıkla karşılaşılan bir durum olan disk üzerinde veya bir bölümündeki dosyaların şifreli olması halinde Passware Kit Forensic (PKF) programının kullanılabilirliğini önermiştir.

Yapılan araştırma sonucu adli analizlerde de sıklıkla kullanılan Passware Kit Forensic (PKF) ücretsiz sürümü kullanılabilirliği sonucuna varılmıştır. İnceleme bilgisayarına PKF yüklendikten sonra PKF’in “files” sekmesinden alınan imaj kopyası eklendi. İmaj kopyadaki şifrelerin kırılması için “Combine Attacks” (içerisinde sözcük saldırısı ve kaba kuvvet saldırısı bulunan) sekme seçilerek şifre kırma süreci başlatılmıştır. Analiz sonucunda “BitLocker” ile şifrelenmiş sürücü başarılı şekilde şifresi kırılmıştır (Şekil 9).



Şekil 9. Passware Kit Forensic (PKF) programıyla şifreli dosyaların şifre kırılma işlemi ekran görüntüsü.

Şekil 9’da PKF programıyla şifreli dosyaların başarılı bir şekilde kırıldığını görülmektedir. Şifre kırma işlemi sonucunda kullanıcı tarafından belirlenen parola “9569695696” tespit edilmiştir. Ayrıca şifreli dosyaların kırılmasını sağlayan kurtarma anahtarı ise “690184-169840-242803-302445-194590-605000-082236-477884” 48 haneden oluştuğu anlaşılmıştır. Bu sonuç BitLocker yöntemiyle şifrelenen dosyalarıda potansiyel şifrelerin oluşan bir tablo (Hellman tabloları veya Gökkuşluğu tabloları gibi) hazırlanması ve denenmesinin uygun bir yöntem olmadığı sonucuna varılabilir. Dahası PKF programı incelenen vaka örneğindeki şifreyi (geçen süre: 2 saniye) oldukça kısa bir sürede başarıyla tamamladığı görülmektedir.

6. SONUÇLAR

Bilgi teknolojilerin kullanımı ve bilginin kullanılmasının artmasıyla yeni sorunların da beraberinde getirilmiş özellikle bilginin güvenli olarak saklanması yetkisiz kişilerin erişimine geçmemesi için güvenlik önlemleri alınması zorunlu hale

getirmiştir. Bilginin öneminin artmasıyla saldırganlarda bilgiyi ele geçirmek için yeni yöntemler geliştirmekte her geçen gün yeni saldırı türlerini piyasaya sürmektedir. Bu durum, e-ticaret şirketler, resmi devlet kurumları ve kişisel kullanıcılar bu saldırılara engel olmak ve yetkisiz kişilerin erişimini engellemek için yeni tedbirler alamaya zorlanmaktadır. Alınan bu tedbirler (yeni şifreleme yöntemlerinin zorluk ve karmaşıklığı) bazı durumlarda şifreli verilere ulaşılmasının zorlaştırmakta hatta kırılması onlarca yıl zaman alabilmektedir. Özellikle şifrelerin unutulması ya da adli incelemelerde şifreli verilerin incelemesi sürecine verilere ulaşılmasında büyük bir engel oluşturmaktadır.

Bilgi sistemleri veya dosyaların şifreleme yöntemler ve kullanılan algoritmalar çok çeşitlilik göstermekle birlikte karşılaşılan olaylara yönelik herkes tarafından kabul gören bir şifre kırma yöntemi bulunmamakla birlikte bazı durumlarda sadece tek bir olaya uygulanabilir bir yaklaşım geliştirilmesi gerekebilmektedir. Bu durum şifre kırma sürecinin en büyük problemlerinden birisidir.

Şifrelemiş verilerin kırılması için yöntemler geliştirilmeye devam etse de mevcut uygulamaların başarı oranı halen tartışma konusudur. Kullanılan yöntem ve şifre kırma programları kadar şifre kırma için kullanılan iş bilgi ayarlarının donanım özellikleri de süreci etkileyen önemli bir etkidir. Bu çalışmada şifreleme saldırıları, yöntemleri ve kullanılan en popüler şifre kırma programları ele alınarak gerçek bir adli inceleme örneği üzerinde şifre kırma süreci detaylı olarak incelenmiştir. Seçilen adli vaka örneği “BitLocker” yöntemi kullanılarak şifrelenmiş olduğu tespit edildikten sonra şifrenin kırılması için yapılan açık kaynak araştırma sonucunda PKF programı şifre kırma işlemlerinde kullanılabilceği öngörülmüştür ve PKF şifre kırma programı tercih edilmiştir. Yapılan şifre kırma işlemleri sonucunda adli vaka örneğinde başarıyla şifrenin kırılmasıyla sonuçlanmıştır. Seçilen adli vaka örneğinde kullanılan yöntemin ve seçilecek şifre kırma programının tercih edilirken adli uzmanları ve diğer araştırmacılar tarafından tekrarlanabilir olması için ücretsiz programlar kullanılmasına özellikle dikkat edilmiştir. Sonuç olarak çalışma sonuçlarıyla adli bilişim uzmanlarına ve bu alanda yapılacak akademik çalışmalara katkı sağlayacağı düşünülmektedir. Adli vaka analizlerinde kullanılan şifre kırma yaklaşımları benzer olmakla beraber farklılıklarda içerebilmektedir. Bu nedenle çalışmanın farklı örneklerle tekrarlanması faydalı olacağı düşünülmektedir.

Hakem Değerlendirmesi: Dış bağımsız.

Çıkar Çatışması: Yazar çıkar çatışması bildirmemiştir.

Etik Komite Onayı: Etik komite onayı alınmıştır.

Finansal Destek: Yazar bu çalışma için finansal destek almadığını beyan etmiştir.

Teşekkür: Yazar, bu çalışmanın analiz kısımlarının hazırlanmasında altyapı desteği sağlayan Çankırı Karatekin Üniversitesi Eldivan Sağlık Hizmetleri Meslek Yüksek Okulu Birimine teşekkür eder.

Peer-review: Externally peer-reviewed.

Conflict of Interest: The author has no conflict of interest to declare.

Ethics Committee Approval: This study approved by an ethical committee

Grant Support: The author declared that this study has received no financial support.

Acknowledgement: The author would like to thank Çankırı Karatekin University Eldivan Health Services Vocational School Unit for providing infrastructural support in the preparation of the analysis parts of this study.

Kaynaklar/References

- Al Fahdi, M., Clarke, N. L., & Furnell, S. M. (2013, August). “Challenges to digital forensics: A survey of researchers & practitioners attitudes and opinions.” In 2013 Information Security for South Africa (pp. 1-8). IEEE.
- Aggarwal, S., Houshmand, S., & Weir, M. (2018). “New technologies in password cracking techniques.” In Cyber Security: Power and Technology (pp. 179-198). Springer, Cham.
- Agostini, E., & Bernaschi, M. (2019). “BitCracker: BitLocker meets GPUs”. arXiv preprint arXiv:1901.01337.
- Beşkirli, A., Özdemir, D., & Beşkirli, M. (2019). “Şifreleme Yöntemleri ve RSA Algoritması Üzerine Bir İnceleme”. Avrupa Bilim ve Teknoloji Dergisi, 284-291.
- Bhanot, R., Hans, R. (2015). “A review and comparative analysis of various encryption algorithms.” International Journal of Security and Its Applications, 9(4): 289-306.
- Billet O., Gilbert, H. (2006). “Cryptanalysis of rainbow.” Security and Cryptography for Networks, 4116:336-347.
- Castelluccia C., Durmuth M., Perito, D. (2012). “Adaptive password-strength meters from Markov models.” Proc. of the Network and Distributed System Security Symposium.
- Dass, A.S., Prabhu, J. (2020). “Hybrid coherent encryption scheme for multimedia big data management using cryptographic encryption methods.” International Journal of Grid and Utility Computing, 11(4):496-508.

- Dürmuth, M., Angelstorf, F., Castelluccia, C., Perito, D., & Chaabane, A. (2015, March). "OMEN: Faster password guessing using an ordered markov enumerator". In International Symposium on Engineering Secure Software and Systems (pp. 119-132). Springer, Cham.
- Guddeti, P., Dharavath, N. (2020). "Analysis of password protected Document." COMPUSOFT: An International Journal of Advanced Computer Technology, 9(7): 3762-3767.
- Harichandran, V. S., Breitinger, F., Baggili, I., & Marrington, A. (2016). "A cyber forensics needs analysis survey: Revisiting the domain's needs a decade later." Computers & Security, 57, 1-13.
- Hassan, N. A. (2019). "Digital Forensics Basics: A Practical Guide Using Windows OS." Apress.
- Hellman, M. (1980). "A cryptanalytic time-memory trade-off." IEEE transactions on Information Theory, 26(4), 401-406.
- Hitaj, B., Gasti, P., Ateniese, G., & Perez-Cruz, F. (2019, June). "Passgan: A deep learning approach for password guessing." In International Conference on Applied Cryptography and Network Security (pp. 217-237). Springer, Cham.
- Hur, U., Park, M., Kim, G., Park, Y., Lee, I., Kim, J. (2019). "Data acquisition methods using backup data decryption of Sony smartphones." Digital Investigation, 31:200890.
- Houshmand S., Aggarwal S. (2017). "Using personal information in targeted grammar-based probabilistic password attacks." In: IFIP International Conference on Digital Forensics. 285-303.
- Kara, İ. (2019). "Kaba Kuvvet Saldırı Tespiti ve Teknik Analizi." Sakarya University Journal of Computer and Information Sciences, 2(2): 61-69.
- Kaya, Ö. F., Öztürk, E. (2017). "Veri ve Ağ Güvenliği İçin Uygulama ve Analiz Çalışmaları." İstanbul Ticaret Üniversitesi Fen Bilimleri Dergisi, 16(31): 85-102.
- Kelley P.G., Komanduri S., Mazurek M.L., Shay R., Vidas, T., Bauer, L., ... Lopez, J. (2012). "Guess again (and again and again): Measuring password strength by simulating password-cracking algorithms." In 2012 IEEE symposium on security and privacy, 523-537.
- Kanta, A., Coisel, I., & Scanlon, M. (2020). "A survey exploring open source intelligence for smarter password cracking." Forensic Science International: Digital Investigation, 35, 301075.
- Kumar, S. (2015). "Digital Evidence-Technical Issues." Advances in Computer Science and Information Technology (ACSIT). 2(11) 42-47.
- Lillis, D., Becker, B., O'Sullivan, T., & Scanlon, M. (2016). "Current challenges and future research areas for digital forensic investigation." arXiv preprint arXiv:1604.03850.
- Lehto, M., & Neittaanmäki, P. (Eds.). (2018). "Cyber Security: Power and Technology" (Vol. 93). Springer.
- Saračević, M. H., Adamović, S. Z., Mišković, V. A., Elhoseny, M., Maček, N. D., Selim, M. M., & Shankar, K. (2020). "Data Encryption for Internet of Things Applications Based on Catalan Objects and Two Combinatorial Structures." IEEE Transactions on Reliability.
- Oechslin, P. (2003). "Making a faster cryptanalytic time-memory trade-off," Advances in Cryptology, 617-630.
- Raza, M., Iqbal, M., Sharif, M., & Haider, W. (2012). "A survey of password attacks and comparative analysis on methods for secure authentication." World Applied Sciences Journal, 19(4), 439-444.
- Thing V.L.L., Ying H.M. (2009). "A Novel Time-Memory Tradeoff Method for Password Recovery."
- Noorunnisa, N.S., Afreen, D.K.R. (2016). "Review on Honey Encryption Technique." International Journal of Science and Research, 2319-7064.
- Madsen W. (1998). "Encryption debate rages again." Network Security, 5: 8-9.
- Milo, F., Bernaschi, M., & Bisson, M. (2011). "A fast, GPU based, dictionary attack to OpenPGP secret keyrings." Journal of Systems and Software, 84(12), 2088-2096.
- Thakur, S., Singh, A.K., Ghrera, S.P., Elhoseny, M. (2019). "Multi-layer security of medical data through watermarking and chaotic encryption for tele-health applications." Multimedia tools and Applications, 78(3):3457-3470.
- Narayanan A., Shmatikov V. (2005). "Fast dictionary attacks on passwords using time-space tradeoff," Proc. of the 12th ACM Conference on Computer and Communications Security, 2005.
- Zhang Y., Monroe F., Reiter M.K. (2010). "The security of modern password expiration: An algorithmic framework and empirical analysis." In Proceedings of the 17th ACM conference on Computer and communications security, 176-186.
- Wang, X. J., Liao, X. F., & Huang, H. Y. (2013). "Improvement of rainbow table technology based on number cutting of reduction function." Comput. Eng, 7, 36.
- Weir M.S., B. Aggarwal de Medeiros., Glodek B. (2009). "Password cracking using probabilistic context-free grammars," Proc. of the 30th IEEE Symposium on Security and Privacy, 391-405.
- Weir M., Aggarwal S., Collins M., Stern, H. (2010). "Testing metrics for password creation policies by attacking large sets of revealed passwords." In Proceedings of the 17th ACM conference on Computer and communications security, 162-175.

Mobil Devlet Kullanımını Etkileyen Faktörlerin Belirlenmesi: Ankara İli Örneği

Determining Factors Affecting Mobile Government Usage: Ankara Province Case

Vildan Ateş¹ 



ÖZ

Kablosuz ve mobil teknolojilerde yaşanan gelişmeler halka hizmet sunmada yeni bir kanal olan mobil devlete (m-devlet) olan ilgiyi arttırmıştır. Bu doğrultuda m-devlet uygulaması üzerinden sunulan hizmetlerin kullanıcılar tarafından benimsenmesi ve kullanılması kritik öneme sahiptir. Bu çalışmanın amacı, m-devlet kullanımını etkileyen faktörlerin belirlenmesidir. Literatür taraması sonucu m-devlet kullanımını etkilediği düşünülen yedi faktörden (bilgi kalitesi, güvenlik algısı, güven, hizmet maliyeti, kullanım kolaylığı, sosyal etki, yararlılık) oluşan kuramsal model ve çalışmanın hipotezleri oluşturulmuştur. Veri toplama aracı olarak M-Devlet Kullanımı Ölçeği (M-DKÖ) geliştirilmiştir. Çalışmanın verileri Ocak 2020 tarihinde Ankara'daki üç devlet üniversitesinde öğrenim gören 183 üniversite öğrencisinden toplanmıştır. M-devlet kullanımını etkileyen faktörlerin belirlenmesi için hazırlanan ölçeğinin faktör yapısı LISREL paket programı kullanılarak analiz edilmiştir. Ölçeğin yapı güvenirliği ve yakınsama geçerliği doğrulayıcı faktör analizi (DFA) ile gerçekleştirilmiştir. Analizler sonucu kuramsal modeldeki yedi faktörün t değerleri, hata varyansları, standart katsayıları ve uyum indeksleri kontrol edilmiştir. Çalışma sonucunda kuramsal modeldeki yedi faktörden biri olan hizmet maliyeti dışındaki diğer altı faktörün m-devlet kullanımını etkilediği görülmüştür.

Anahtar Kelimeler: M-Devlet, M-Devlet Kullanımı, Yapısal Eşitlik Modeli, Türkiye

ABSTRACT

The developments in wireless and mobile technologies have increased the interest in the mobile government (m-government), which is a new channel in providing services to the citizens. It is critical users adopt and use the services offered through the m-government application. The purpose of this study is to determine the factors affecting the use of m-government. As a result of the literature review, a theoretical model consisting of seven factors (information quality, security perception, trust, service cost, ease of use, social impact, usefulness) was created and hypotheses of the study were presented. M-Government Usage Scale (M-GUS) was developed as a data collection tool. The data of the study was collected from 183 university students studying in three public universities in Ankara in January 2020. The scale was prepared to determine the factors affecting the use of m-government. The factor structure was analyzed by using the LISREL package program. The construct reliability and convergence validity of the scale were performed by confirmatory factor analysis (CFA). As a result of the analysis, t values, error variances, standard coefficients and fit indexes of the seven factors in the theoretical model were checked. As a result of the study, it was observed that six factors except the service cost, which is one of the seven factors in the theoretical model, affect the use of m-government.

Keywords: M-Government, Usage Of M-Government, Structural Equation Model, Turkey

¹(Dr. Öğr. Üyesi), Ankara Yıldırım Beyazıt Üniversitesi, İşletme Fakültesi, Yönetim Bilişim Sistemleri, Ankara, Türkiye

ORCID: V.A. 0000-0002-8855-8556

Corresponding author:

Vildan ATEŞ
Ankara Yıldırım Beyazıt Üniversitesi, İşletme Fakültesi, Yönetim Bilişim Sistemleri, Ankara, Türkiye
E-mail address: vvirginates@gmail.com

Submitted: 23.01.2021

Revision Requested: 01.02.2021

Last Revision Received: 19.02.2021

Accepted: 23.02.2021

Published Online: 12.05.2021

Citation: Ates, V. (2021). Mobil devlet kullanımını etkileyen faktörlerin belirlenmesi: Ankara ili örneği. *Acta Infologica*, 5(1), 39-52. <https://doi.org/10.26650/acin.866945>

1. GİRİŞ

Mobil teknoloji insanların teknolojiyi kullanma biçimlerini değiştirerek sunulan hizmetlerinden her yerde ve her zaman yararlanma fırsatı sunmaktadır. Bu nedenle mobil cihazlar kullanıcıya hizmet sunmada yeni bir yöntem olarak kullanılmaya başlanmıştır. Mobil teknolojilerde yaşanan bu gelişmeler mobil devlet (m-devlet) uygulamalarının geliştirilmesine yol açmıştır. M-devlet, vatandaş ve devlet arasındaki iletişimi kullanıcı hareket halinde dahi olsa hızlı bir şekilde sağlayacak şekilde uygulamaların mobil cihazlar üzerinden kullanılması olarak tanımlanabilir. M-devlet kamu hizmetlerinin sunulmasında mobil teknolojiler kullanılarak e-devlet uygulamalarının yaygınlaştırılması ve geliştirilmesi şeklinde düşünülen yeni nesil bir hizmet sunumu olarak görülmektedir. E-devletin bir alt kümesi ve tamamlayıcısı olan m-devlet, e- devlet uygulamalarını daha geniş kitlelere ulaştırabilmesi, kalitesini arttırması, zamandan tasarruf sağlaması ve kullanılabilirliğini arttırması özellikleriyle e-devleti destekleyen önemli bir unsurdur (Alotaibi, Houghton & Sandhu, 2016).

Türkiye’de m-devlet uygulaması 2012 yılından beri kullanılmaktadır. M-devlet uygulaması içerisinde 2869 hizmet sunulmakla birlikte e-devlet kapsamında yer alan hizmetlerin tamamı bu uygulama içerisinde bulunmamaktadır (E-Devlet Portalı, 2021). 2019 yılı sonu itibarı ile e-devlet portalına genel giriş sayısı 1.179.624.439 iken bu girişlerin yaklaşık %40’ının mobil uygulama üzerinden yapıldığı görülmektedir (Önal, 2020). Bu oran 2012 yılında sadece % 0,5 iken sonraki yıllarda mobil uygulama üzerinden e-devlet kullanımının arttığını ve gelecek yıllarda daha da artacağı öngörülebilir. M-devlet uygulamasının özellikle genç nüfus (18-24 yaş) tarafından kullanılması ve benimsenmesi uygulamanın amacına ulaşması ve geleceği için kritik bir öneme sahiptir. Bu doğrultuda m-devlet kullanıcısı olan üniversite öğrencilerinin görüşleri alınarak m-devlet kullanımını etkileyen faktörler araştırılacaktır.

Literatür taraması sonucunda farklı ülkelerde m-devlet kullanımını etkileyen faktörler ve engellerin neler olduğunu araştıran çalışmaların olduğu görülmüştür (Abdelghaffar & Magdy, 2012; Almarashdeh & Alsmadi, 2017; Abu-Shanab & Haider, 2015; Al-khamaysch, Lawrence & Zmijewska, 2006; Hung, Chang & Kuo, 2013; Ishengoma, Mselle & Mongi, 2019; Roubah, Qurban, Al-Qirim & Tarhini, 2018; Shareef, Archer & Dwivedi, 2012; Saxena, 2017; Tan, Ooi, Sim & Phusavat, 2012; Wang, 2014). Türkiye’de ise m-devlet hakkında yapılan çalışmaların çok az sayıda olduğu ve bu çalışmalarında m-devlet hizmetlerini açıklamaya yönelik olduğu görülmektedir (Arslan, 2012; Güler & Dövertaş, 2009; Gürses & Engin, 2016; Öztürkcan, Kasap & Eryaysoy, 2012; Üste & Güzel, 2012; Yıldırım & Panayırıcı, 2012). M-devlet kullanımını etkileyen faktörleri araştıran ulusal düzeyde bir çalışmanın olmaması bu araştırmanın motivasyon kaynağıdır. Buna ilaveten bu çalışma bilişim teknolojilerine daha yatkın ve yoğun kullanan 18-24 yaş aralığındaki kullanıcılar ile gerçekleştirildiği için de önemlidir.

Bu çalışmanın amacı, Ankara’daki üniversite öğrencilerinin m-devleti kullanmalarını etkileyen faktörlerin belirlenmesidir. Çalışma sonucunda m-devlet kullanımını etkileyen faktörlerin neler olduğunun belirlenmesi ve bu faktörlerin etki gücünün ortaya çıkarılması hedeflenmektedir.

Çalışma altı bölümden oluşmaktadır. İkinci bölümde literatür taraması sunulmuştur. Üçüncü bölümde ise m-devlet kullanımını etkileyebileceği düşünülen faktörler ve araştırmanın hipotezlerine yer verilmiştir. Dördüncü bölümde yöntem detaylı bir şekilde açıklandıktan sonra beşinci bölümde çalışmanın bulguları rapor edilmiştir. Altıncı bölümde çalışmanın sonucu ve önerilere yer verilmiştir.

2. LİTERATÜR TARAMASI

Literatür taraması bölümü iki alt bölümden oluşmaktadır. İlk olarak dünyadaki farklı ülkelerde m-devlet kullanımını etkileyen faktörlerini araştıran çalışmalar sunulmuştur. İkinci bölümde de Türkiye’de m-devlet hakkında yapılan çalışmalar incelenmiştir.

Farklı ülkelerde m-devlet kullanımını etkileyen faktörlerin araştırıldığı çalışmalar görselliği arttırmak ve okunabilirliği kolaylaştırmak için Tablo 1’de sunulmuştur. Çalışmalar kronolojik olarak sıralanmıştır. Tablo 1’de görüldüğü gibi çalışmalar 2006 ila 2018 yılları arasında gerçekleştirilmiştir. Tablo 1’de bulunan tüm faktörler değerlendirildiğinde özellikle kullanılabilirlik, kullanım kolaylığı, risk, güven ve güvenlik faktörlerinin diğer faktörlere göre daha fazla çalışmada m-devlet kullanımını etkileyen faktörler olarak rapor edildiği görülmektedir.

Literatür taramasının ikinci bölümünde Türkiye’de m-devlet konusunda yapılan çalışmalar incelenmiştir. Güler ve Dövertaş (2009) tarafından yapılan çalışmada Türkiye’de sunulan m-devlet hizmetleri merkezi ve yerel yönetim düzeyinde incelenmiştir. Diğer iki çalışmada da Türkiye’deki belediyelerin sunduğu m-devlet hizmetleri kullanımı hakkında olduğu görülmektedir (Arslan, 2012; Gürses & Engin, 2016). Türkiye’deki m-devlet uygulamaları hakkında bilgiler veren bir çalışma da Üste ve Güzel (2012) tarafından gerçekleştirilmiştir. Aynı yıl Yıldırım ve Panayırıcı (2012) tarafından gerçekleştirilen çalışmada Türkiye’de mobil devlete geçişteki çabalar ve yeniliklere vurgu yapılmıştır. Sadece Öztürkcan ve diğerleri (2018) tarafından yapılan çalışmada m-devlet kullanıcı kabulü açısından incelendiği görülmektedir. Bu çalışmada veriler telefon görüşmesi ile toplanmıştır. Çalışma sonucunda kullanıcıların m-devlet ile aldıkları hizmetleri cep telefonu ile oy kullanmak, vergiyi öğrenmek ve ödemek, emniyet hizmetlerine erişmek, sağlık hizmetleri almak, adres değişikliği yapmak şeklinde rapor etmişlerdir.

M-devletin vatandaş devlet iletişimde büyük bir etkisi bulunmasına rağmen Türkiye’de m-devlet üzerine yapılan az araştırma olduğu görülmektedir. M-devlet kullanımını etkileyen faktörlerin kapsamlı bir şekilde araştırıldığı bir çalışma ise bulunmamaktadır.

Tablo 1

Önceki çalışmaların özeti

Yazar/lar	Yıl	Ülke	M-Devlet Kullanımını Etkileyen Faktörler
Alkhamayseh et al.	2006	Global	Altyapı, e-devletin kabulü, gizlilik ve güvenlik, kalite, kullanıcı ihtiyaçları, kullanılabilir uygulamalar, maliyet
Abdelghaffar & Magdy	2012	Mısır	Farkındalık, kullanılabilirlik, sosyal etki, yüz-yüze iletişim
Shareef et al.	2012	Hindistan	Empati, güvenlik, kullanım kolaylığı, sağlanan avantaj
Tan et al.	2012	Malezya	Kişisel kurallar, kullanım kolaylığı, yararlılık
Hung et al.	2013	Tayvan	Dışsal etki, etkileşim, kullanılabilirlik, kullanım kolaylığı, kolaylaştırıcı koşullar
Wang	2014	Çin	Değer algısı, kullanılabilirlik, güven
Abu-Shanab & Haider	2015	Ürdün	Cevap verme yeteneği, kullanılabilirlik, kullanım kolaylığı, uygunluk
Alotaibi et al.	2016	Suudi Arabistan	Deneyim, farkındalık, güven, hoşlanmak
Almarashdeh & Alsmadi	2017	Suudi Arabistan	Hizmet maliyeti, kullanılabilirlik, kullanım kolaylığı, sosyal etki, teknolojiye güven
Saxena	2017	Hindistan	Güven, kullanılabilirlik, kullanım kolaylığı, tutum, öz yeterlilik
Roubah et al.	2018	Kuveyt	Ara yüz tasarımı, beklenti, bilgi kalitesi, değer algısı, farkındalık, hizmet kalitesi, güven, tasarım, memnuniyet, risk ve gizlilik algısı
Ishengoma et al.	2019	Tanzanya	Altyapı, erişim, güvenlik, kişisel faktörler, kullanılabilirlik, maliyet

3. MOBİL DEVLET KULLANIMINI ETKİLEYEN FAKTÖRLER VE HİPOTEZLER

M-devlet kullanımını etkilediği düşünülen yedi faktöre yapılan literatür taraması sonucu ulaşılmıştır. Bu faktörler Bilgi Kalitesi (BK), Güvenlik Algısı (GG), Güven (GÜ), Hizmet Maliyeti (HM), Kullanım Kolaylığı (KK), Sosyal Etki (SE) ve Yararlılık (YA) şeklindedir. Aşağıda bu faktörler açıklanmış ve bu faktörler ile M-Devlet Kullanma Niyeti (KN) arasındaki çalışmanın hipotezleri sunulmuştur.

Bilgi kalitesi, m-devlette sunulan hizmetlerdeki bilgilerin doğruluğunu, eksiksizliğini ve güncelliğini göz önünde bulundurur. M-devlette sunulan bilgilerin doğru, doğrulanabilir, tam, ilgili ve zamanında olması kullanıcılar için önemli kriterlerdir. Dolayısıyla, aşağıdaki hipotez önerilmektedir:

H1: Bilgi kalitesi (KK) m-devlet kullanma niyetini (KN) doğrudan ve olumlu etkileyecektir.

Güvenlik ise kullanıcıların bilgilerinin izinsiz veya yetkisiz bir biçimde erişim, kullanım, değiştirilme, ortadan kaldırılma ve hasar verilmesini önlemek olarak tanımlanabilir. Gizlilik güvenliğinin alt bileşenlerinden biri olup bilginin yetkisiz kişilerin eline geçmeme ve yetkisiz erişime karşı korunmasıdır (Çolak, 2020). Gizlilik ve güvenlik m-devletin uygulanması aşamasında karşılaşılan en önemli konulardır. Yapılan çalışmalarda da gizlilik ve güvenliğinin başarılı m-devlet için ayırt edici özellikler

olduğu vurgulanmıştır (Kushchu & Kuscu, 2003; Sandy & McMillan, 2005; Wang, 2014). Buna ilaveten güvenlik algısının yüksek olması m-devlet kullanımını olumlu yönde etkilediği çalışma sonuçlarında da görülmektedir (Shareef et al., 2012; Tan et al., 2012; Wang, 2014). Bu sonuçlara göre, aşağıdaki hipotez önerilmektedir:

H2: Güvenlik algısı (GG) m-devlet kullanma niyetini (KN) doğrudan ve olumlu etkileyecektir.

Diğer bir faktör güven olup bir tarafın beklentilerini karşılayacak şekilde karşı tarafın beklenen işlemi gerçekleştirme yönündeki öznel değerlendirmesidir. Bu nedenle güven, başkalarının gelecekteki davranışları hakkındaki belirsizliğin neden olduğu emin olamamayı içerirken karşı tarafın da kötü niyetli olmayacağına olan inancıdır (Colesca, 2009; Tan et al., 2012). Güven yapılan birçok e-devlet çalışmasında e-devlet kullanımını etkileyen faktörler arasında yer almaktadır (Khasawneh, Rabayah & Abu-Shanab, 2013; Wang & Lo, 2013; Rehman, Kamal & Esichaikul, 2012; Bannister & Connolly, 2011). Aynı doğrultuda m-devlet kullanımını da etkileyeceği açıktır. Bu doğrultuda aşağıdaki hipotez önerilmektedir:

H3: Güven (GÜ) m-devlet kullanma niyetini (BI) doğrudan ve olumlu etkileyecektir.

Hizmet maliyeti kullanıcıların m-devleti kullanma ve hizmet almak istemeleri sonucu karşılaştıkları maliyettir. Kullanıcılar m-devlet uygulamasını kullanabilmek için akıllı mobil cihaza, cihaz üzerinde uygulamayı indirebilmek için alana, kablolu ya da kablosuz internet bağlantısına ihtiyaç duymaktadırlar. Bu maliyet birimlerinin herhangi birisini tedarik edememe durumu m-devlet kullanma niyetleri olumsuz etkileyebilecektir. Dolayısıyla aşağıdaki hipotez önerilmektedir:

H4: Hizmet maliyeti (HM) m-devleti kullanma niyetini (BI) doğrudan ve olumsuz etkileyecektir.

Kullanım kolaylığı bir sistemin kullanımı esnasında kullanıcının minimum çaba ile başarabileceğine olan inancıdır. Teknolojiyi kullanmak kolay ise teknoloji önündeki engeller de kalkacaktır (Davis, 1989). Kullanım kolaylığı inancının derecesi yükseldikçe kullanıcının sistemi kullanma niyeti de artacaktır. Çalışma sonuçlarında kullanım kolaylığının m- devlet kullanımını etkilediği belirtilmiştir (Shareef et al., 2012; Saxena, 2017). Bu nedenle aşağıdaki hipotez önerilmektedir:

H5: M-devletin kullanım kolaylığı (KK) m-devleti kullanma niyetini (BI) doğrudan ve olumlu etkileyecektir.

Sosyal etki kişilerin görüşlerini dikkate aldığı kişilerin düşünceleri sonucu davranışlarını değiştirmesidir. Başkaları tarafından kabul edilme ve beklentilerini karşılama arzusu sonucu kişiler m-devlet hizmetlerini kullanmaya başlamaktadırlar (Abu-Shanab & Haider, 2015; Venkatesh & Davis, 2000). Buna dayanarak aşağıdaki hipotez önerilmektedir:

H6: Sosyal etki (SE) m-devlet kullanma niyetini (BI) doğrudan ve olumlu etkileyecektir.

Yararlılık, kullanıcıların m-devlet uygulamasını kullanmaları durumunda sağlayacakları faydalardır. M-devleti kullanmak herhangi bir fayda ya da değer eklediği zaman kullanıcılar m-devleti kullanmak istemeyeceklerdir. Buna dayanarak aşağıdaki hipotez önerilmektedir:

H7: Yararlılık (YA) m-devlet kullanma niyetini (BI) doğrudan ve olumlu etkileyecektir. Sonuç olarak çalışmada yukarıda açıklanan yedi tane hipotez bulunmaktadır.

4. METODOLOJİ

Bu çalışmada nicel araştırma yöntemi kullanılmıştır. Bu bölümde sırasıyla araştırmanın çalışma grubu, kuramsal modeli, veri toplama aracı, verilerin toplanması ve verilerin analizi sunulmuştur.

4.1 Çalışma Grubu

Araştırmaya katılacak üniversite öğrencileri seçilirken seçkisiz olmayan örnekleme yöntemlerinden uygun örnekleme yöntemi kullanılmıştır. Araştırmaya hız kazandırması ve erişilebilmesi kolay olan katılımcılara olanak sağladığı için bu yöntem tercih edilmiştir. Bu araştırmanın çalışma grubunu Ankara'da üç devlet üniversitesinde 2019-2020 öğretim yılı güz döneminde öğrenim gören 183 üniversite öğrencisi oluşturmaktadır. M-devlet kullanımını konusunda yapılan benzer çalışmalarda da Rana ve diğerleri (2019) 81, Abdelghaffer ve Magdy 100, Abaza ve Saif (2015) 120 ve Almarashdeh ve Alsmadi (2017) 468 katılımcı sayıları ile çalışmalarını gerçekleştirmişlerdir. Bu nedenle bu araştırmanın çalışma grubunun sayısının yeterli olduğu düşünülmüştür. Ayrıca öğrencilerin mutlaka m-devleti kullanıyor olmalarına dikkat edilmiştir. Üniversite öğrencilerinin

seçilmiş olmasının sebebi Türkiye İstatistik Kurumu 2019 yılı Hanehalkı Bilişim Teknolojileri Kullanım Araştırmasına göre 18-24 yaş grubunun bilgisayar ve interneti en yoğun (%90,8) olarak kullanan yaş aralığı olmasıdır (Bayrak, 2020). Öğrenciler Endüstri ve Makine Mühendisliği, Fen Bilgisi Öğretmenliği ve Yönetim Bilişim Sistemleri bölümleri öğrencileri olup, araştırmaya gönüllü olarak katılmışlardır. Tablo 2’de araştırmaya katılan çalışma grubunun demografik özellikleri görülmektedir.

Tablo 2 incelendiğinde katılımcıların büyük çoğunluğunun (%94) 18-24 yaş aralığında olduğu görülmektedir. Buna ilaveten katılımcıların %65’i 3. ve 4.sınıf öğrencileridir. Buna ilaveten mühendislik ve yönetim bilişim sistemleri öğrencilerinin çoğunlukta (%60) olması bu öğrencilerin teknolojik gelişmelere ve bilişim teknolojilerine daha ilgili olmalarından dolayı araştırma için avantaj olarak değerlendirilebilir.

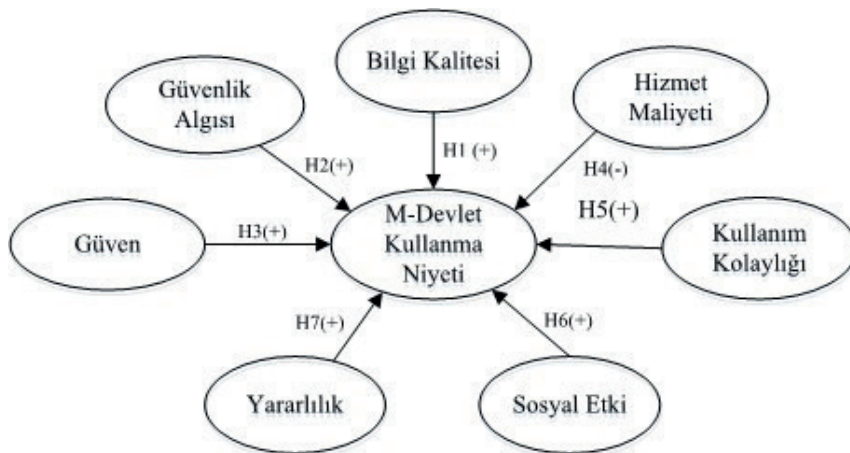
Tablo 2

Çalışma grubunun demografik özellikleri

Cinsiyet	Sayı	%
Kadın	109	59
Erkek	74	41
Yaş	Sayı	%
18-20	40	22
21-23	130	71
24-27	13	7
Sınıf	Sayı	%
1.Sınıf	22	12
2.Sınıf	41	22
3.Sınıf	46	25
4.Sınıf	74	41
Bölüm	Sayı	%
Makine Mühendisliği	11	6
Endüstri Mühendisliği	26	14
Fen Bilgisi Öğretmenliği	73	40
Yönetim Bilişim Sistemleri	73	40

4.2 Araştırmanın Kuramsal Modeli

Kuramsal modeli geliştirebilmek için ilk olarak literatür taraması yapılmıştır. Yapılan tarama sonucunda m-devlet kullanımını etkilediği düşünülen faktörler tespit edilmiştir. M-devlet kullanımını etkilediği düşünülen faktörler: bilgi kalitesi, güvenlik algısı, güven, hizmet maliyeti, kullanım kolaylığı, sosyal etki, yararlılıktır. Kuramsal modelde yedi tane faktör gözlenemeyen gizil değişkenler olup bir tane bağımlı gözlenemeyen değişken (m-devleti kullanma niyeti) bulunmaktadır. Bu çalışmadaki yedi faktör gizil değişkenler olup bu değişkenler ile m-devlet kullanımı arasındaki nedensel ilişkiler açıklanmaya çalışılmıştır. Bu gizil değişkenler arasındaki ilişkiler ve hipotezler Şekil 1’de sunulmuştur.



Şekil 1. Kuramsal model ve hipotezler

4.3. Veri Toplama Aracı

M-devlet kullanımını etkileyen faktörleri belirlemek için bir ölçek hazırlanmıştır. Hazırlanan ölçekteki maddelerin belirlenmesi için kapsamlı bir literatür taraması yapılarak farklı ülkelerin m-devlet kullanımını etkileyen faktörlerin araştırıldığı araştırma makaleleri incelenmiştir. Ölçekte yer alması uygun görülen maddeler için araştırmacılarla elektronik posta aracılığıyla iletişime geçilmiş ve gerekli izinler alınmıştır. Bu çalışmalar İngilizce olduğundan ölçek maddelerinin Türkçeye çevrilmesi süreci üç aşamada gerçekleştirilmiştir. Birinci aşamada dilsel eş değeri sağlamak için ölçek maddelerini Türkçeye çevirecek kişiler belirlenmiştir. Çeviri işleminde üç kişi bulunmaktadır. İki kişi yüksek lisans ve doktorasını yurt dışında tamamlamış iki öğretim üyesi olup üçüncü kişi İngilizce eğitim verilen bir üniversitedeki öğretim üyesidir. Üç kişi ölçek maddelerini çevirdikten sonra bir araya getirilerek gerekli uyumlaştırma çalışmaları yapılarak düzeltmeler yapılmıştır. İkinci aşamada Türk Dili ve Edebiyatı bölümü öğretim üyesi tarafından dil bilgisi kuralları ve anlaşılabilirliği dikkate alınarak değerlendirilmiştir. Öneriler doğrultusunda gerekli düzeltmeler yapılmıştır. Üçüncü aşama da on üniversite öğrencisi ile ölçek maddeleri tartışılmış ve ölçek maddelerinin son haline ulaşılmıştır. Veri toplama aracı olarak 30 maddeden oluşan M-Devlet Kullanımı Ölçeği (M-DKÖ) geliştirilmiştir. Tablo 3’de ölçme aracındaki faktörler, madde sayıları ve maddelerin derlendiği kaynaklar görülmektedir. Ölçekte 5’li Likert ölçeği kullanılmış ve 1-5 şeklinde (1. Kesinlikle katılmıyorum 5. Kesinlikle katılıyorum) ölçeklendirilmiştir. Tablo 4’te ölçek maddeleri görülmektedir.

Tablo 3

Ölçme aracındaki faktörler ve madde sayıları

Faktör adı	Madde sayısı	Kaynaklar
Bilgi kalitesi	5	Abu-Shanab & Haider, 2015; Colesca, 2009
Güvenlik algısı	4	Colesca, 2009
Güven	4	Almarashdeh & Alsmadi, 2017
Hizmet kalitesi	3	Almarashdeh & Alsmadi, 2017; Saxena, 2017
Kullanım kolaylığı	4	Colesca, 2009; Saxena, 2017
Sosyal etki	3	Abu-Shanab & Haider, 2015; Saxena, 2017
Yararlılık	4	Almarashdeh & Alsmadi, 2017
M-devlet kullanma niyeti	3	Almarashdeh & Alsmadi, 2017; Saxena, 2017

Tablo 4

M-Devlet kullanımı ölçeği maddeleri

Bilgi kalitesi	1	M-devlet hizmetler hakkında kapsamlı bilgiler sunmaktadır.
	2	M-devlet hizmetler hakkında doğru bilgiler sunmaktadır.
	3	M-devlet uygulamasındaki bilgiler günceldir.
	4	M-devletin kullanımı şeffaflık düzeyini artırmaktadır.
	5	M-devletin kullanımı vatandaşlar arasında bilgiye ulaşım konusunda fırsat eşitliği sağlamaktadır.
Hizmet maliyeti	6	M-devlet hizmetlerini kullanmak için gerekli akıllı telefon ve benzeri mobil cihazların pahalı olduğunu düşünüyorum.
	7	M-devlet hizmetlerine erişimin pahalı olduğunu düşünüyorum.
	8	M-devlet hizmetleri işlem maliyetinin pahalı olduğunu düşünüyorum.
Güvenlik algısı	9	M-devlet gizliliğimi korumaktadır.
	10	M-devlet uygulamasında kişisel bilgilerim korunmaktadır.
	11	M-devlet sisteminde kişisel bilgilerime benim dışımda kimse ulaşamaz.
	12	M-devlet sistemi siber saldırılara karşı korunmaktadır.
Güven	13	Devletin mobil hizmetleri güvenli bir şekilde sunacağına inanıyorum.
	14	M-devletin internet altyapısına güveniyorum.
	15	M-devlet hizmeti alırken kullandığım mobil cihaza güveniyorum.
	16	M-devlet hizmeti veren kamu kurumlarına güveniyorum.
Sosyal etki	17	Karar verme süreçlerinde görüşlerini dikkate aldığım kişiler m-devleti kullanmam gerektiğini düşünüyor.
	18	Çevremdeki önem verdiğim insanlar, m-devleti kullanabileceğimi düşünüyor.
	19	Çevremdeki önem verdiğim kişiler, m-devleti kullanmamın faydalı olacağını düşünüyor.

Kullanım kolaylığı	20	M-devlet hizmetlerinin sunulduğu uygulamaların kullanımı kolaydır.
	21	M-devlette sunulan yeni bir hizmet uygulamasının kullanımını öğrenmek benim için kolaydır.
	22	M-devlette sunulan hizmetler açık ve anlaşılardır.
	23	M-devlette ihtiyaç duyduğum her hizmeti almak benim için kolaylık olurdu.
Yararlılık	24	M-devlet uygulamasındaki hizmetleri kullanmak, işlerimi daha hızlı yapmama yardımcı oluyor.
	25	M-devlet uygulamasındaki hizmetleri alabilmek hayatımı kolaylaştırıyor.
	26	M-devlet uygulamasındaki hizmetleri günlük yaşamım için faydalı buluyorum.
	27	M-devlet uygulamasındaki hizmetleri kullanmak verimliliğimi artırıyor.
M-devlet kullanma niyeti	28	İhtiyacım olduğunda m-devlet kullanmayı düşünüyorum.
	29	Hizmet almak için m-devlet uygulamalarını sık sık kullanmayı düşünüyorum.
	30	Fırsat verildiğinde m-devlet kullanmayı düşünüyorum.

4.4. Verilerin Toplanması

Verileri toplama sürecine başlamadan önce Ankara Yıldırım Beyazıt Üniversitesi Sosyal ve Beşeri Bilimler Etik Kurulu'ndan (27.12.2019 tarih ve 53 karar no) onay alınmıştır. Onaydan sonra verilerin toplanması Ocak 2020 yılında Ankara'daki üç devlet üniversitesinde gerçekleştirilmiştir. Veri toplama süreci araştırmacı tarafından gerçekleştirilmiştir.

4.5. Verilerin Analizi

M-devlet kullanımını etkileyen faktörlerin belirlenmesi için hazırlanan ölçeğin faktör yapısı LISREL 8.72 paket programı kullanılarak analiz edilmiştir. Ölçeğin geçerlik çalışmaları doğrulayıcı faktör analizi (DFA) ile gerçekleştirilmiştir. Şekil 1'deki kuramsal model Yapısal Eşitlik Modeli (YEM) yöntemiyle analiz edilmiştir. Verilerin analizi sonuçlarında maddelerin t değerleri, hata varyansları ve standart yükleri kontrol edilmiştir. Kuramsal modelin gerçek verilerle uyumunu değerlendirmek için Yaklaşık Hataların Ortalama Karekökü (Root Mean Square Error of Approximation, RMSEA), Normlandırılmış uyum indeksi (The Normed Fit Index, NFI), Normlandırılmamış uyum indeksi (The Non-Normed Fit Index, NNFI), Karşılaştırmalı uyum indeksi (Comparative Fit Index, CFI), Karşılaştırmalı Uyum İndeksi (Comparative Fit Index, CFI), Uyum iyiliği indeksi (Goodness-of-fit index, GFI) ve Düzeltilmiş İyilik Uyum İndeksleri (Adjusted Goodness of Fit Index, AGFI) olmak üzere altı uyum indeksi değeri kontrol edilmiştir (Cole & Maxwell, 2003). Buna ilaveten ki-kare (χ^2) değeri örneklem büyüklüğünden etkilendiği ve geniş örneklemelerde olumlu sonuç verdiği için analizlerde ki-karenin serbestlik derecesine (χ^2 /sd) oranına bakılmıştır (Tabachnick, Fidell & Osterlind, 2007:506).

5. BULGULAR

Bu bölümde sırasıyla ölçeğin güvenirlik ve geçerlik ile kuramsal modelin test edildiği YEM yöntemi analiz sonuçları sunulmuştur.

Ölçeğin güvenirliği için güvenirlik katsayısı hesaplama tekniklerinden iç tutarlılık yöntemi seçilmiştir. Bunun içinden de Cronbach alfa iç güvenirlik katsayısı yöntemi kullanılmıştır. Ölçek için hesaplanan Cronbach alfa iç güvenirlik katsayısı ($\alpha=0,89$) 0,89 olup bu değer ölçeğin yüksek derecede güvenilir olduğunu göstermektedir. Her faktöre ait Cronbach alfa iç güvenirlik katsayıları ise; bilgi kalitesi için 0,88, güvenlik algısı ve güven için 0,83, sosyal etki için 0,87, kullanım kolaylığı için 0,84, hizmet maliyeti ve yararlılık için 0,73'tür. Bilgi kalitesi, güvenlik algısı, güven, sosyal etki, kullanım kolaylığı faktörleri yüksek derece güvenilir olup hizmet maliyeti ve yararlılık faktörleri ise oldukça güvenilirdir.

Verilerin analizi için YEM uygulanmış ve DFA LISREL 8.72 bilgisayar programında gerçekleştirilmiştir. Verilerin hazırlanması süreci üç adımda gerçekleştirilmiştir. Birinci adımda veriler tanımlanmıştır. İkinci adımda DFA için kovaryans matrisi oluşturulmuştur. Üçüncü adımda da gerçekleştirilecek her bir analiz için sözdizimleri LISREL programında yazılarak verilerin analizi süreci başlatılmıştır.

Şekil 1'de görüldüğü gibi bu çalışmanın kuramsal modelinde sekiz ölçme modeli bulunmaktadır. İlk olarak her ölçme modelinin faktör yapısı DFA ile test edilmiştir. Yapılan analizleri değerlendirmek için ilk olarak t değerleri kontrol edilmiştir. Her ölçme modelindeki maddelerin tümünün t değerinin 2,56'dan büyük olduğu için 0,01 düzeyinde istatistiksel olarak

anlamli olduđu sonucuna varılmıřtır (Çokluk, Şekerçiođlu & Büyüköztürk, 2010:223; Nunnally, 1994). İkinci olarak hata varyanslarına bakılmıř ve hata varyanslarının da düşük (<90) olduđu görölmüřtür. Son olarak her bir ölçme modelinin dođrulanıp dođrulanmadığını test etmek için uyum indeksleri (χ^2/sd , RMSEA, NFI, NNFI, CFI, GFI, AGFI) deđerlendirilmiřtir. Deđerlendirmeler sonucunda kuramsal modelde bulunan sekiz ölçme modeli dođrulanmıřtır. Daha sonra Şekil 1'de görölen kuramsal modele DFA uygulanmıřtır. Tablo 5'te kuramsal modelin DFA sonuçları görölmektedir.

Tablo 5 incelendiđinde tüm maddelerin hata varyanslarının 0,23 ile 0,66 arasında deđiřtiđi görölmektedir. Bu durum yüksek hata varyansının olmadıđı řeklinde yorumlanıp istenen bir durumdur. Diđer incelenmesi gereken deđer maddelerin path katsayıları (standartlařtırılmıř faktör yükleri) olup 0,64 ile 0,88 arasında deđiřmektedir ve bu path katsayıları parametrik sınamada t deđerlerine göre istatistiksel olarak anlamlı bulunmuřtur. Diđer taraftan Tablo da göröldüđu gibi altı madde (BK1, GG1, GU1, SE1, KK1 ve YA1) için t deđeri hesaplanmamıřtır. Buna rađmen düşük hata varyansları ve yüksek path katsayılarına sahip oldukları için ölçme modelinden çıkarılmamıřlardır.

Tablo 5'te göröldüđu gibi her bir faktörün yapı güvenilirliđi hesaplanmıřtır. Her bir faktör için yapı güvenilirliđi 0,70'den büyük olduđundan bu deđerler ölçme sonuçlarının güvenilir olduđunu göstermektedir (Nunnally, 1994). Buna ilaveten yakınsama geçerliđi için faktör yükleri ve ortalama açıklanan varyans (OAV) deđerleri incelenmiřtir. Faktör yüklerinin ve OAV deđerlerinin hepsinin 0,50'den büyük olması yakınsama geçerliliđinin sađlandığını göstermektedir (Fornell & Larcker, 1981). Ayrıca yapısal güvenilirlik deđerlerinin her biri de OAV deđerlerinden büyüktür ve bu durum da yakınsama geçerliliđinin bir diđer kanıtı olarak gösterilebilir.

Fornel ve Larcker (1981), ayırt edici geçerlik için kuramsal modelin faktörleri arasındaki korelasyonlardan ve OAV deđerlerinin karekökünden yararlanılması gerektiđini belirtmiřlerdir. Herhangi boyuttaki OAV'ın karekökünün o faktörün diđer faktör ile aralarındaki korelasyondan ve aynı zamanda 0,50 deđerinden küçük olmaması gerektiđini vurgulamıřlardır. Kuramsal modelin faktörleri arasındaki korelasyonlar ve OAV deđerlerinin karekök deđerleri Tablo 6'da görölmektedir.

Tablo 5

Kuramsal modelin DFA sonuçları

Maddeler	Hata Varyansları	Path Katsayıları	t Deđerleri ¹	Yapı Güvenirliđi	Ortalama Açıklanan Varyans	R ²
Bilgi Kalitesi						
BK1	0,24	0,84	-			0,71
BK2	0,23	0,88	14,57**			0,77
BK3	0,33	0,82	13,14**	0,80	0,56	0,67
BK4	0,54	0,69	10,26**			0,47
BK5	0,49	0,71	10,83**			0,51
Hizmet Maliyeti						
HM1	0,62	0,61	3,46**			0,38
HM 2	0,26	0,86	7,98**	0,82	0,60	0,74
HM 3	0,30	0,83	8,06**			0,70
Güvenlik Algısı						
GG1	0,24	0,87	-			0,76
GG2	0,23	0,87	13,49**			0,75
GG3	0,33	0,63	9,07**	0,83	0,56	0,40
GG4	0,66	0,54	8,25**			0,34
Güven						
GU1	0,46	0,74	-			0,54
GU2	0,29	0,84	10,78**			0,71
GU3	0,59	0,64	8,20**	0,91	0,58	0,41
GU4	0,36	0,80	10,31**			0,64

Sosyal Etki						
SE1	0,27	0,85	-			0,73
SE2	0,32	0,83	12,85**	0,90	0,70	0,68
SE3	0,31	0,83	12,99**			0,69
Kullanım Kolaylığı						
KK1	0,25	0,86	-			0,75
KK2	0,39	0,78	12,54**	0,86	0,61	0,61
KK3	0,46	0,73	11,38**			0,54
KK4	0,45	0,74	11,62**			0,55
Yararlılık						
YA1	0,36	0,80	-			0,64
YA2	0,24	0,87	13,48**	0,89	0,68	0,76
YA3	0,28	0,85	12,99**			0,72
YA4	0,40	0,78	11,55**			0,60
M-Devlet Kullanma Niyeti						
KN1	0,36	0,80	12,72**			0,64
KN2	0,37	0,79	12,66**	0,84	0,64	0,63
KN3	0,34	0,81	13,05**			0,66
	Standartlaştırılmış katsayılar (beta)		t değerleri	R ² değerleri		
BK → KN	0,69		8,93**	0,48		
HM → KN	-0,14		-1,70	0,02		
GG → KN	0,62		7,93**	0,38		
GU → KN	0,80		6,95**	0,64		
SE → KN	0,71		9,18**	0,51		
KK → KN	0,89		12,18**	0,79		
YA → KN	0,90		11,31**	0,81		

¹t-değerleri *p < .05 ; **p < .01 düzeyinde anlamlıdır.

Tablo 6

Faktörler arasındaki korelasyon ve OAV değerlerinin karekökü

Faktörler	Sembol	[1]	[2]	[3]	[4]	[5]	[6]	[7]
Bilgi Kalitesi	[1]	0,75						
Hizmet Maliyeti	[2]	-0,10	0,78					
Güvenlik Algısı	[3]	0,43	-0,09	0,75				
Güven	[4]	0,55	-0,11	0,49	0,76			
Sosyal Etki	[5]	0,49	-0,10	0,44	0,57	0,84		
Kullanım Kolaylığı	[6]	0,61	-0,13	0,55	0,71	0,63	0,78	
Yararlılık	[7]	0,62	-0,13	0,56	0,72	0,64	0,78	0,80

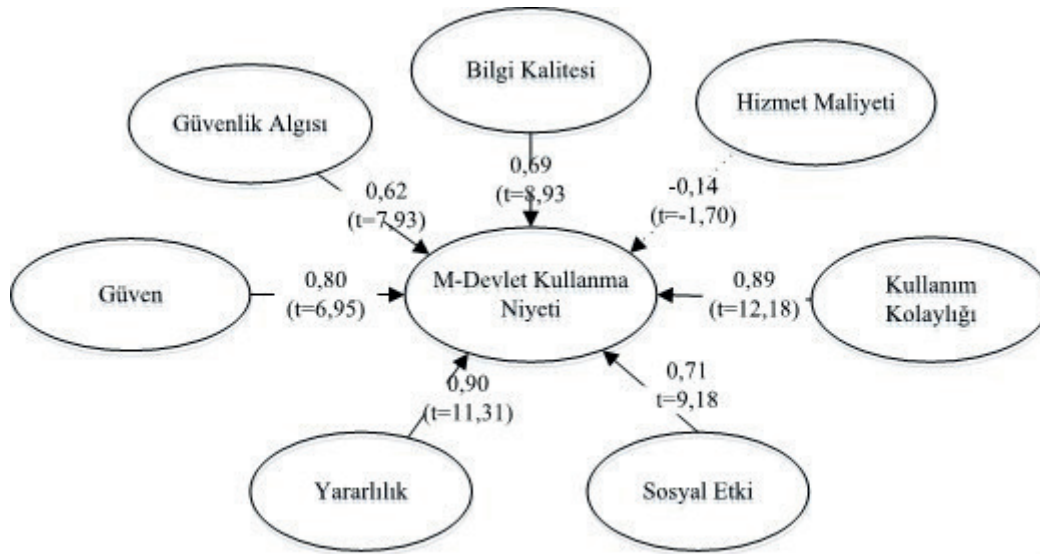
Tablo 6 incelendiğinde korelasyon matrisinin köşegen elemanları OAV'ın karekökünü (koyu olan değerler) ve köşegen dışı elemanlar ise boyutlar arasındaki korelasyon değerlerini göstermektedir. Tablo 6'da görüldüğü gibi her bir boyuttan hesaplanan OAV değerlerinin karekökü diğer alt faktörler ile olan korelasyon değerlerinden daha büyüktür. Örneğin sosyal etki faktörünün OAV karekökü 0,84 olup bu faktörün diğer faktörlerle olan korelasyonundan daha büyük olduğu görülmektedir. Bu durum diğer tüm boyutlar için de geçerli olup sadece kullanım kolaylığı faktörünün OAV değerinin karekökü ile yararlılık arasındaki korelasyon değeri eşittir. Bu sonuçlar doğrultusunda ölçeğin ayırt edicilik geçerliği sağladığı söylenebilir.

Kuramsal model için hesaplanan uyum iyiliği istatistikleri de Tablo 7'de görülmektedir. Kuramsal model için Tablo 7'deki uyum ölçütleri incelendiğinde RMSEA indeksi sadece zayıf uyum gösterirken, CFI, GFI ve AGFI kabul edilebilir uyum, X²/d, NFI ile NNFI iyi uyum göstermektedir. Sonuç olarak modelin kabul edilebilir uyum gösterdiği söylenebilir.

Tablo 7

Kuramsal modelin uyum ölçüleri

Uyum Ölçüsü	Değeri	Uyum
X ² /(d=398)	2,7	İyi uyum
RMSEA	0,09	Zayıf uyum
NFI	0,92	İyi uyum
NNFI	0,94	İyi uyum
CFI	0,91	Kabul edilebilir uyum
GFI	0,87	Kabul edilebilir uyum
AGFI	0,85	Kabul edilebilir uyum



Şekil 2. Araştırma modeli ve DFA sonuçları

Şekil 2’de kuramsal modelde yer alan bağımlı gizil değişken ve bağımsız gizil değişkenler arasındaki nedensel ilişkiler görülmektedir. Şekil 2 incelendiğinde BK ile KN arasındaki katsayısının 0,69’dur. Öğrenciler için bilgi kalitesi önemli olup bilgi kalitesi arttığında m-devlet kullanma niyetlerinde de artış olacağı yönünde yorumlanabilir. Sonuç olarak H1 hipotezi desteklenmiştir. Diğer bir faktör olan GG ile KN arasındaki katsayı 0,62’dir. Öğrencilerin güvenlik algısı olumlu arttığında m-devlet kullanma niyetleri artacak olup bu durumu ifade eden H2 hipotezi de desteklenmiştir. Çalışmanın H3 hipotezi GU ile KN arasında olup aralarındaki katsayı değeri 0,80’dir. Bu değer öğrencilerin güvenlerinin artması durumunda m-devlet kullanma niyetlerinin de artacağı şeklinde yorumlanabilir. H3 hipotezi de istatistiksel olarak desteklenmiştir. Diğer taraftan HM ile KN arasındaki katsayı -0,14 olup manidar t değeri de (-1,70) vermediğinden dolayı H4 hipotezi desteklenememiştir. KK ile KN arasındaki ilişki katsayısı 0,89 olduğu Şekil 2’de görülmektedir. KK’daki bir puanlık artış KN’de 0,89 puanlık yükselişe neden olacak olup H5 hipotezi de desteklenmiştir. Çalışmanın altıncı hipotezi SE ile KN arasındadır. Bu iki değer arasındaki ilişki katsayısı 0,71 olup H6 hipotezi de desteklenmiştir. En yüksek ilişki katsayısının 0,90 ile YA ile KN arasında olduğu görülmektedir. Bu değer, öğrencilerin m-devleti kullanmalarının yararlı olduğunu düşündüklerinde kullanmalarının artacağını göstermektedir. YA’daki bir puanlık artış KN’de 0,90 puanlık yüksek bir artışa neden olacaktır. Bu ilişkinin ifade edildiği H7 hipotezi de istatistiksel olarak desteklenmiştir.

Sonuç olarak H1, H2, H3, H5, H6 ve H7 hipotezleri desteklenirken H4 hipotezi desteklenememiştir.

6. SONUÇ VE ÖNERİLER

Bu çalışmada üniversite öğrencilerinin m-devlet kullanımını etkileyen faktörler araştırılmıştır. Çalışmanın verileri Ankara’daki üç devlet üniversitesindeki öğrencilerden toplanmıştır. Verilere uygulanan DFA sonucu kavramsal modelde önerilen yedi faktörden (bilgi kalitesi, güvenlik algısı, güven, hizmet maliyeti, kullanım kolaylığı, sosyal etki, yararlılık) sadece biri hariç

(hizmet maliyeti) diğer altısının m-devlet kullanımını etkilediği görülmüştür. Çalışma sonuçlarına göre faktörlerin m-devleti kullanma niyetine etkisine göre büyükten küçüğe şöyle sıralanmaktadır: Yararlılık, kullanım kolaylığı, güven, sosyal etki, güvenlik algısı, bilgi kalitesidir. Bu çalışmanın belirleme katsayısı (R^2) %69,4 olup bu altı faktör birlikte m-devlet kullanımını etkileyen faktörlerin %69,4'ünü açıklamaktadır.

Çalışma sonuçları literatür ile paralellik göstermektedir. İlk üç sırada m-devlet kullanımını etkilediği görülen faktörler yararlılık, kullanım kolaylığı ve güvendir. Mobil Hizmet Kabul Modeli'ne göre de yararlılık, kullanım kolaylığı ve güven mobil hizmeti kullanma niyetini direk etkilemektedir (Gao, Krogstie & Gransæther, 2008). Bir diğer çalışma sonucunda da bu üç faktöre ilaveten erişim, altyapı ve kullanılabilirlik de m-devlet kullanmayı etkileyen faktörler olarak rapor edilmiştir (Ishengoma et al., 2019). Farklı ülkelerde yapılan araştırma sonuçlarının büyük çoğunluğunda da bu üç faktörün etkili olduğu Tablo 1'de görülmektedir.

Çocukların ve gençlerin çevrelerindeki kişilerden etkilenmeleri ve davranışlarını bu doğrultuda düzenlemeleri sıklıkla rastlanılan bir durumdur. Bu çalışma sonucunda da üniversite öğrencilerinin çevrelerindeki insanların m-devlet kullanımında etkili olduğu görülmüştür. Literatürde bulunan teknoloji ya da hizmet kabul modelleri (TAM2, UTAUT, TPB) içerisinde de sosyal etkinin bulunduğu görülmektedir (Davis, 1989; Venkatesh & Davis, 2000; Venkatesh et al., 2003). Buna ilaveten Almarashdeh ve Alsmadi tarafından 2016 yılında yapılan ve kullanıcıların m-devleti benimseme ve kullanma faktörlerini araştıran çalışma sonucunda da en fazla etkili olan faktör sosyal etki olarak rapor edilmiştir.

Güvenlik algısının olumlu yönde olması da m-devlet kullanımını olumlu etkilemektedir. Güvenlik algısının m-devlet kullanımını etkileyen faktörler arasında beşinci sırada olması dikkat çekicidir. Bu durum günümüzde bilişim sistemlerinin güvenliğinin sağlanması kullanıcılar tarafından bir zorunluluk olarak görüldüğü şeklinde yorumlanabilir. Literatürle (Al-khamayseh et al., 2006; Roubah, et al., 2018) paralellik gösteren bu sonuç aynı zamanda kullanıcıların Türkiye'de m-devlet ile güvenli bağlantılar kurdukları algısına sahip olduklarını da göstermektedir. Üniversite öğrencilerinin güvenlik algılarının olumlu olması m-devlet için bir başarı olup geleceğini de olumlu etkileyecektir.

Bilgi kalitesi m-devlet kullanımını etkileyen bir diğer faktör olup yapılan çalışmalarda da bilişim sistemleri, e-devlet ve m-devlet kullanıcı niyetini etkilediği rapor edilmiştir (Koivumaki, Ristola & Kesti, 2008; Veeramootoo, Nunkoo & Dwivedi, 2018; Wang, 2008).

Bu araştırmanın bulgularına göre hizmet maliyeti faktörü m-devlet kullanımını etkilememektedir. Bu sonuç değerlendirildiğinde öğrencilerin m-devlet hizmetine erişimi, erişebilecekleri cihazı ve işlem maliyetini kolaylıkla karşılayabildikleri söylenebilir. Literatürde yer alan araştırma sonuçları değerlendirildiğinde bazı çalışma sonuçlarında hizmet maliyetinin m-devlet kullanımını etkilediği görülürken (Al-Hadidi & Rezgui, 2010; Al-khamayseh et al., 2006; Ishengoma et al., 2019) bazıları da etkilemediğini belirtmektedir. (Abu-Shanab & Haider, 2015; El-Kiki & Lawrence, 2006; Hung et al., 2013; Rana et al., 2019). Sonuç olarak çalışmanın bulguları etkili ve etkin bir m-devlet sistemi kurmaktan sorumlu olan birimler için yol gösterici olacaktır. Başarılı bir m-devlet uygulamasına tüm paydaşların katkılarıyla ulaşılabileceği unutulmamalıdır. Bu doğrultuda en önemli paydaş olan kullanıcıların m-devleti kullanmalarını etkileyen faktörlerin bilinmesi ve bu doğrultuda çalışmaların yapılması kritik öneme sahiptir.

Her çalışmanın olduğu gibi bu çalışmanın da bazı sınırlıkları bulunmaktadır. Çalışma Ankara ilinde ve devlet üniversitelerinde gerçekleştirilmiş nicel bir araştırmadır. Gelecekte buna benzer çalışmaların kapsamı yüksek lisans öğrencileri ile özel üniversitelerde okuyan üniversitelerde okuyan öğrencilerle görüşmeler yapılarak genişletilebilir. Bu çalışmada m-devlet kullanımını etkileyen yedi temel faktör değerlendirilmiştir. Ayrıca m-devlet kullanımını etkileyebilecek muhtemel aracı değişkenler araştırılarak sunulan modele eklenebilir.

Hakem Değerlendirmesi: Dış bağımsız.

Çıkar Çatışması: Yazar çıkar çatışması bildirmemiştir.

Finansal Destek: Yazar bu çalışma için finansal destek almadığını beyan etmiştir.

Peer-review: Externally peer-reviewed.

Conflict of Interest: The author has no conflict of interest to declare.

Grant Support: The author declared that this study has received no financial support.

Kaynaklar/References

- Abaza, M., & Saif, F. (2015). The adoption of mobile government services in developing countries. *IJCSI International Journal of Computer Science Issues*, 12(1), 137-145.
- Abdelghaffar, H., & Magdy, Y. (2012). The adoption of mobile government services in developing countries: The case of Egypt. *International Journal of Information and Communication Technology Research*, 2(4), 333-341.
- Abu-Shanab, E., & Haider, S. (2015). Major factors influencing the adoption of m-government in Jordan. *Electronic Government an International Journal*, 11(4), 223-240.
- Al-Hadidi, A., & Y. Rezgui, Y. (2010). Adoption and diffusion of m-government: challenges and future directions for research. *In Working Conference on Virtual Enterprises* (pp. 88-94). Berlin, Germany.
- Al-khamayseh, S., Lawrence, E., & Zmijewska, A. (2006). Towards understanding success factors in interactive mobile government, *Second European Conference on Mobile Government*, Brighton, England.
- Almarashdeh, I., & Alsmadi, M. (2016). Heuristic evaluation of mobile government portal services: An experts' review. *In 2016 11th International Conference for Internet Technology and Secured Transactions (ICITST)* (pp. 427-431), Barcelona, Spain.
- Almarashdeh, I., & Alsmadi, M.K. (2017). How to make them use it? Citizens acceptance of m-government. *Applied Computing and Informatics*, 13(2), 194-199.
- Alotaibi, R., Houghton, L., & Sandhu, K. (2016). Exploring the potential factors influencing the adoption of m-government services in Saudi Arabia: A qualitative analysis. *International Journal of Business and Management*, 11(8), 56-71.
- Arslan, A. (2012). M-government service provisions among Turkish local governments. *İnternet Uygulamaları ve Yönetimi Dergisi*, 3(2), 5-26.
- Bannister, F., & R. Connolly, R. (2011). Trust and transformational government: a proposed framework for research. *Government Information Quarterly*, 28(2), 137-147.
- Bayrak, H. (2020). *TÜİK hanehalkı bilişim teknolojileri kullanım araştırması*, 18 Ocak 2020 tarihinde <https://dijilopedi.com/2019-tuik-hanehalki-bilisim-teknolojileri-kullanim-arastirmasi/> adresinden alındı.
- Cole, D.A., & Maxwell, S.E. (2003). Testing mediational models with longitudinal data: Questions and tips in the use of structural equation modeling. *Journal of Abnormal Psychology*, 112(4), 558-577.
- Colesca, S. E. (2009). Understanding trust in e-government. *Engineering Economics*, 63(4), 7-15.
- Çokluk, Ö., Şekercioğlu, G., & Büyükoztürk, Ş. (2010). *Sosyal Bilimler için Çok Değişkenli İstatistik SPSS ve LISREL uygulamaları*. Ankara, Türkiye: Pegem Yayıncılık.
- Çolak, M. (2020). *Gizlilik ve Güvenlik*, 18 Ocak 2020 tarihinde <http://bilgisayarbilim.com/gizlilik-ve-guvenlik/> adresinden alındı.
- Davis, F. (1989). *A Technology acceptance model for empirically testing new end user information systems: theory and results* (Doctoral dissertation). MIT Sloan School of Management, Cambridge, MA.
- E-devlet Portalı, (2021). *Mobil Hizmet Sayısı*, 12 Ocak 2021 tarihinde <https://www.turkiye.gov.tr/> adresinden alındı.
- El-Kiki, T., & Lawrence, E. (2006). Mobile user satisfaction and usage analysis model of m-government services, *In Proceedings of EURO mGOV* (pp. 91-102). Brington, United Kingdom.
- Fornell, C., & Larcker, D. F. (1981). Evaluating structural equation models with unobservable variables and measurement error, *Journal of Marketing Research*, 18(1), 39-50.
- Gao, S., Krogstie, J., & Gransæther, P.A. (2008). Mobile services acceptance model. *2008 International Conference on Convergence and Hybrid Information Technology*, IEEE (pp. 446-453). Daejeon, Korea (South).
- Güler, M., & Dövertaş, E. (2009). Elektronik devletten (e-devlet) mobil devlete (m-devlet) geçişte Türkiye'de yerel yönetim uygulamaları. *Hitit Üniversitesi Sosyal Bilimler Enstitüsü Dergisi*, 2(1), 25-49.
- Gürses, F., & Engin, M. (2016). Türkiye'de yerel yönetimlerde mobil devlet uygulamaları: büyükşehir belediyeleri üzerine ampirik bir araştırma. *Süleyman Demirel Üniversitesi İktisadi ve İdari Bilimler Fakültesi Dergisi*, 21(1), 223- 234.
- Hung, S.Y.; Chang, C.M., & Kuo, S. R. (2013). User acceptance of mobile e-government services: An empirical study. *Government Information Quarterly*, 30(1), 33-44.
- Ishengoma, F., Mselle, L., & Mongi, H. (2019). Critical success factors for m-government adoption in Tanzania: A conceptual framework. *The Electronic Journal of Information Systems in Developing Countries*, 85(1), e12064.
- Khasawneh, R.T., Rabayah, W. A., & Abu-Shanab, E. A. (2013). E-government acceptance factors: trust and risk, *In The 6th International Conference On Information Technology (ICIT 2013)* (pp. 5-10). Amman, Jordan.

- Koivumäki, T., Ristola, A., & Kesti, M. (2008). The effects of information quality of mobile information services on user satisfaction and service acceptance—empirical evidence from Finland, *Behaviour ve Information Technology*, 27(5), 375-385.
- Kushchu, I., & Kusu, H. (2003). From e-government to m-government: facing the inevitable. *In the 3rd European Conference on e-Government* (pp.253-260). MCIL Trinity College Dublin, Ireland.
- Nunnally, J. C., & Bernstein, I. H. (1994). The assessment of reliability. *Psychometric Theory*, 3(1), 248-292.
- Önal, M. F. (2020). *E-devlet kapısı istatistikleri ve değerlendirmeler*, 20 Ocak 2020 tarihinde <http://www.egovturkey.com/e-devlet-kapisi-istatistikleri-ve-degerlendirmeler-04032020> adresinden alındı.
- Öztürkcan, S., Kasap, N., & Eryarsoy, E. (2012). M-devlet kullanıcı kabul potansiyeli: kümeleme analizi ve karar ağacı yaklaşımı. *Cumhuriyet Üniversitesi İktisadi ve İdari Bilimler Dergisi*, 13(2), 87-111.
- Rana, N., Janssen, M., Sahu, G.P., Baabdullah, A., & Dwivedi, Y. (2019). Citizens' perception about m-government services: Results from an exploratory survey. *In Proceedings of the 52nd Hawaii International Conference on System Sciences* (pp. 3356-3365). Hawaii, USA.
- Rehman, M., Kamal, M., & Esichaikul, V. (2012). Determinants of trust in e-government adoption: a case study of Pakistan. *Proceedings of the Eighteenth Americas Conference on Information Systems* (pp.1-12). Seattle, Washington, USA.
- Roubah, K., Qurban, H., Al-Qirim, N., & Tarhini, A. (2018). Understanding mobile government success in an arab country: findings from a qualitative study. *Issues in Information Systems*, 19(2), 185-198.
- Sandy, G.A., & McMillan, S. (2005). A success factors model for m-government. *EURO mGOV* (pp. 349-358). Brighton, England.
- Saxena, S. (2017). Enhancing ICT infrastructure in public services: factors influencing mobile government (m-government) adoption in India. *The Bottom Line*, 30(4), 279-296.
- Shareef, M. A., Archer, N., & Dwivedi, Y.K. (2012). Examining adoption behavior of mobile government. *Journal of Computer Information Systems*, 53(2), 39-49.
- Tabachnick, B.G., Fidell, L.S., & Osterlind, S. J. (2007). *Using Multivariate Statistics*. Boston, MA, USA: Pearson Publishing.
- Tan, G. W. H., Ooi, K.B., Sim, J.J., & Phusavat, K. (2012). Determinants of mobile learning adoption: an empirical analysis. *Journal of Computer Information Systems*, 52(3), 82-91.
- Üste, R. B., & Güzel, B. (2012). From e-governance through m-governance: Turkey applications. *International Journal of eBusiness and eGovernment Studies*, 4(1), 47-58.
- Veeramootoo, N., Nunkoo, R., & Dwivedi, Y.K. (2018). What determines success of an e-government service? Validation of an integrative model of e-filing continuance usage. *Government Information Quarterly*, 35(2), 161-174.
- Venkatesh, V., & F.D. Davis, F.D. (2000). A theoretical extension of the technology acceptance model: Four longitudinal field studies. *Management Sciences*, 46(2), 186-204.
- Venkatesh, V., Morris, M.G., Davis, G.B., & Davis, F.D. (2003). User acceptance of information technology: Toward a unified view. *MIS Quarterly*, 27(3), 423-478.
- Wang, H.J., & Lo, J. (2013). Determinants of citizens' intent to use government websites in Taiwan. *Information Development*, 29(2), 123-137.
- Wang, C. (2014). Antecedents and consequences of perceived value in mobile government continuance use: An empirical research in China. *Computers in Human Behavior*, 34, 140-147.
- Wang, Y.S. (2008). Assessing e-commerce systems success: a respecification and validation of the Delone and Mclean model of IS success. *Information Systems Journal*, 18(5), 529-557.
- Yıldırım, F., & Panayırıcı, U.C. (2012). Governmental mobile technology usages during promotion efforts of reaching their citizens. *International Journal of eBusiness and eGovernment Studies*, 4(1), 35-43.

Tedarikçilerin TOPSIS ile Seçilmesi ve Gelişim Durumlarının Sezgisel Optimizasyon ile Belirlenmesi*

Selecting the Suppliers with TOPSIS and Determining Their Development Status by Using Heuristic Optimization

Ceylan Oğuz¹ , İhsan Pençe² , Melike Şişeci Çeşmeli³ , Özlem Çetinkaya Bozkurt⁴ 



*Bu çalışma 9-11 Aralık 2020 tarihlerinde gerçekleştirilen 7. Uluslararası Yönetim Bilişim Sistemleri Konferansında sözlü olarak sunulan bildirinin genişletilmiş halidir.

¹(Yüksek Lisans Öğrencisi), Burdur Mehmet Akif Ersoy Üniversitesi, Bucak Zeliha Tolunay Uygulamalı Teknoloji ve İşletmecilik Yüksekokulu, Yönetim Bilişim Sistemleri Yüksek Lisans Programı, Burdur, Türkiye

²(Dr. Öğr. Üyesi), Burdur Mehmet Akif Ersoy Üniversitesi, Bucak Teknoloji Fakültesi, Yazılım Mühendisliği Bölümü, Burdur, Türkiye

³(Dr. Öğr. Üyesi), Burdur Mehmet Akif Ersoy Üniversitesi, Bucak Teknoloji Fakültesi, Yazılım Mühendisliği Bölümü, Burdur, Türkiye

⁴(Prof. Dr.), Burdur Mehmet Akif Ersoy Üniversitesi, Bucak İşletme Fakültesi, İşletme Bölümü, Burdur, Türkiye

ORCID: C.O. 0000-0002-1073-0766;
İ.P. 0000-0003-0734-3869;
M.Ş.Ç. 0000-0001-9541-2590;
Ö.Ç.B. 0000-0002-6218-2570

Corresponding author:

Ceylan OĞUZ
Burdur Mehmet Akif Üniversitesi, Bucak Zeliha Tolunay Uygulamalı Teknoloji ve İşletmecilik Yüksekokulu, Yönetim Bilişim Sistemleri Yüksek Lisans Programı, Burdur, Türkiye
E-mail address: ceylanoguz070@gmail.com

Submitted: 26.01.2021
Revision Requested: 17.02.2021
Last Revision Received: 30.03.2021
Accepted: 02.04.2021
Published Online: 31.05.2021

Citation: Oğuz, C., Pence, I., Sisece Ceshmeli, M. ve Cetinkaya Bozkurt, O. (2021). Tedarikçilerin TOPSIS ile Seçilmesi ve gelişim durumlarının sezgisel optimizasyon ile belirlenmesi.

Acta Infologica, 5(1), 53-64.
https://doi.org/10.26650/acin.868427

ÖZ

Küreselleşmenin etkisiyle Türkiye’de üretim yapan işletmelerin sayısı gün geçtikçe artmaktadır. Bu durum perakende satış yapan işletmelerin seçebileceği tedarikçi sayısında artışlara neden olmaktadır. Ancak bazı işletmeler bulunduğu bölgede satışını yaptıkları ürünün markası ile tanındıkları için tedarikçi değişikliğinden kaçınmaktadır. İlgili markanın tedarikçisi ile arasındaki işbirliği bozulursa işletmenin satışlarında da düşüşler yaşanabilmekte olup, tedarikçi değişikliğine sıcak bakılmamaktadır. İşletme tedarikçisini değiştiremediği için tedarikçisinin kendisini geliştirmesini istemektedir. Bu çalışmada Burdur ilinde ev tekstili ürünleri satışı yapan bir işletme ele alınarak, işletmenin işbirliği yaptığı tedarikçiler arasından en iyisinin belirlenmesi ve diğer tedarikçilerin de kendilerini ne ölçüde geliştirmesi gerektiği problemi üzerinde durulmuştur. Analitik Hiyerarşi Süreci ve TOPSIS kullanılarak en iyi tedarikçi belirlendikten sonra diğer tedarikçilerin hangi ölçüde kendilerini geliştirmeleri gerektiği ise sezgisel optimizasyon ile belirlenmiştir. Toplamdaki 11 kriter belirlenirken 8 adet tedarikçi firma görüşü ile birlikte literatürden de yararlanılmıştır. Çalışmanın sonucunda ağırlığı en yüksek olan kriter “ürünün kalitesi” olarak belirlenirken en iyi tedarikçinin tespiti ve diğer tedarikçilerin optimum durumu da başarı ile tespit edilebilmiştir.

Anahtar kelimeler: Analitik Hiyerarşi Süreci, TOPSIS, Sezgisel Optimizasyon, Tedarikçi İyileştirmesi

ABSTRACT

The number of enterprises engaged in production in Turkey is increasing day by day with the impact of globalization. This situation causes an increase in the number of suppliers that retailers can choose. However, some businesses avoid changing suppliers because they are known for the brand of the product they sell in their region. If cooperation between the suppliers of the relevant brand is disrupted, the enterprise’s sales may also decrease and the change of the supplier is not looked upon favorably. Since the business cannot change its supplier, it requests that its supplier improve itself. In this study, which focuses on an enterprise selling home textile products in the province of Burdur, the problems of determining the best among the suppliers with which the enterprise cooperates and the extent to which other suppliers should improve themselves are emphasized. After determining the best supplier using the Analytical Hierarchy Process and TOPSIS, the extent to which other suppliers should improve themselves is determined by heuristic optimization. The literature and the opinions of eight suppliers are used to determine 11 criteria in total. As a result of the study, while the criterion with the highest weight is determined to be “quality of the product,” the determination of the best supplier and the optimum status of other suppliers are also successfully determined.

Keywords: Analytical Hierarchy Process, TOPSIS, Heuristic Optimization, Supplier Improvement

1.GİRİŞ

Tekstil ürünleri insanoğlunun varoluşundan itibaren önemli bir ihtiyaç olmuştur. Eski zamanlara bakıldığında tekstil giyim olarak başlangıç gösterir ancak zamanla değişen ihtiyaçlar doğrultusunda ev tekstili alanı oluşmuştur. Ev tekstilinde farklı özelliklerde ve çok sayıda ürün olması ve müşterilerin markalaşmış ürünleri tercih etmesi sebebiyle satıcının tedarikçi seçiminde zorlandığı durumlar ortaya çıkmaktadır. Tedarikçi seçimi bir işletme için karar verilmesi zor olan konulardan bir tanesidir. Eğer işletme tedarikçi seçiminde hata yaparsa bunun sonucunda satışlarda azalma, kâr oranında düşüşler, müşteri memnuniyetinde azalış ve maliyetlerde artış olma olasılığı oldukça yüksektir. Tedarikçi seçim problemi gerekli olan ürün veya hizmetleri sağlayabilecek kapasitede olan tedarikçiler arasından hangilerinin seçileceği sorundur (Özdemir, 2007). Günümüzde firmalar ile tedarikçileri arasındaki ilişkiler değişmiş olup, artık alıcı ile satıcı firmalar aralarındaki ilişkileri güçlendirme yolunda ilerlemektedir. Tedarikçi seçimine etki eden kriterler alıcı ve satıcı arasındaki ilişkinin kazan-kaybet modelinden kazan-kazan modeline dönüşmesine katkıda bulunmaktadır (Vatansever, 2013). Bu noktada da işletmeyi en uygun tedarikçi veya optimum sonuca yönlendirecek AHP (Analitik Hiyerarşi Süreci), TOPSIS (Technique for Order Preference by Similarity to Ideal Solution) ve optimizasyon gibi çözümler gerekmektedir.

Günümüzde tedarikçi seçimi konusu, son yıllarda kapsamlı bir şekilde incelenmekte olan popüler bir sorun haline gelmiştir. Chai vd., (2013) karar verme tekniklerine dayalı olarak 2008-2012 yılları arasında yayınlanan tedarikçi seçim problemi hakkında 123 makaleyi inceledikten sonra tedarikçileri seçmek için kullanılan metodolojileri üç grupta sınıflandırmışlardır. İncelenen bu makalelere göre tedarikçi seçimi konusunda kullanılan yöntemlerin başında %44,4 ile AHP, ANP (Analitik Ağ Süreci) ve TOPSIS yöntemlerinin geldiği görülmektedir. Dolayısıyla literatürde tedarikçi seçiminde AHP ve TOPSIS yöntemlerinin birlikte kullanıldığı pek çok çalışma bulunmaktadır. Örneğin Wang vd., (2009), tedarikçi seçimi probleminde bulanık AHP ve TOPSIS yöntemlerini birlikte kullanmışlardır. Benzer şekilde Fazlollahtabar vd., (2011), araştırmalarında tedarikçi seçiminde AHP, TOPSIS ve çok amaçlı lineer olmayan programlamayı kullandıkları görülmektedir. Pramanik vd., (2017) makalelerinde, esnek bir tedarikçi seçme süreci geliştirmek için bulanık-çok kriterli bir grup karar verme yaklaşımı tasarlanmışlardır. Önerilen yöntemde, mesafe tabanlı optimizasyon metodolojisi, yani bulanık sistem ile entegre TOPSIS genel seçim kriterlerinin özelliklerini belirlemektedir. Benzer şekilde Günay ve Ünal (2016) telekomünikasyon sektöründeki bir firmayı ele almışlardır. Literatür taramasından ve uzman görüşlerinden yararlanarak belirlenen kriterleri AHP yöntemi ile önem derecesine göre sıralayıp, TOPSIS yöntemi ile de dört tedarikçi firmadan en uygun olanı belirlemişlerdir. Ağırlığı en yüksek olan kriter ‘ürün kalitesi’ kriteri olarak belirlenmiştir. Özçakar ve Demir (2011) ise bir gıda işletmesindeki karar vericilerle yapılan mülakatlar sonucunda edinilen bilgilerden yararlanarak yapılacak tedarikçi seçimini bulanık TOPSIS yöntemiyle gerçekleştirmişlerdir. Supçiller ve Çapraz (2011) bir işletmenin tedarikçi seçimini AHP-TOPSIS yöntemleriyle yapabildiği için model önermişlerdir. Önerilen modelde TOPSIS yöntemi ile tedarikçiler sıralanmış ve en iyi tedarikçi belirlenmiştir. Tedarikçilerin değerlendirilmesinde kullanılan kriterlerden kalite, fiyat, teslimat ve hizmet kriterleri ana kriterler olarak belirlenmiştir. İlgili çalışmada önem derecesi en yüksek kriter olarak ‘ürünün kalitesi’ kriteri belirlenmiştir. Vatansever (2013) tekstil sektöründe çalışan büyük işletmelere fason üretim yapan bir firmanın tedarikçi seçimini Bulanık TOPSIS yöntemiyle değerlendirmiştir. Öztürk (2019) Ordu ili Organize Sanayi Bölgesinde bulunan bir hazır giyim işletmesinin en uygun kumaş hammadde tedarikçisi seçiminde AHP-TOPSIS yöntemlerini kullanmıştır. Okwu ve Tartibu (2020) perakende sektöründe sürdürülebilir tedarikçi seçimi probleminin çözümü için TOPSIS ve ANFIS (Adaptive Neuro-Fuzzy Inference Systems) yöntemi kullanmışlardır. Avcı ve Çınaroğlu (2018), 5 havayolu firmasının 4 yıllık süre içerisinde göstermiş oldukları finansal performansları AHP ve TOPSIS yöntemlerini kullanarak sıralamışlardır. Şahin ve Akyer (2011) ise çalışmalarında kamu kurumlarında kullanılmak için alınacak olan 4×4 arama kurtarma aracı seçimi için AHP ve TOPSIS yöntemlerini kullanmışlardır. Literatürde sezgisel optimizasyon yöntemi ile ilgili çalışmalar da yapılmıştır. Akbel ve Kahraman (2020), çalışmalarında çok amaçlı sezgisel optimizasyon algoritmalarını test etmek için bir platform tasarlamışlardır. Literatürde bulunan 8 tane meta-sezgisel optimizasyon algoritmasından faydalanarak 24 tane birbirinden farklı problem üzerinde performans değerleri ölçülmüş ve performans karşılaştırmaları yapılmıştır. Kaya Keleş ve Keleş (2017) ise çalışmalarında literatürde son 5 yılda inşaat sektörünün temeli olan yapım yönetimi konusunda veri madenciliği ve sezgisel optimizasyon algoritmaları ile ilgili çalışmaları inceleyip ayrıntılı olarak tablo üzerinde kronolojik olarak göstermişlerdir. Niranjan vd., (2021), çalışmalarında farklı talep senaryolarında çoklu ürün akışı için yeşil tedarikçi seçimi ve sipariş tahsisi için bir model

önerip ilgili problemi yeni bir meta-sezgisel algoritma kullanarak çözüm getirmişlerdir. Özçelik ve Gündüz (2019) araştırmalarında bir otomobil fabrikasında otomobil montajında kullanılan parçaların raflardaki yerleşimini optimum şekilde kullanılmasını sağlayacak sezgisel algoritmalar kullanarak en iyiye yakın raf yerleştirme işlemini gerçekleştirilmiştir. Fırat ve Alpaslan (2019) ise bir işletmede üretim sürecindeki kesme ve paketleme sorununa ilişkin alt sol dolgu, ilk uygun azalan ve uygun olmayan çokgen sezgisel algoritmaları kullanarak iki boyutlu dikdörtgen şerit paketleme problemine çözüm sunulmaya çalışılmıştır. Tempelmeier (2002) araştırmasında endüstriyel satın alma uygulamasıyla ilgili varsayımlar altında dinamik sipariş boyutlandırma ve tedarikçi seçimi problemi için yeni model formülasyonları ve sezgisel bir çözüm yöntemi geliştirmiştir.

Görüldüğü üzere literatürde AHP ve TOPSIS yöntemlerinin tedarikçi seçimi probleminin çözümü için oldukça yaygın bir şekilde kullanıldıkları görülmektedir. Ancak AHP ve TOPSIS ile tedarikçi seçimi yapılan çalışmalarda tedarikçilerin birbirleri ile durumları hakkında bir değerlendirme yapılmayıp sadece en iyi tedarikçinin belirlenmesine odaklanılmıştır. Bu durum tedarikçilerin işletmeler için nasıl tercih edilebilir olabilecekleri ile ilgili öneriler sunmamaktadır. Bu çalışmada ise literatürde yer alan diğer çalışmalardan farklı olarak AHP ve TOPSIS yöntemleri ile en iyi tedarikçi belirlendikten sonra diğer tedarikçilerin kendilerini en iyi tedarikçiye göre ne ölçüde geliştirmeleri gerektiği problemi sezgisel optimizasyon kullanılarak gerçekleştirilmiştir. Bunun yanında tedarikçilerin birbirleri ile kıyaslamalarının yapılması ve araştırmanın ev tekstili sektöründe olması da ilgili çalışmayı diğer çalışmalardan farklı kılmaktadır. Sonuç olarak çalışmada, Burdur ilinde ev tekstil ürünleri satışı yapan bir firmanın tedarikçi seçimi sorunu ele alınmıştır. Amaç işletmenin çalıştığı tedarikçileri arasından en iyiyi belirleyip diğer tedarikçileri de en iyi tedarikçi olma yolunda yönlendirerek hem tedarikçilere hem de işletmeye fayda sağlayacak adımlar atılmasını sağlamaktır. İşletmenin tedarikçi seçiminde işletme sahibi ile görüşülüp tedarikçi seçiminde kullanılan kriterler belirlenip derecelendirilmiştir. Kriterler ve ağırlıklar belirlendikten sonra en uygun değer durum ise sezgisel optimizasyon ile tespit edilmiştir.

2. YÖNTEM

Bu çalışmada en uygun tedarikçinin belirlenmesinde çok kriterli karar verme modellerinden olup, kriter ağırlıklandırma da kullanılan AHP ve sıralama yapılırken kullanılan TOPSIS yöntemlerinden faydalanılırken, optimum kriter ve değerlerin belirlenmesi ise sezgisel bir yöntem olan Kümeleme Tabanlı Global Optimizasyon (KTGO) ile gerçekleştirilmiştir.

2.1. Analitik Hiyerarşi Süreci

Thomas L. Saaty (1980) tarafından geliştirilen ve çok kriterli karar verme yöntemlerinden biri olan AHP matematiksel bir yöntemdir. Bu yöntemin tercih edilmesindeki en büyük etken birden çok kriterin önem düzeylerinin belirlenmesine olanak sağlamasıdır (Saatly ve Özdemir, 2003). AHP karar veren kişilere öznel değerlendirmeleri nicel kriterlere dönüştürme fırsatı sunar.

AHP yönteminin ilk adımında modelin belirlenmesi işlemi bulunmakta olup, bu adım karar vericinin amacına yönelik kriterlerin, alt kriterlerin ve alternatiflerin belirlenmesi sürecini içermektedir. Alınacak olan kararın hatasız olması için belirlenen kriterler iyi seçilmelidir (Pençe vd., 2017). İkinci adımda, ikili karşılaştırma matrisinin oluşturulması aşaması bulunmakta olup belirlenmiş kriterler ve alternatifler kullanılarak kriterler ve alternatiflerin ikili karşılaştırma matrisleri 1-9 önem skalası ile oluşturulur. Üçüncü adımda ise, kriterlerin öz vektörlerinin hesaplanması aşaması yer alıp karar vericilerin tecrübeleri ile belirledikleri matriste yer alan kriterlere göre önem derecesini ifade eden öz vektör hesabı gerçekleştirilmektedir (Supçiller ve Çapraz, 2011). Dördüncü adımda, ikili karşılaştırmalarda kıyaslama yapılırken tutarlılık oranlarının hesaplanması yer almakta olup, tutarlılık oranının 0.10 değerinden küçük olması gerekmektedir. Bu oranın 0.10 değerinden daha büyük olması ya AHP'deki bir hesaplama hatasını ya da karar vericinin ikili karşılaştırmalarının tutarlı olmadığı anlamına gelmektedir (Saatly, 1980). Beşinci ve son adım ise, sonuç ağırlıkları bulunup genel alternatifler ile sıralama yapılmaktadır. İkili karşılaştırmalar matrisinde bulunan değerler birleştirilerek en alt seviyedeki alternatiflerin sonuç ağırlıkları bulunur. Elde edilen sonuçların sıralanmasıyla karar verici en iyi alternatifi seçebilir (Kerkhoff, 2018).

2.2. TOPSIS Yöntemi

TOPSIS yöntemi 1981'de Hwang ve Yoon tarafından geliştirilmiş olan çok kriterli karar verme yöntemlerinden biridir. Bu yöntem pozitif-ideal çözüme en yakın mesafedeki ve negatif-ideal çözüme en uzak mesafedeki alternatifleri bulmaya çalışır (Shyjith vd., 2008). TOPSIS yönteminin uygulama adımları şu şekildedir;

1. adım: karar matrisinin oluşturulması:

TOPSIS yönteminin ilk adımı, kriterlerin belirlenmesi ve sonrasında karar matrisinin oluşturulmasıdır. Karar matrisinin sütununda karar vermede kullanılacak değerlendirme faktörleri, satırında ise üstünlükleri sıralanmak istenen alternatifler yer almaktadır. Denklem (1)'de karar matrisini ifade eden A gösterilmektedir (Dumanoğlu ve Ergül, 2010).

$$A_{ij} = \begin{bmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \dots & a_{mn} \end{bmatrix} \quad (1)$$

2. adım: karar matrisinin normalleştirilmesi:

Oluşturulmuş olan karar matrisinin verilerinden yararlanılarak sütundaki her bir değer aynı sütundaki değerlerin kareleri toplamının kareköküne bölünmesiyle normalleştirme işlemi Denklem (2)'ye göre gerçekleştirilmektedir (Yayar ve Baykara, 2012).

$$r_{ij} = \frac{a_{ij}}{\sqrt{\sum_{i=1}^m a_{ij}^2}} \quad (2)$$

Denklem (2)'de yer alan i ; 1.2.3... n kriter sayısını ifade ederken, j ; 1.2.3... m ise alternatif sayısını göstermektedir.

3. adım: normalize edilmiş karar matrisinin oluşturulması:

İlk olarak değerlendirme faktörlerine bağlı ağırlık değerleri hesaplanır, sonrasında ise ikinci adımda oluşturulan matrisin her bir sütunundaki elemanlar bulunan ağırlık değeriyle çarpılarak normalize edilmiş karar matrisini ifade eden V , Denklem (3)'teki gibi belirlenir (Yayar ve Baykara, 2012).

$$V_{ij} = \begin{bmatrix} v_1 r_{11} & v_2 r_{12} & \dots & v_n r_{1n} \\ v_1 r_{21} & v_2 r_{22} & \dots & v_n r_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ v_1 r_{m1} & v_2 r_{m2} & \dots & v_n r_{mn} \end{bmatrix} \quad (3)$$

4. adım: pozitif ve negatif ideal çözümlerin oluşturulması:

İdeal çözüm, ağırlıklı normalleştirilmiş karar matrisinin en iyi performans değerlerinden oluşurken, negatif ideal çözüm ise, ağırlıklı normalleştirilmiş karar matrisinin en kötü değerlerinden oluşmaktadır. İdeal çözümler Denklem (4)'teki gibi elde edilmektedir (Shyjith vd., 2008).

$$\begin{aligned} A^+ &= \left\{ \left(\max_i v_{ij} \mid j \in J \right), \left(\min_i v_{ij} \mid j \in J' \right) \right\} \\ A^- &= \left\{ \left(\min_i v_{ij} \mid j \in J \right), \left(\max_i v_{ij} \mid j \in J' \right) \right\} \end{aligned} \quad (4)$$

5. adım: ayırım ölçülerinin hesaplanması:

Bu adımda alternatifler arasındaki ayırım hesaplanmaktadır. Pozitif ideal çözüm kümesinden sapma değeri hesaplanırken Denklem (5), negatif ideal çözüm kümesinden sapma değeri hesaplanırken ise Denklem (6) kullanılmaktadır (Ertuğrul ve Özçil, 2014).

$$S_i^+ = \sqrt{\sum_{j=1}^n (v_{ij} - v_j^+)^2} \quad (5)$$

$$S_i^- = \sqrt{\sum_{j=1}^n (v_{ij} - v_j^-)^2} \quad (6)$$

6. adım: ideal çözüme görelî yakınlığın hesaplanması:

Bu adımda alternatiflerin ideal çözüme görelî yakınlığını ifade eden C_i^* hesaplanmakta olup Denklem (7)' de verilmiştir. İlgili hesaplama yapılırken pozitif ve negatif ideal ayırım ölçülerinden faydalanılmaktadır. Negatif ideal ayırım ölçüsünün, toplam ayırım ölçüsü içindeki payı yakınlık katsayısı değerini vermektedir (Ertuğrul ve Özçil, 2014).

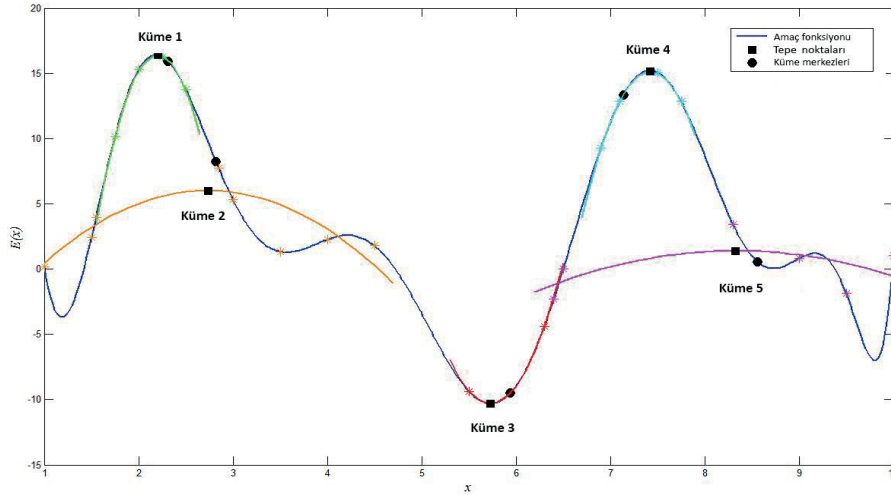
$$C_i^* = \frac{S_i^-}{S_i^- + S_i^+} \quad (7)$$

2.3. Sezgisel Optimizasyon

Optimizasyon, matematiksel olarak herhangi bir fonksiyonu veya problemi minimize veya maksimize etmektir. Optimizasyon temelinde her zaman en iyiye ulaşmayı amaçlamaktadır. Günümüze kadar farklı problemlerin çözümü için birçok optimizasyon yöntemi geliştirilmiştir (Çunkaş, 2004).

Yapay zekâ evrendeki varlıkların akıllı davranışlarını yapay olarak taklit eden bir süreç olup, bünyesinde kümeleme ve sınıflandırma problemleri için birçok algoritma barındırmaktadır (Atalay ve Çelik, 2017). Kümeleme analizi birbirine benzeyen verilerin aynı gruplarda olmasını sağlamaktadır (Çakmak vd., 2005). Kümeleme analizi veri madenciliğinin yanı sıra diğer alanlarda da kullanılmakta olup sezgisel optimizasyondaki kullanım alanında KTGO yöntemi ön plana çıkmaktadır. KTGO yöntemi popülasyon elemanlarını kümeleme işlemi sonucunda gruplayarak ilgili kümede yer alan elemanlara parabolik eğriler geçirmek suretiyle daha iyi noktalara ulaşmasıyla ön plana çıkan bir sezgisel optimizasyon yöntemidir (Pençe vd., 2016).

KTGO yönteminin en önemli adımlarından olan küme merkezlerinin belirlenmesi ve parabol uydurma işlemleri Şekil 1'de görülmektedir.



Şekil 1. KTGO yönteminin küme belirleme ve parabol uydurma işlemi (Pençe vd., 2016)

KTGO yöntemine ait algoritma ise aşağıda verilmiştir.

KTGO Algoritması:

1. Önceden belirlenen popülasyon büyüklüğüne göre ilk popülasyonu rastgele elemanlardan oluştur ve popülasyondaki her elemanın hata değerini hesapla.
2. Popülasyon-Hata oranı ve önceden belirlenen küme sayısına göre Bulanık C-Ortalamlar Kümeleme yöntemi ile kümeleme işlemi yap.

3. Fonksiyon değerlerine göre kümeleri sırala ve küme sayısının 1/3 oranında en iyi kümeleri seç.
4. Seçilen her bir kümeye ait elemanları kullanarak her boyut için ayrı ayrı parabolik eğri uydurma gerçekleştir ve oluşan eğrilerin iç bükey veya dış bükey olduğunu tespit et.
5. İç bükey eğrilerde minimum nokta hata değerinden daha küçük ise küme merkezini bu minimum nokta ile değiştir. Dış bükey eğriler için ise küme merkezini değiştirmeden kullan.
6. Önceki popülasyonun en iyi elemanları ile birlikte küme merkezi etrafından rastgele seçilen yeni elemanlarla yeni popülasyonu oluştur. Çıkarılan kümelere ait bireyler yerine yeni noktalar ekle.
7. Yeni popülasyondaki eleman değerlerini önceden belirlenmiş sınırlarda tut.
8. Yeni popülasyondaki her eleman için hata değeri hesapla.
9. Hata değeri önceden belirlenen eşik değerinden küçükse veya hedeflenen tekrara erişilmişse algoritmayı sonlandır değilse 2. adıma git.

Literatürdeki diğer bazı yöntemlere kıyasla KTGO yöntemi hızlı yapısı ve global optimuma ulaşmadaki başarısı ile ön plana çıkmaktadır (Pençe vd., 2016).

3. DENEYSEL ÇALIŞMALAR

Araştırma Burdur ilinde ev tekstili ürünleri satışı yapan bir işletmede gerçekleştirilmiştir. İşletmede havlu, bornoz, battaniye, perde, yastık, yorgan, halı gibi çok fazla ürün çeşidi bulunmaktadır. Bundan dolayı da işletme çok sayıda tedarikçi ile işbirliği içindedir. Yıllardır aynı tedarikçilerle çalışan işletme sahibi, tedarikçi firmalarla arasındaki özel ilişkilerden ve işletmenin o tedarikçinin markası ile anılmasından dolayı tedarikçisini değiştirmek istememektedir. Bunun yerine tedarikçilerine kendilerini geliştirmeleri ve ilişkilerinin daha karlı olabilmesi için yol göstermektedir. İşletme sahibinin tedarikçilerini doğru yönlendirebilmesi için daha bilimsel kararlar vermesi gerekmektedir. Bu çalışmada kullanılan yöntemler işletme sahibinin fikrini daha nesnel olarak belirlemesine yardımcı olacak düzeydedir. Çalışmada işletmenin en çok mal aldığı tedarikçiler olan 8 firma incelenmiştir. Bu firmalar Firma A, Firma B, Firma C, Firma D, Firma E, Firma F, Firma G, Firma H olarak gösterilmiştir. Tedarikçi firmaları değerlendirmek ve matematiksel bir model ortaya koyabilmek için işletmenin sahibi ve yöneticisi pozisyonunda olan kişiler ile birlikte ilgili kriterler belirlenmiştir. Kriterler belirlenirken literatürden de faydalanılmıştır. Supçiller ve Çapraz (2011), çalışmalarında gerçekleştirdikleri literatür taraması sonucunda en uygun tedarikçinin değerlendirilmesi ve seçilmesi için karar vericiler tarafından düşünülen kriterleri araştırmışlardır. Buna göre en popüler kriterler sırasıyla kalite, fiyat, teslimat, servis, esneklik, teknoloji, teknik, uzaklık vb. olarak belirlenmiştir. Bu çalışmadaki kriterler belirlenirken de işletmenin sahibi ve yöneticisi pozisyonunda olan kişilere bu kriterlerin işletme için uygun olup olmadığı önerilmiş olup, bu kriterler ile işletmeye özgü kriterler birlikte değerlendirilip belirlenmiştir. Bu çalışmada yararlanılan kriterler ve açıklamaları Tablo 1'de gösterilmektedir.

Tablo 1

Kriterler ve açıklamaları

Kriter No	Kriter Adı	Kriter Açıklaması
1	Ürün kalitesi (ÜK)	Alınan ürünün kalitesini göstermektedir.
2	Defosuz ürün miktarı (DÜM)	Alınan ürünün ne kadarının defosuz olduğunu göstermektedir.
3	Zamanında teslimat (ZT)	Teslimatın verilen süre içinde yapıp yapılmadığını göstermektedir.
4	Paketleme Yeteneği (PY)	Ürünün nasıl paklendiğini göstermektedir.
5	İletişim kolaylığı (İK)	Tedarikçi ile iletişimin nasıl olduğunu göstermektedir.
6	Sorunlara yaklaşım (SY)	Tedarikçinin sorunlara yaklaşımını göstermektedir.
7	Şikâyet politikaları (ŞP)	Tedarikçinin şikâyetler karşısında tutumunu göstermektedir.
8	Ürün-fiyat uygunluğu (ÜF)	Ürünün satış fiyatına uygunluğunu göstermektedir.
9	İskonto miktarı (İM)	Tedarikçinin sağladığı iskonto miktarını göstermektedir.
10	Ödeme vadesi (ÖV)	Yapılan ödemelerin vadesini göstermektedir.
11	Ödeme şekli (ÖŞ)	Yapılan ödemelerin şeklini (kredi nakit) göstermektedir.

Tablo 1’de yer alan kriterlerden [kriter no: 1-10] olanlar Supçiller ve Çapraz (2011)’a ait çalışmada belirlenen kriterleri oluştururken ÖŞ kriteri ise işletme yöneticileri tarafından önerilip çalışmaya dahil edilmiştir. Çalışmanın amacı Tablo 1’de verilen kriterlerden yararlanılarak tedarikçi firmalar içindeki en iyi firmayı belirleyip diğer firmaların da bu firmaya göre kendilerini ne yönde geliştirmeleri gerektiğini tespit etmektir. Bunun için ise en iyi firma AHP ve TOPSIS yöntemleri ile tespit edilmiş, sonrasında ise en optimum gelişim özellikleri sezgisel bir optimizasyon yöntemi olan KTGO ile belirlenmiştir. İlgili optimizasyon geliştirilen tedarikçi firmalar için rekabet dinamiği ele alınmamıştır.

4. BULGULAR

4.1. AHP Yönteminin Uygulanması

AHP’nin ilk aşaması olan karar matrisi oluşturma adımında Tablo 1’de yer alan 11 adet kriter ve alternatifler kullanılmıştır. Karşılaştırma matrisinin oluşturulması için bu kriterlerin önem derecelerinin belirlenmesi gerekmektedir. Bunun için Saaty (1980) tarafından geliştirilen 1-9 önem skalası ölçeği kullanılmıştır. Karşılaştırma matrisi Tablo 2’de verilmiştir.

Tablo 2

Kriterlerin karşılaştırma matrisi

	ÜK	DÜM	ZT	PY	İK	SY	ŞP	ÜFU	İM	ÖV	ÖŞ
ÜK	1	3	4	5	6	6	6	3	4	5	5
DÜM	0.33	1	3	4	5	5	5	1	3	4	4
ZT	0.25	0.33	1	2	3	3	3	0.5	4	5	5
PY	0.2	0.25	0.5	1	2	3	3	0.25	0.33	0.5	0.5
İK	0.16	0.2	0.33	0.5	1	1	1	0.2	0.33	0.5	0.5
SY	0.16	0.2	0.33	0.33	1	1	1	0.2	0.33	0.5	0.5
ŞP	0.16	0.2	0.33	0.33	1	1	1	0.2	0.33	0.5	0.5
ÜFU	0.33	1	2	4	5	5	5	1	0.5	0.33	0.33
İM	0.25	0.33	0.25	3	3	3	3	2	1	2	2
ÖV	0.2	0.25	0.2	2	2	2	2	3	0.5	1	1
ÖŞ	0.2	0.25	0.2	2	2	2	2	3	0.5	1	1

Karşılaştırma matrisinden sonra kriterlerin ağırlıklarının belirlenmesi için normalize edilmiş karşılaştırma matrisi oluşturulmuştur. Normalize karşılaştırma matrisi için her sütun değeri aynı sütunun değeri toplamına bölünürken, ağırlıkların belirlenmesi için de normalize edilmiş matrisin satır değerlerinin ortalaması alınır. Kriterlerin ağırlıkları Tablo 3’deki gibi elde edilmiştir. Ağırlıklar belirlendikten sonra tedarikçilerin sıralamalarının oluşturulması için TOPSIS yöntemi kullanılırken tedarikçilerin kendilerini iyileştirmeleri içinde yine bu ağırlıklardan yararlanılmıştır.

Tablo 3

Kriterlerin ağırlıkları

Kriter	ÜK	DÜM	ZT	PY	İK	SY	ŞP	ÜFU	İM	ÖV	ÖŞ
Ağırlık	0.255	0.163	0.124	0.047	0.027	0.026	0.026	0.107	0.086	0.065	0.065

Tablo 3’te belirtilen kriterlerin ağırlıkları hesaplandıktan sonra tutarlılık analizi uygulanmıştır. Buna göre tutarlılık göstergesi 0.156 olarak hesaplanırken, bu değer random gösterge değeri olan 1.51’e bölünmesiyle tutarlılık oranı 0.10 olarak elde edilmiştir. Tutarlılık oranı 0.10’dan büyük olmadığı için ağırlıklar tutarlı kabul edilerek TOPSIS ve KTGO için kullanılmıştır.

4.2. TOPSIS Yönteminin Uygulanması

AHP ile belirlenen alternatifler ve kriterler kullanılarak karar matrisi oluşturulmuştur. Oluşturulan karar matrisinin satırlarında alternatifler yani tedarikçi firmalar yer alırken sütunlarında ise belirlenen kriterler yer almaktadır. Tedarikçi firmaların değerlendirildiği karar matrisi Tablo 4’te gösterilmiştir.

Tablo 4

Karar matrisi

	ÜK	DÜM	ZT	PY	İK	SY	ŞP	ÜFU	İM	ÖV	ÖŞ
A	8	8	7	7	8	8	8	5	7	5	6
B	8	7	7	8	7	7	7	7	5	6	6
C	7	7	5	7	8	8	8	7	8	7	8
D	9	8	8	8	8	8	7	4	6	4	5
E	7	6	5	7	7	7	7	8	7	6	6
F	8	8	5	7	8	8	8	7	7	6	6
G	7	6	5	9	7	7	7	8	6	6	6
H	8	7	7	7	8	8	8	7	7	6	8

Tablo 4'te yer alan karar matrisinin normalleştirme işlemi ise sütundaki her bir değer ilgili sütunun değerlerinin kareleri toplamının kareköküne bölünmesi ile yapılmaktadır. Elde edilen değerlerin, AHP ile Tablo 3'teki gibi tespit edilen kriter ağırlıkları ile çarpılması ile de ağırlıklandırılmış karar matrisi bulunmuştur. Son olarak ağırlıklar kullanılarak tedarikçi firmalar için görelî yakınlıklar Tablo 5'teki gibi elde edilmiştir.

Tablo 5

Görelî yakınlıklar

Firma	A	B	C	D	E	F	G	H
Yakınlık	0.52685	0.54295	0.44863	0.56199	0.40122	0.51681	0.39068	0.62103

Tablo 5'te de görüldüğü gibi, elde edilen sonuçlara göre en iyi firma H olarak belirlenmiş olup, bu firma optimizasyon için en optimum değer kabul edilip en az maliyetle diğer firmaların çabaları KTGO yöntemiyle optimize edilmeye çalışılmıştır.

4.3. Optimizasyon Yönteminin Uygulanması

KTGO yöntemi ile TOPSIS'te bulunan en iyi tedarikçiyi temel alarak diğer tedarikçilerin en iyiye olan uzaklığı bulunmuştur. Burada yer alan karar matrisinde ilgili kriterin alt ve üst sınırları kullanılmıştır. Bunun sebebi tedarikçilerin aynı kriteri dönem dönem farklı seviyelerde sağlayabilmeleridir. Kriterlerin etkinlik durumu mevsim değişikliklerinden, enflasyondan ve çalışanların davranışlarından etkilenmektedir. İskonto miktarı, teslimat zamanı ve ödeme vadesi gibi sık sık değişebilen kriterler için firmaları puanlandırmak çok doğru bir değerlendirme kriteri olarak düşünülmemektedir. Bu sebeple 1-9 önem ölçüğü yerine işletme sahibinin kişisel görüşlerinden yararlanılarak alt ve üst sınırlar belirlenmiş ve Tablo 6'da gösterilmiştir.

Tablo 6

Optimizasyon karar matrisi

	ÜK	DÜM	ZT	PY	İK	SY	ŞP	ÜFU	İM	ÖV	ÖŞ
A	7-9	7-9	7-8	6-8	8-9	8-9	7-9	4-6	6-8	4-6	5-7
B	7-9	6-8	7-8	7-9	6-8	6-8	6-8	6-8	4-6	6-7	5-7
C	6-8	6-8	4-6	6-8	8-9	8-9	7-9	6-8	7-9	7-8	7-9
D	8-10	8-9	7-9	7-9	8-9	8-9	6-8	3-5	5-7	3-5	4-6
E	6-8	5-7	4-6	6-8	6-8	6-8	6-8	7-9	6-8	6-7	5-7
F	7-9	7-9	4-6	6-8	8-9	8-9	7-9	6-8	6-8	6-7	5-7
G	6-8	5-7	4-6	7-9	6-8	6-8	6-8	7-9	5-7	6-7	5-7
H	7-9	6-8	7-8	6-8	8-9	8-9	7-9	6-8	6-8	6-7	7-9

Tablo 6'da yer alan optimizasyon için oluşturulan karar matrisinden ve Tablo 3'te belirtilen kriter ağırlıklarından yararlanılarak firmaların kendilerini iyileştirme miktarlarını tespit edilebilecek matematiksel bir model Denklem (8)'de ki gibi oluşturulup, KTGO ile de optimize edilmiştir.

$$\max f(x_1, x_2, \dots, x_{11}) = 0.255x_1 + 0.163x_2 + 0.124x_3 + 0.047x_4 + 0.027x_5 + 0.026x_6 + 0.026x_7 + 0.107x_8 + 0.086x_9 + 0.065x_{10} + 0.065x_{11} \quad (8)$$

Denklem (8)'de yer alan x_i 'ler belirlenen kriterleri ifade etmektedir. Firmaların bir alanda iyileştirme miktarlarının gerçeği yansıtması ve diğer firmaların ortalama değerlerinin altında kalmaması için optimizasyon problemi için amaç fonksiyonunun kısıtları Denklem (9)'da ki gibi belirlenmiştir.

$$\text{s.t. } x_3 + x_5 + x_6 + x_7 \geq 28.875 \quad (9)$$

$$\frac{x_1 + x_2 + x_4}{3} \leq \frac{x_8 + x_9 + x_{10} + x_{11}}{4}$$

Kısıtlar arasında yer alan 28.875 değeri, $x_3 + x_5 + x_6 + x_7$ kriterlerinin her bir sütun toplamının 8 tedarikçi firmaya ortalamasına göre belirlenmiştir. İlgili kısıtlar belirlenirken kriterler aralarındaki uyum durumuna göre kategorilere ayrılmıştır. “Ürünün kalitesi, defosuz ürün miktarı ve paketleme yeteneği” birinci kategori, “ürün fiyat uygunluğu, iskonto miktarı, ödeme vadesi ve ödeme şekli” ikinci kategori, “zamanında teslimat, iletişim kolaylığı, sorunlara yaklaşım ve şikâyet politikaları” da üçüncü kategoriye oluşturmaktadır. Ayrıca tedarikçi bir firmanın belirlenen kriterlerde yapacağı değişimin 2’den küçük olması sınır değer olarak önerilmektedir. Bunun sebebi de işletmede yapılacak büyük değişikliklerin devasa maliyetlere neden olmasıdır. Burada değişimi asıl isteyen tedarikçi değil mal sattığı perakendeci olup büyük değişiklikler yerine daha makul ve kolay değişiklikleri uygulama eğiliminde olduğundan, kriter değişiminin en fazla 2 olması daha uygundur.

TOPSIS yöntemi kullanılarak bulunan en iyi tedarikçi olan Firma H ile diğer tedarikçilerin kıyaslaması KTGO ile elde edilen sonuçlara göre yapılmış olup tedarikçiler için elde edilen en uygun değerler Tablo 7’de görülmektedir.

Tablo 7

KTGO ile elde edilen sonuçlar

KRİTERLER	Firma A	Firma B	Firma C	Firma D	Firma E	Firma F	Firma G
Ürünün Kalitesi	8,23	8,53	8,27	9,02	10,00	8,31	10,00
Defosuz Ürün Miktarı	8,03	7,26	7,03	8,00	9,00	8,09	9,00
Zamanında Teslimat	7,95	7,94	5,08	8,03	9,00	5,41	9,00
Paketleme Yeteneği	7,00	8,03	7,29	8,03	9,00	7,08	9,00
İletişim Kolaylığı	8,02	7,04	8,05	8,02	9,00	8,07	9,00
Sorunlara Yaklaşım	8,01	7,08	8,00	8,02	9,00	8,15	9,00
Şikâyet Politikaları	8,00	7,06	8,03	7,00	9,00	8,06	9,00
Ürün-Fiyat Uygunluğu	5,06	7,00	7,01	4,00	9,00	7,08	9,00
İskonto Miktarı	7,06	5,00	8,01	6,12	9,00	7,14	9,00
Ödeme Vadesi	5,02	6,01	7,01	4,03	8,00	6,05	8,00
Ödeme Şekli	6,50	6,00	8,00	6,66	9,00	6,43	9,00

Tablo 7’de KTGO ile elde edilen sonuçlar ile Tablo 4’teki değerler karşılaştırıldığında; Firma A için “zamanında teslimat” kriterini neredeyse 1 puan geliştirmesi gerektiği ve en büyük yatırımı bu alana yapması sonucuna ulaşılmıştır. Sonrasında ise “ödeme şekli” kriterine yatırım yapması önerilmiştir. Diğer kriterlerden olan “paketleme yeteneği” ve “şikâyet politikaları” kriterlerinde herhangi bir iyileştirmeye ihtiyaç duyulmamıştır. Diğer kriterler için de küçük miktarlarda yatırımlar yapılmalıdır. Firma B için ise en fazla yatırım “zamanında teslimat” kriterine yapılmalı olup, bunu “ürünün kalitesi” kriteri takip etmektedir. “ürün fiyat uygunluğu”, “iskonto miktarı” ve “ödeme şekli” kriterlerinde bir değişim söz konusu değildir ancak diğer kriterlere de yatırım yapılması gerekmektedir. Firma C’ye bakıldığında “ürünün kalitesi” 1 puandan fazla arttırılmalıdır. “Sorunlara yaklaşım” ve “ödeme şekli” kriterlerinin arttırılmasına gerek duyulmamıştır. Geri kalan kriterlerde az miktarda değişime gidilmesi önerilmiştir. Firma D’nin ise “ödeme şekli” kriterini 1.5 puandan fazla arttırması ve “şikâyet politikaları” ile “ürün fiyat uygunluğu” kriterlerinde herhangi bir değişime gitmemesi önerilmiştir. Diğer kriterler içinde küçük yatırımlar yapılmalıdır. Firma E için ise, KTGO yöntemi yatırım yapılacak kriterlerde maksimum değerleri önermektedir. Sınır değer 2 olarak belirlenmesine rağmen bu firma için 2 puanlık bir yatırım yetersiz kalmaktadır. Daha köklü değişikliklere gidilmesi ilgili firma için önerilmektedir. Firma F’ye bakıldığında ise bir kriterde büyük miktarda yatırım yapmak yerine tüm kriterlere küçük miktarlarda yatırım yapılması önerilmiştir. Firma G’nin de Firma E gibi maksimum değerlerde yatırımlar yapması gerekip bu iki firma için küçük değişimler istenilen faydayı sağlayamamaktadır.

5. TARTIŞMA ve SONUÇ

Türkiye’de perakende satış yapan işletmelerin sayısı gün geçtikçe artış göstermektedir. Bu durum tedarikçi firmalar için zorlu bir rekabet ortamı yaratmaktadır. Bundan dolayı tedarikçi firmalar perakendeci işletmelerin isteklerini kolay kolay reddedemezler. Perakendeci işletmeler bu durumun farkında olmasına rağmen tedarikçisinden de ürünü maliyetine satmasını isteyemezler. Bunun için iki tarafın da tatmin olabileceği bir orta yol bulunup ilgili istekler makul ölçüde olmalıdır.

Bu çalışmada hem işletmenin hem de tedarikçinin optimum kâr elde edebilmesi için en uygun çözümü içeren, tedarikçilerin yatırım yapması gereken alanların optimizasyonu gerçekleştirilmiştir. İlgili çalışmada, Burdur ilinde perakende çalışan bir işletme ele alınmış olup bu işletmenin tedarikçilerinin yatırım alanları ve oranları bulunmuştur. Çalışmanın sonucunda ağırlığı en yüksek olan Firma H'nin muadili olarak Firma D gösterilebilmektedir. Ürün portföyü en geniş olan Firma D ele alınan tüm tedarikçilerin muadili konumundadır. A, B, C, D, G firmaları birbirinin muadili iken E, D ve F firmaları da kendi aralarında birbirinin muadilidir. Önerilen matematiksel model ve optimizasyon ile tedarikçiler kendilerini ne ölçüde geliştirmeleri gerektiğini belirleyebileceklerdir. Bu durum sonucunda hem tedarikçi firmalar hem de yönlendiren işletmeler en optimum kazancı elde etmiş olacaklardır. Tedarikçisini değiştirmek istemeyen işletmelerin önerilen modeli kullanarak tedarikçilerini geliştirmeleri önerilmektedir. Tedarikçilere ise ağırlığı en yüksek olan 'ürün kalitesi' konusunda kendilerini geliştirmeleri önerilmektedir.

Günay ve Ünal (2016)'a ait bir telekomünikasyon şirketinin hizmet alımı yaptığı dört tedarikçi firmanın sıralaması çalışmasında elde edilen ağırlığı en yüksek olan kriter "kalite" kriteri olarak bulunmuş olup, bu da bu çalışmada elde edilen sonuçla paralellik göstermektedir.

Çalışmanın kısıtı olarak, geliştirilen modelin tek bir işletme için oluşturulmuş olması gösterilebilir. Diğer firmaların modeli kullanabilmeleri kendi belirleyecekleri tedarikçilerini ve gerekli gördükleri kriterleri ele alarak modeli güncellemeleri gerekmektedir. İleriki çalışmalarda belirli bir sektörde yer alan birçok firma ele alınarak ortak bir model geliştirilebilir.

Hakem Değerlendirmesi: Dış bağımsız.

Çıkar Çatışması: Yazarlar çıkar çatışması bildirmemiştir.

Finansal Destek: Yazarlar bu çalışma için finansal destek almadığını beyan etmiştir.

Yazar Katkıları: Çalışma Konsepti/Tasarım- C.O., İ.P., M.Ş.Ç., Ö.Ç.B.; Veri Toplama- C.O.; Veri Analizi/Yorumlama-C.O., İ.P., M.Ş.Ç., Ö.Ç.B.; Yazı Taslağı- C.O., Ö.Ç.B.; İçeriğin Eleştirel İncelemesi-C.O., İ.P., M.Ş.Ç., Ö.Ç.B.; Son Onay ve Sorumluluk- C.O., İ.P., M.Ş.Ç., Ö.Ç.B.

Peer-review: Externally peer-reviewed.

Conflict of Interest: The authors have no conflict of interest to declare.

Grant Support: The authors declared that this study has received no financial support.

Author Contributions: Conception/Design of Study- C.O., İ.P., M.Ş.Ç., Ö.Ç.B.; Data Acquisition- C.O.; Data Analysis/Interpretation- C.O., İ.P., M.Ş.Ç., Ö.Ç.B.; Drafting Manuscript- C.O., Ö.Ç.B.; Critical Revision of Manuscript- C.O., İ.P., M.Ş.Ç., Ö.Ç.B.; Final Approval and Accountability- C.O., İ.P., M.Ş.Ç., Ö.Ç.B.

Kaynaklar/References

- Akbel, M. ve Kahraman, H.T. (2020). Çok Amaçlı Meta-Sezgisel Optimizasyon Algoritmalarının Performanslarının Karşılaştırılması. *Mühendislik Bilimleri ve Tasarım Dergisi*, 8(5), 185-199.
- Atalay, M., & Çelik, E. (2017). Büyük Veri Analizinde Yapay Zeka ve Makine Öğrenmesi Uygulamaları. *Mehmet Akif Ersoy Üniversitesi Sosyal Bilimler Enstitüsü Dergisi*, 9(22), 155-172.
- Avcı, T. ve Çınaroğlu, E. (2018). AHP Temelli TOPSIS Yaklaşımı İle Havayolu İşletmelerinin Finansal Performans Değerlemesi. *C.Ü. İktisadi ve İdari Bilimler Dergisi*, 19(1), 316-335.
- Chai, J., Liu, J. N., & Ngai, E. W. (2013). Application of decision-making techniques in supplier selection: A systematic review of literature. *Expert systems with applications*, 40(10), 3872-3885.
- Çakmak, Z., Uzgören, N., & Keçek, G. (2005). Kümeleme Analizi Teknikleri ile İllerin Kültürel Yapılarına Göre Sınıflandırılması ve Değişimlerinin İncelenmesi. *Dumlupınar Üniversitesi Sosyal Bilimler Dergisi*, 12, 15-36.
- Çunkaş M. (2004). Elektrik Makinalarının Genetik Algoritmayla Optimizasyonu. Doktora Tezi, Selçuk Üniversitesi, Fen Bilimleri Enstitüsü, Konya.
- Dumanoğlu, S., & Ergül, N. (2010). İMKB'de İşlem Gören Teknoloji Şirketlerinin Mali Performans Ölçümü. *Muhasebe ve Finansman Dergisi*, 48, 101-111.
- Ertuğrul, İ. & Özçil, A. (2014). "Çok Kriterli Karar Vermede TOPSIS ve VIKOR Yöntemleriyle Klima Seçimi", *Çankırı Karatekin Üniversitesi İktisadi ve İdari Bilimler Fakültesi Dergisi*, 4(1), 267-282.
- Fashoto, S. G., Akinnuwesi, B., Owolabi, O., & Adelekan, D. (2016). Decision support model for supplier selection in healthcare service delivery using analytical hierarchy process and artificial neural network. *African journal of business Management*, 10(9), 209-232.
- Fazlollahtabar, H., Mahdavi, I., Ashoori, M.T., Kaviani, S., Mahdavi-Amiri, N., (2011). A multi-objective decision-making process of supplier selection and order allocation for multi-period scheduling in an electronic market. *The International Journal of Advanced Manufacturing Technology*, 52, 1039-1052.
- Fırat, H., & Alpaslan, N. (2019). Sezgisel Algoritmalar Kullanılarak İki Boyutlu Dikdörtgen Şerit Paketleme Probleminin Çözümü. *Avrupa Bilim ve Teknoloji Dergisi*, (17), 315-322.

- Günay, Z., & Ünal, Ö. F. (2016). Ahp-Topsis Yöntemi İle Tedarikçi Seçimi (Bir telekomünikasyon şirketi örneği). *PESA Uluslararası Sosyal Araştırmalar Dergisi*, 2(1), 37-53.
- Hwang, C. L., & Yoon, K. (1981). Methods for multiple attribute decision making. In *Multiple attribute decision making*. Springer, Berlin, Heidelberg, 58-191.
- Kaya Keleş, M. ve Keleş, A. E. (2017). Veri Madenciliği Uygulamalarının ve Sezgisel Optimizasyon Algoritmalarının Yapım Yönetimindeki Yeri. *Çukurova Üniversitesi Mühendislik Mimarlık Fakültesi Dergisi*, 32(1), 235-242.
- Kerkhoff, E. (2018). Çok Kriterli Karar Verme Yöntemleriyle Tedarikçi Seçimi, Yayınlanmamış Yüksek Lisans Tezi, *İstanbul Ticaret Üniversitesi Sosyal Bilimler Enstitüsü*, İstanbul.
- Niranjan, T., Singaravel, B., & Raju, S. S. (2021). Integrated Fuzzy Criteria Evaluation with Metaheuristic Optimization for Green Supplier Selection and Order Allocation. In *IOP Conference Series: Materials Science and Engineering*, 1057(1), 012074.
- Okwu, M. O., & Tartibu, L. K. (2020). Sustainable supplier selection in the retail industry: A TOPSIS-and ANFIS-based evaluating methodology. *International Journal of Engineering Business Management*, 12, 1-14.
- Özçakar, N., & Demir, H.H. (2011). Bulanık TOPSIS Yöntemiyle Tedarikçi Seçimi. *İstanbul Üniversitesi İşletme Fakültesi İşletme İktisadi Enstitüsü Yönetim Dergisi*, 22(69), 25-44.
- Özçelik, T. Ö., & Gündüz, G. (2019). Sezgisel Algoritmaları Kullanarak Raf Optimizasyonu Çalışması ve Bir Yazılım Uygulaması. *Avrupa Bilim ve Teknoloji Dergisi*, (16), 977-982.
- Özdemir, A. (2007). Tedarikçi Seçiminde Karar Modelleri ve Bir Uygulama Denemesi. Doktora Tezi, *Uludağ Üniversitesi Sosyal Bilimler Enstitüsü*, 158s., Eskişehir.
- Öztürk, D. (2019). Selection of Suppliers Using AHP and TOPSIS Methods: An Application in the Ready-Made Clothing Sector. *Journal of Textiles and Engineer*, 26(115), 299-308.
- Pençe, İ., Tarhan, L., & Çetinkaya Bozkurt, Ö. (2017). Türk Eğitim Vakfı Bursu Verilecek Uygun Adayların AHP ve TOPSIS Yöntemi Kullanılarak Belirlenmesi: Mehmet Akif Ersoy Üniversitesi Eğitim Fakültesi Örneği. *Mehmet Akif Ersoy Üniversitesi Uygulamalı Bilimler Dergisi*, 1(1), 37-49.
- Pençe, İ., Şişeci Çeşmeli, M., Senel, F. A., & Cetisli, B. (2016). A new unconstrained global optimization method based on clustering and parabolic approximation. *Expert Systems with Applications*, 55, 493-507.
- Pramanik, D., Haldar, A., Mondal, S. C., Naskar, S. K., & Ray, A. (2017). Resilient supplier selection using AHP-TOPSIS-QFD under a fuzzy environment. *International Journal of Management Science and Engineering Management*, 12(1), 45-54.
- Saaty, T.L., & Özdemir, M.S. (2003). Negative Priorities in the Analytic Hierarchy Process. *Mathematical and Computer Modelling*, 37(9-10), 1063-1075.
- Saaty, T.L. (1980). *The Analytic Hierarchy Process*. McGraw-Hill International Book Company, New York.
- Shyjith, K., Ilangkumaran, M., & Kumanan, S., (2008). Multi-criteria Decision-making Approach to Evaluate Optimum Maintenance Strategy in Textile Industry. *Journal of Quality in Maintenance Engineering*, 14(4), 375-386.
- Supçiller, A., & Çapraz, O. (2011). AHP-TOPSIS Yöntemine Dayalı Tedarikçi Seçimi Uygulaması. *İstanbul Üniversitesi İktisat Fakültesi Ekonometri ve İstatistik Dergisi*, 13, 1-22.
- Şahin, Y. ve Akyer, H. (2011). Ülke Kaynaklarının Verimli Kullanımı: 4×4 Arama ve Kurtarma Aracı Seçiminde AHS ve TOPSIS Yöntemlerinin Uygulanması. *Süleyman Demirel Üniversitesi Vizyoner Dergisi*, 3(5), 72-87.
- Tempelmeier, H. (2002). A simple heuristic for dynamic order sizing and supplier selection with time-varying data. *Production and Operations Management*, 11(4), 499-515.
- Vatansever, K. (2013). Tedarikçi Seçim Kararlarında Bulanık TOPSIS Yönteminin Kullanımı ve Bir Uygulama. *Anadolu Üniversitesi Sosyal Bilimler Dergisi*, 13(3), 155-168.
- Wang, J.-W., Cheng, C.-H., Kun-Cheng, H., (2009). Fuzzy hierarchical TOPSIS for supplier selection. *Applied Soft Computing*, 9, 377-386.
- Yayar, R., & Baykara, H.V. (2012). TOPSIS Yöntemi İle Katılım Bankalarının Etkinliği ve Verimliliği Üzerine Bir Uygulama. *Business and Economics Research Journal*, 3(4), 21-42.

Information Technology Governance Practices, Challenges and Effects on Enterprise Performance*

Bilgi Teknolojileri Yönetişim Uygulamaları, Güçlükler ve Kurumsal Performansa Etkileri

Murat Paşa Uysal¹ , Evrim Çetinkaya² 



*This submission appeared as part of Evrim Çetinkaya's master thesis, titled "The Effects of IT Governance Practices on Key Performance Indicators of Businesses" with advisor Murat Paşa Uysal

¹(Prof. Dr.), Baskent University, Faculty of Commercial Sciences, Department of Management Information Systems, Ankara, Turkey

²Ufuk University, Institute of Social Sciences, Master of Business Administration Department, Ankara, Turkey

ORCID: M.P.U. 0000-0002-8349-9403
E.Ç. 0000-0003-0599-0077

Corresponding author:
Murat Paşa UYSAL
Baskent University, Faculty of Commercial Sciences, Department of Management Information Systems, Ankara, Turkey
E-mail address: mpuysal@baskent.edu.tr

Submitted: 27.11.2020
Revision Requested: 16.01.2021
Last Revision Received: 19.02.2021
Accepted: 23.02.2021
Published Online: 18.05.2021

Citation: Uysal, M. P., & Çetinkaya, E. (2021). Information technology governance practices, challenges and effects on enterprise performance. *Acta Infologica*, 5(1), 65-78.
<https://doi.org/10.26650/acin.832561>

ABSTRACT

Information Technology Governance (ITG) is important for enterprises in many respects. Acquiring better outputs from Information Technologies (IT) investments and meeting various expectations in business can be shown as primary examples. Another issue in this context is how an enterprise can manage its own IT resources according to its business strategy and functions. IT Governance Practices (ITGP), which is also described as best practices, are one of the main approaches that can be used to handle these issues. On the business side, Critical Success Factors (CSF) and Key Performance Indicators (KPI) are amongst the other tools used by enterprises for performance management. However, it is seen that there are not enough research studies exploring ITGP implementation challenges, their effects on CSF, KPI, and their relationships. In this paper, we present the findings of a qualitative research conducted according to the guidelines of Systematic Literature Review and Systematic Mapping Methods. The research outputs are promising though there are limitations that should be considered as well. Our study can be viewed as an initial attempt to the enhancement of ITGP and enterprise performance knowledge domains.

Keywords: IT Governance Practice, Enterprise Performance, Critical Success Factors, Key Performance Indicators

ÖZ

Bilgi Teknolojileri Yönetimi (BTY) birçok açıdan organizasyonlar için önemlidir. Bilişim Teknolojileri (BT) yatırımlarından daha iyi sonuç alınması ve işletmelerdeki çeşitli beklentilerin BT tarafından karşılanması bunlara örnek gösterilebilir. Bu kapsamdaki bir diğer konu ise işletmelerin BT uygulamalarına göre kendi bilişim teknolojilerini nasıl yönetebildikleridir. En iyi uygulamalar (best practices) olarak nitelendirilen Bilgi Teknolojileri Yönetim Uygulamaları (BTYU), bu konulara çözüm getirmek üzere kullanılan yaklaşımların başında gelmektedir. Öte yandan, Kritik Başarı Faktörleri (KBF) ve Anahtar Performans Göstergeleri (APG) ise işletmeler tarafından kurumsal performans yönetiminde kullanılan araçlar arasındadır. Ancak, alan yazın taraması sonucunda BTYU'nun uygulama güçlüklerini, KBF ve APG'ne olan etkilerini ve bunlar arasındaki ilişkileri inceleyen yeterli sayıda çalışmanın bulunmadığı gözlenmiştir. Söz konusu araştırma problemine yönelik olarak bu çalışmada Sistematik Haritalama ile Sistematik Literatür Taraması yöntemlerinin bütünlendirilmesiyle gerçekleştirilen araştırmaya ait bulgular sunulmuş, sonuçlar BTYU ve kurumsal temeller çerçevesinde tartışılmıştır. Araştırmamız BTYU'nun iyileştirilmesine yönelik bir girişim olarak nitelendirilebilir.

Anahtar kelimeler: Bilgi Teknolojileri Yönetimi, Kurumsal Performans, Kritik Başarı Faktörleri, Anahtar Performans Göstergeleri

1. INTRODUCTION

Enterprises have various goals and objectives, which are managed and organized with limited resources and usually dependent on internal and external factors (Khudyakova et al. 2020). The resources, such as, human, capital and materials, should be managed in an effective and efficient way in order to achieve strategic organizational goals. The ability to organize, plan, manage and control enterprise or business functions for achieving these goals is called enterprise management (Blokdyk, 2020). Marketing and sales, accounting and finance, production and manufacturing, human resources are the core business functions. In this context, Information Technologies (IT) and Information Systems (IS) have become the integral part of organizational activities, which acquire, record, transform and distribute data and information. Thus, IT and IS (ITIS) together add great value to enterprises by bringing innovative solutions to organizational problems (Rehm et al., 2015). It is possible to state that there is an increasing and strong dependency between IT and enterprises. Not only are key, core and important business processes conducted by ITIS, but also most of the daily processes are connected with digital networks, managed by digital solutions and software applications. Consequently, IT has become a change agent for transforming businesses (Ash and Burn, 2003).

Organizations usually invest in technology to develop or determine their ITIS-based strategies (Chen et al., 2011). It is hoped that managers can make better decisions, increase organizational performance as well as business profitability. ITIS has become an important instrument for realizing new products, services and business models (Arnold and Voigt, 2017). However, effective use of ITIS requires a concrete and sound understanding of organization, management and technology. Changes in strategy, organization, processes or operations may also require significant changes in IT, hardware, software and IT infrastructure (ISACA, 2020). IT has become so important for enterprises that IT Management (ITM) has already been regarded as an important business service (ITIL, 2021).

Information Technologies Management (ITM) is defined as planning and managing ITIS, collecting, processing, recording the information of businesses while transmitting them from one location to another when needed (AIS, 2020). It has a set of strategies, methods, tools and principles for the use of IT in an organization. In other words, ITM determines the authority and responsibilities in decision making processes related to ITIS. Increasing productivity, providing quality goods and services to customers, reducing costs, developing new products, assuring competitiveness of enterprises are among the important contributions provided by ITM. As a result, ITM has become one of the highest priorities for businesses and it is one of the major factors affecting the overall state of any business.

Technology investments are not only costly but also, they carry significant risks that would affect the organizational situation (Vincent et al., 2017; Okul et al., 2019). Therefore, it is of great importance to determine ITIS, not intuitively, but by following purposeful and comprehensive ITM strategies, methods and techniques. One of the concerns of managers and boards is naturally to expect a better return on investment (ROI), and therefore, they expect that ITIS will deliver what is needed to enhance the organization (NCC, 2005). As in all sectors of the industry, there are regulatory requirements and control frameworks for ITM to help enterprises manage, assess, monitor and improve their IT related processes (Barafort et al., 2017). Enterprises seek to optimize the costs by driving these standardized approaches. Moreover, they need to assess how they are performing against IT standards in view of organizational aspects (ITGI, 2020). Within this scope, it is necessary to determine, manage and control the risks arising from the use of IT. They have to provide effective and efficient information and knowledge management, which would meet business goals in an agile, continuous and secure manner (Al-Alawi et al., 2007). ITM should be based on solid foundations in order for an enterprise to manage personnel and technology efficiency, assure competitive advantage and better marketing position. Therefore, ITM has become important more than ever for improving enterprise performance.

Performance is defined as the ability or level that may be achieved by a person, group, organization or process. Performance Management (PM) provides a review of enterprise performance and contribution of business activities (Franceschini, 2007). It also verifies the achievement of strategic goals set by an organization. However, this is not an easy task. PM and control frameworks require the determination of certain areas and activities, which should also receive continuous attention from

the management (Parmenter, 2015). Therefore, various methods, techniques and tools exist for measuring enterprise performances. Critical Success Factors (CSF) and Key Performance Indicators (KPI) are well known techniques (Trkman, 2010). CSF are the areas, which may be the conditions, characteristics or variables that can have significant impacts on the success of an organization (Samsonowa, 2012). KPI are the performance measure(s) associated with CSF, which also link the activities and day-to-day organizational performances to their corresponding CSF.

In the same sense, effective IT performance management and measurement is also an essential and integral part of the organizational success. To that aim, IT Governance (ITG) practices provide organizations with the required means to assure that the investments in IT would generate expected business value. (Haes and Grembergen, 2009). These practices' main purpose is to develop and implement organizational strategies by defining well-defined roles for the responsibility of IT processes, data, information, application and technical infrastructure (ITGI, 2020). While aligning IT with business objectives, resource management, risk management and value delivery are the other key components of ITG practices. They ensure that "IT related decisions and management are made and driven by the business and not vice versa (NCC, 2005)." Therefore, there have been various attempts to use or adopt these practices, i.e., ITIL, COBIT, CMMI. This is widespread in IT businesses; however, it often requires significant investments and has important impacts on enterprise performances. To be more concise and formally representative, the aforementioned core concepts and our research claims are described by using the formal methods as follows:

$$\forall x \text{ CSF}(x_{1..n}) \rightarrow \exists x \text{ KPI}(x_{1..n}) \quad (1)$$

$$\text{ITG}(P_{1..n}) \rightarrow \text{CSF}(x_{1..n}) \quad (2)$$

$$\text{ITG}(P_{1..n}) \rightarrow \exists x \text{ KPI}(x_{1..n}) \quad (3)$$

- Equation (1) indicates that for all performance measures $x_{1..n}$ included in the CSF function, there exists one or more associated KPI;
- The equation (2) states that different types of ITG practices ($P_{1..n}$) should be in relation to the organizational CSF.
- Consequently, the equation (3) asserts that there should be one or more CSF-associated KPI that are also related with ITG practices.

However, it is possible to state that there is not much work that is researching the effects of ITG best practices on enterprise KPI and CSF (Haes and Grembergen, 2009). Therefore, the main purpose of this research is to explore the studies on ITG practices, implementation challenges and their effects on enterprise performances (Çetinkaya, 2017). The study is conducted by a systematic way to determine the research gaps and trends. A review protocol is developed after identifying the research questions. This protocol covers keywords to be used, databases to be scanned, search statements, paper selection criteria, methods and techniques for extracting and synthesizing data as well as research validity and reliability measures. The primary research findings are presented and they are discussed according to the theoretical background that exists in ITG and PM knowledge domains. The following sections contain the theoretical background, method, results and discussion and conclusion parts of this paper.

2. THEORETICAL BACKGROUND

2.1. COBIT (Control Objectives for Information and Related Technology)

COBIT is one of the most widely used best practices adopted as an IT management and control framework (ISACA, 2020). It plays a significant role in achieving the enterprises' goals, supporting IT and business objectives, designing IT activities using process models, determining IT resources and defining management control objectives. COBIT offers an IT management framework, a control and governance system developed by obtaining a large number of expert opinions (Figure 1). It models the IT processes that would ensure the management and control of IT resources necessary for enterprises to achieve their strategic goals. In COBIT, IT activities are conducted by IT processes that exist in four main domains. These are Align, Plan and Organize (APO), Build, Acquire and Implement (BAI), Deliver, Service and Support (DSS), Monitor, Evaluate and

Assess (MEA). There are a total of 35 processes in these domains. APO provides a plan and organization of IT solutions and services, BAI is responsible for the realization of this plan, DSS ensures that relevant IT solutions are implemented and made available for enterprise users, and finally, MEA provides monitored and controlled IT services aligned with the plan.

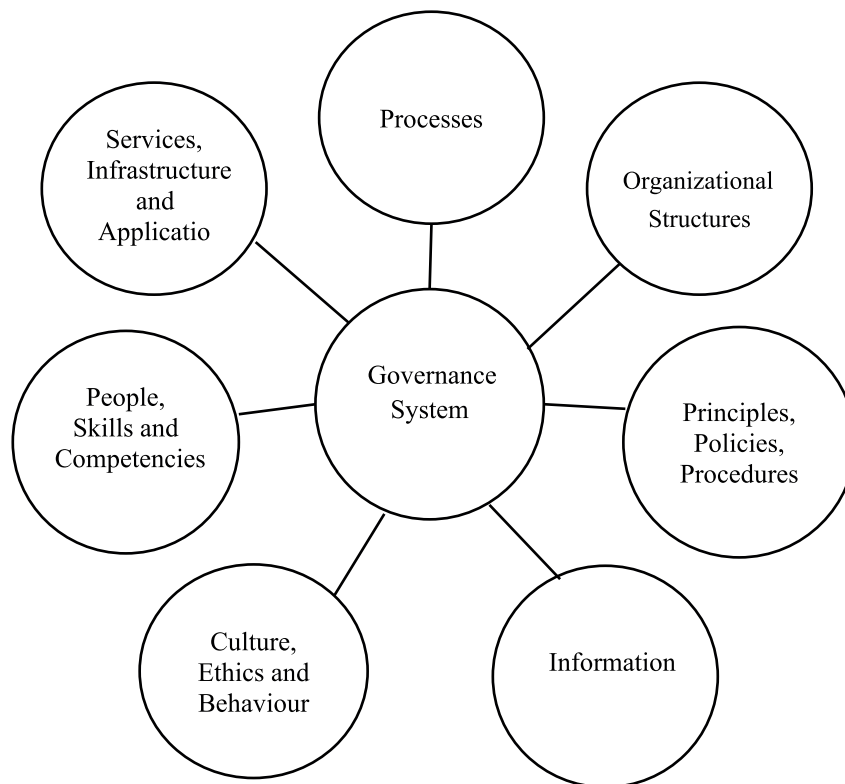


Figure 1. COBIT Components of a Governance System (ISACA, 2020)

2.2. ITIL (Information Technology Infrastructure Library)

ITIL is an IT service management practice, which is used to align business processes and objectives with IT management strategy that aims to provide efficient and effective IT services (ITIL, 2021). The bottom line and main purpose of IT service management in ITIL is to provide and support IT-based information services to meet business goals. It includes comprehensive descriptions of 26 processes arranged along a service lifecycle: Service Strategy (SS), Service Design (SD), Service Transition (ST), Service Operation (SO), Continual Service Improvement (CSI) respectively (Figure 2). SS determines the services that an IT organization is to offer and the capabilities needed for developing them. SD includes the design of new services as well as changes or improvements required for existing services. ST builds and deploys IT services while tries to make sure that the changes to services and their management processes are carried out in a harmonized way. SO is responsible for operational tasks and assuring that IT services are delivered efficiently and effectively. Finally, CSI improves the efficiency and effectiveness of processes and services and it uses the methods of quality management. It is known that ITIL and COBIT complement each other.

2.3. CMMI (Capability Maturity Model Integration)

CMMI is a process improvement framework developed by Software Engineering Institute (SEI) (CMMI, 2021). It includes three different core models; CMMI for Development (CMMI-Dev), CMMI for Acquisition (CMMI-Aqc) and CMMI for Services (CMMI-Svc). CMMI-Dev is used as a process model for management and improvement of processes in software development organizations. CMMI-Aqc is a model for organizations, which have to manage and initiate the acquisition of products or services. CMMI-Svc is, again, a process model for deploying and managing services. CMMI has two different representation models for process improvement: A staged representation and a continuous representation. The staged representation focuses on the overall organization and it provides a series of stages, which are also called as maturity levels

indicating the maturity of an organization's set of process areas. The continuous process model, however, focuses on an individual process area and how it could be improved. It is different from the staged model in that it provides less guidance on the order of implementation and improvement of process areas. The continuous representation of CMMI uses capability levels to measure the process improvement. These levels are 0 for incomplete, 1 for performed, 2 for managed, 3 for defined, 4 for quantitatively managed and 5 for optimizing. On the other hand, the maturity levels are applied to various sets of process areas. They are represented as 1 for initial, 2 for managed, 3 for defined, 4 for quantitatively managed, 5 for optimizing respectively.

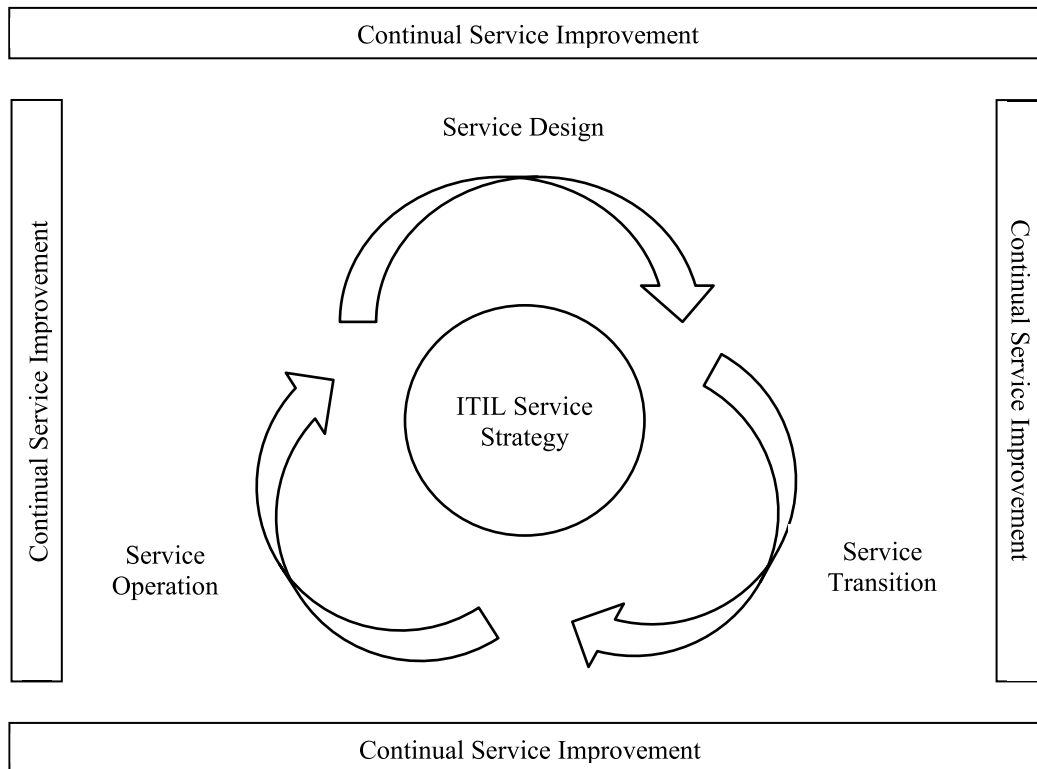


Figure 2. ITIL v-3 Service Lifecycle

2.4. IT Governance

Information Technologies Governance (ITG) is a discipline that performs IT process, risk and value chain management to ensure the effective and efficient use of IT resources in achieving strategic enterprise goals (ITGI, 2020). To this end, COBIT and ITIL are at the top of IT best practices used within the ITG knowledge domain. While COBIT sets out IT-related control objectives and performance metrics, ITIL describes what needs to be done to achieve these goals. A large part of the research related to ITG aims at the application of IT practices and standards to the organization and determining the skills and maturity levels related to IT processes (Haes and Grembergen, 2009). In some of these applications, the general situation of organizations regarding IT is determined, while in some others IT processes are explored. With the development of IT technologies, information security and risk analysis are among the popular topics. In these studies; IT management models and applications, performance, risk and security, strategic and organizational goals, IT performance measurements are examined at various enterprise levels.

2.5. Enterprise Performance, CSF and KPI

Today, enterprises have intensive IT or IS-dependent processes, however, the most of strategic and operational decisions are not fact-based and they are still made intuitively (Dimon, 2013). Thus, managers need performance management frameworks for analyzing, designing and implementing such systems as well as aligning ITIS with their organizational performances (Blokdyk, 2020). Therefore, CSF and KPI play an important role in this issue. CSF are the areas, which may be the conditions, characteristics or variables that can have significant impacts on the success of an organization (Samsonowa, 2012). The core

concept is based on the notion that a limited number of factors exist, which can make a significant contribution to any enterprise’s success. They are actually the few such key areas where “things must go right” for any business. Otherwise, business outcomes would be less than desired. In the IS domain, CSF are viewed as something that must happen if an IT-related project, service, process or activity is to succeed.

KPIs focus on the aspects of performance that are the most critical for both the current and future state of organizations and they measure what has happened in terms of organizational performance. They are actually the set of performance indicators agreed on and selected upfront by management as the most representative of r indicators (De Sousa, 2004). KPIs are also the measures that show how teams are performing daily and/or weekly. Thus, KPIs link the activities and day-to-day performance in the workplace to the organization’s CSF that are associated with one or more KPI. While contributing to the ownership of process improvement, they provide a basis for recognizing achievements, identifying problem areas and determining organizational priorities. Parmenter (2015) suggests that KPIs may be non-financial, team-based, timely measures as well as expressing what action needs to be taken to make a difference in the current situation. Therefore, performance measures and KPIs have to be carefully developed from organizations’ CSF.

3. METHOD

The research goal is to search, find and synthesize the studies published in the research area of ITG practices and to identify the relationships between CSF, KPI and ITG practices. Therefore, the study is conducted by integrating the Systematic Mapping (Petersen et al., 2008) and Systematic Literature Review (Kitchenham, 2004) methods and techniques. Figure 3 presents the research model.

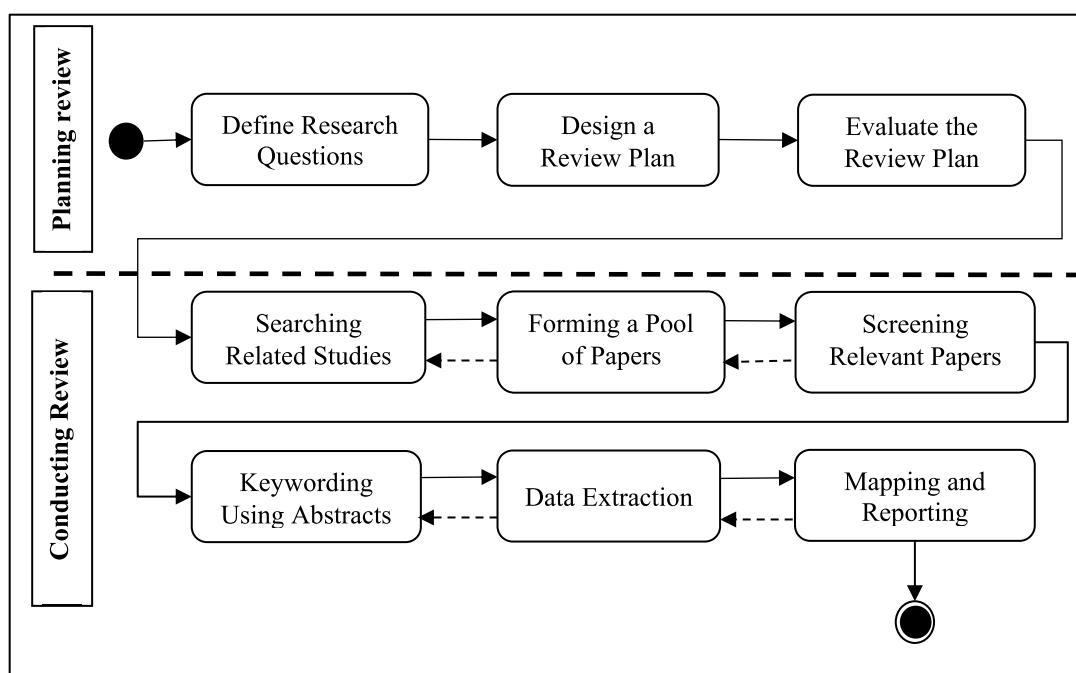


Figure 3. Research Model

3.1. Planning Review

The review process is a means of searching, identifying and interpreting available studies relevant to research questions. Depending on research limitations and constraints, automated search in electronic databases and the use of digital resources are determined as the main search strategy (Table 1).

Table 1

Online databases

	Database	URL
1	Science Direct	http://www.sciencedirect.com
2	Taylor & Francis	http://www.tandfonline.com
3	CiteSeerX	http://citeseer.ist.psu.edu
4	Google Scholar	http://scholar.google.com

Asking experienced researchers and experts, checking the reference lists of primary or secondary studies, accessing other web resources are determined as the ancillary search techniques. The research questions are formulated as below:

RQ-1: What are the CSF and KPI used in enterprises?

RQ-2: What are IT governance / management practices and their implementation challenges?

3.2. Conducting Review

The search strings and statements are developed according to the research questions along with the possible interventions, methods and factors in the context of ITG practices, CSF and KPI (Table 2):

Table 2

Search Strings and Statements

Nu	Search String and Statement
1	ORGANIZATION AND (PERFORMANCE OR IMPROVEMENT) AND INDICATOR
2	ORGANIZATION AND "CRITICAL SUCCESS FACTOR"
3	ORGANIZATION AND "KEY PERFORMANCE INDICATOR"
4	"KEY PERFORMANCE INDICATOR" AND (EFFECT OR IMPACT) AND ORGANIZATION
5	"CRITICAL SUCCESS FACTOR" AND (EFFECT OR IMPACT) AND ORGANIZATION
6	((FACTOR OR VARIABLE) AND (EFFECT OR IMPACT) "ON" AND "KEY PERFORMANCE INDICATOR")
7	((“IT GOVERNANCE” OR “IT SERVICE MANAGEMENT”) AND (MODELS OR STANDARDS) AND ORGANIZATION)
8	((“IT GOVERNANCE” OR “IT SERVICE MANAGEMENT”) AND (ITIL OR COBIT OR ITSM OR “ISO/IEC 20000” OR “ISO/IEC 38500” OR “CMMI”))
9	((“IT GOVERNANCE” OR “IT SERVICE MANAGEMENT”) AND (“KEY PERFORMANCE INDICATOR” OR “PERFORMANCE INDICATOR” OR “INDICATOR”))
10	((“IT GOVERNANCE” OR “IT SERVICE MANAGEMENT”) AND (EFFECT OR IMPACT) AND “ON” AND ORGANIZATION AND (PERFORMANCE OR IMPROVEMENT))

The downloaded studies formed the initial pool of papers in an online repository. It is the Google Drive Service, which is both for paper storage and information exchange between the researchers. In the next stage, the inclusion-exclusion criteria are applied for screening papers.

3.3. Screening Relevant Papers

The selection strategy consists of resource selection, application of the exclusion and inclusion criteria to screen relevant papers. The research questions are the main determinant of these criteria. For example, the abstracts explicitly mention ITG practices/models and KPI are directly included. The studies presenting the use of ITG for organizational improvement are also included. However, these studies are carefully reviewed in the next iterations and screened for theoretical and practical correspondence for the knowledge domain of ITG and KPI. Otherwise, they are excluded. Additionally, the papers lying outside the intersected research domain of ITG, KPI are directly excluded.

3.4. Keywording and Data Extraction

Data extraction and keywording is a two-step process as a way to develop a classification scheme and ensure that the existing studies were taken into account. The researchers read the abstracts and looked for similar concepts and keywords i.e. KPI, IT Governance, IT Service Management, which also reflect the context and contributions of the reviewed studies. While forming sets of keywords from the relevant papers, they were combined to establish a higher-level understanding. Then, the final sets of keywords were clustered and used to create a classification scheme for forming the categories related to KPI

and IT Governance. Therefore, these categories were useful for interpreting and classifying without evaluating each study in detail. The data relevant to the research questions were extracted and uploaded to the online repository in addition to the descriptive information about the papers.

4. FINDINGS AND DISCUSSION

We were initially able to reach 460 studies using the review protocol and consequently excluded 111 papers after a screening process. The included studies were classified as 97 magazine papers, 124 journal papers, 68 conference papers, 55 research reports, 3 technical reports and 2 books. Only major observations and textual data of review study are presented here due to space constraints and research limitations. Research findings are as follows:

RQ-1: What CSF and KPI are used in enterprises?

Table 3 and Table 4 below present the number of CSF and KPI, which are classified according to business functions.

Table 3
The number of CSF classified according to business functions

Sales & Marketing	Finance & Accounting	Manufacturing & Production	Human Resources
Careful package selection	Financial risk assessment	Adequate resources	Allocation of sufficient resources
Partnership with vendor	Financial risk planning	Operations and maintenance	Education on new business processes
Vendor support	Financial risk analysis	Project management	Employee empowerment and morale
Internal customer satisfaction	Financial risk management	Design and engineering	Experienced project manager-leadership
Community partnership	Responses to financial risk	Research & development	Knowledge and skills
External customer satisfaction	-	Energy management	Leadership
Closer customer relationships	-	Dedicated resources	Project Team Competence
Community engagement and partnership	-	Project management	Training
-	-	Flexible Operations	Team skills
-	-	ERP implementation strategy	Incentives
-	-	Formalized project plan	Management expectations
-	-	Good project scope management	
-	-	Improve facility energy awareness	-
-	-	Energy information	-
5 Studies	2 Studies	4 Studies	4 Studies

As discussed in the introduction and theoretical background sections, KPI links organizational activities and daily performances to the organizations’ CSF. Performance measures and KPI also have to be developed from these organizations’ CSF. Consequently, CSF should be associated with both KPI and ITG practices. However, our findings indicate that only a few studies explore and present the CSF (Table 3). On the other hand, most papers explore KPI (Table 4). This is not aligned with the research claims and is also contrary to expectations. Table 4 presents the number of KPI classified according to the business functions and sectors:

Table 4

The number of KPI classified according to the business functions and sectors

Business Functions & Sectors	Sales & Marketing	Finance & Accounting	Manufacturing & Production	Human Resources	Total
Finance	8	6	4	3	21
IT	3	18	2	61	84
E-Trade	6	2	35	1	44
Construction	12	14	14	6	46
Health	2	9	-	3	14
Telecom.	22	1	5	1	29
Retail	3	9	6	-	18
Production	20	34	83	17	154
Others	17	12	17	13	59
Total	93	105	166	105	469

In addition to these data, a set of CSF and KPI is presented in Appendix 1, which are grouped according to the business functions referenced in the reviewed papers. As can be seen from Table 4 and Appendix 1, Manufacturing & Production and Business Function have the highest number with 166 different KPI. The numbers of Human Resources and Finance & Accounting KPI are equal. However, the Sales & Marketing Business Function has the least amount of KPI. As for the sectors, the Production Sector is first with 154 different KPI. Then IT comes next as the second sector, in which 84 KPI are used for performance measurement. Contrary to expectations, the Finance and Health Sectors have the least amount of KPI studied, in which, in fact, the performance measurement should have a high priority when considering their mission critical roles played in our society (Dede 2016; Koyuncu, 2016).

Figure 4 below presents this knowledge in the form of a scatterplot with bubbles in order to be more powerful in giving an overview of business functions and KPI. The size of bubbles is proportional to the number of KPI that are in the pair of categories, which also corresponds to the sectors and business functions respectively.



Figure 4. The Systematic Map for KPI classified according to the business functions and sectors

RQ-2: What are IT governance / management practices and their implementation challenges?

Table 5 presents each ITG practice found in the reviewed papers. As can be seen, there are 14 different practices, some of which are used not only for ITG implementations, but also for software engineering process and IT management. Although these are IT management frameworks competing for effective IT governance, ITIL, COBIT and ITSM have the highest rate of an industry-wide implementation. ITIL focuses on the IT service delivery, continuous quality improvement and it provides an IT service life cycle. COBIT, on the other hand, separates governance from management, applies a single integrated framework and adopts a holistic approach. Although they are different in nature, they have great potential for complementing each other (Uysal, 2012).

Table 5

IT governance and/or management practices

IT Governance and Management Practice			
1	ITIL: Information Technology Infrastructure Library	8	ISO/IEC 9000: International Standard for Quality Management
2	COBIT: Control Objectives for Information and Related Technologies	9	SIX SIGMA: Quality Management
3	CMMI: Capability Maturity Model Integration	10	PRINCE2: Projects in Controlled Environments
4	ISO/IEC 20000: IT Service Management	11	PMBOK: Project Management Body of Knowledge
5	ISO/IEC 38500: International Standard for Corporate Governance of Information Technology	12	TOGAF: The Open Group Architecture Framework
6	ISO/IEC 27001: International Standard for Information Security Management	13	MOF: Microsoft Operations Framework
7	ISO/IEC 31000: International Standard for Risk Management	14	SQUARE: Software Quality Requirements and Evaluation

Table 6 below gives the ITG implementation challenges facing organizations. Lack of sufficient management support and commitment, ITG planning and its complexity and resistance to organizational change are among important challenges.

Table 6

IT governance/management practices and their implementation challenges

Nu.	Implementation Challenge	IT Governance Practice(s)	The Number of Studies Referenced
1	Top management resistance to change in organizational culture	ITIL, COBIT	3 studies
2	Employees' resistance to change	ITIL, CMMI	7 studies
3	Process complexity and lack of understanding	ITIL, COBIT	8 studies
4	Lack of sufficient resources for implementation	ITIL, COBIT, CMMI	9 studies
5	Lack of detailed guidelines and prescriptions	ITIL	6 studies
6	Issues on communication between different types of employees	ITIL	3 studies
7	Unpreparedness, not being ready to plan and implement IT governance practice	COBIT	1 study
8	Not allowing the skipping of the implementation steps	COBIT	1 paper
9	Not knowing how to use or implement process metrics and measures	COBIT	1 study
10	Issues on monitoring and controlling IT governance implementation processes	ITIL, COBIT	3 studies
11	Implementation issues stemming from regulatory environment	ITIL, CMMI	3 studies
12	Improper implementation	ITIL, CMMI, COBIT	4 studies
13	Issues on capability maturity levels of processes	CMMI	1 study
14	Lack of knowledge and skills	ITIL, ISO/IEC 38500, ISO/IEC 20000, COBIT	7 studies
15	Unclear definition of roles and responsibilities	ITIL	3 studies
16	Lack of focus and priority during implementation	ISO/IEC 20000	4 studies
17	Unrealistic expectations	ITIL	2 studies
18	Lack of organizational support and employees	ITIL, CMMI	4 studies
19	Unqualified personnel.	ITIL	1 study
20	Event management.	ITIL	2 studies
21	Lack of commitment	ITIL	3 studies
22	Service desk	ITIL	1 study
23	Lack of application tools and software	ITIL	5 studies

IT governance is important in that organizations expect better and clear returns from IT investments to achieve strategic goals. The regulatory requirements for IT management and control as well as performance measurement against various IT standards also add much to the ITG's significance. On the business side, CSF and KPI are the frequently referenced tools for business performance management., we claim that the ITG practices should be in relation to both CSF and their associated KPI. However, the literature on ITG cannot provide sufficient evidence exploring the relationships between ITG practices, CSF and KPI.

5. CONCLUSION

In this paper, we present the primary findings of a systematic research study. It is conducted according to the integrated principles and guidelines of Systematic Mapping and Systematic Literature Review Methods. Our research purpose is to

identify CSF and KPI used in organizations, ITGP and their implementation challenges. Consequently, the prominent findings and noteworthy points are as follows:

- Although KPIs have to be related to one or more CSF, this is not the case according to the research findings.
- KPIs are mostly used in the Manufacturing and Production Industry, which are also associated with the CSF for different organizations.
- There are approximately 14 types of ITGP practices and standards, together with various software engineering tools and methods, however, ITIL and COBIT practices are the foremost.
- Lack of top management support/sufficient resources, ITGP knowledge and skills, ITGP process complexity and resistance to organizational change are considered as the main implementation challenges.
- Although there are various ITGP domain-specific IT-related KPI, such as for COBIT, there is no sufficient work presenting the empirical evidence investigating the effects of ITGP on the business CSF and KPI at the time this paper was written .

The findings are encouraging though there are limitations that should be considered as well. This research can be viewed as an initial attempt to enhance ITGP, CSF and KPI knowledge and application domains. Therefore, our future research efforts will focus on the research gaps and trends put forth by this paper. It is hoped that this study may help researchers to obtain an overview of existing approaches to ITGP and enterprise performance.

Peer-review: Externally peer-reviewed.

Conflict of Interest: The authors have no conflict of interest to declare.

Grant Support: The authors declared that this study has received no financial support.

Author Contributions: Conception/Design of Study- M.P.U.; Data Acquisition- E.Ç.; Data Analysis/Interpretation- M.P.U.; Drafting Manuscript- M.P.U., E.Ç.; Critical Revision of Manuscript- M.P.U.;Final Approval and Accountability- M.P.U., E.Ç

Hakem Değerlendirmesi: Dış bağımsız.

Çıkar Çatışması: Yazar çıkar çatışması bildirmemiştir.

Finansal Destek: Yazar bu çalışma için finansal destek almadığını beyan etmiştir.

Yazar Katkıları: Çalışma Konsepti/Tasarım- M.P.U.; Veri Toplama- E.Ç.; Veri Analizi/Yorumlama- M.P.U.; Yazı Taslağı-M.P.U., E.Ç.; İçeriğin Eleştirel İncelemesi- M.P.U.; Son Onay ve Sorumluluk- M.P.U., E.Ç.

Kaynaklar/References

- AIS, (2020, November 1). The Association for Information Systems. Retrieved from <https://aisnet.org>
- Al-Alawi, A.I., Al-Marzooqi, N.Y, Mohammed, Y.F. (2007) Organizational culture and knowledge sharing: Critical success factors. *Journal of Knowledge Management*, 11(2), 22-42.
- Arnold Christian, Voigt Kai-Ingo (2017). Ecosystem effects of the industrial internet of things on manufacturing companies. *Acta Infologica*, 1(2), 99-108.
- Ash, C.G. & Burn, J.M (2003). Assessing the benefits from e-business transformation through effective enterprise management. *European Journal of Information Systems*, 12(4), 297-308.
- Barafort, B., Mesquida, A.-L. & Mas, A. (2017). Integrating risk management in IT settings from ISO standards and management systems perspectives. *Computer Standards and Interfaces*, 54(1), 176-185.
- Blokdyk G. (2020). Enterprise performance management a complete guide - 2020 edition. 5STARCOoks, USA.
- Chen, Y.-C & Wu, J.-H (2011). IT management capability and its impact on the performance of a CIO. *Information and Management*, 48(5), 145-156.
- Çetinkaya, E. (2017). The effects of IT governance practices on key performance indicators in organizations. Unpublished MS Thesis, Ufuk University, Turkey.
- CMMI (2021, February 6). *Capability Maturity Model Integration*. Software Engineering Institute, Carnegie Mellon University, USA. Retrieved from <https://www.sei.cmu.edu/>
- de Haes, S., van Grembergen (2009). An exploratory study into IT governance implementations and its impact on business /IT alignment. *Information Systems Management*, 26(2), 123-137.
- De Sousa, J. M. E. (2004). *Definition and analysis of critical success factors for ERP implementation projects*. Unpublished PhD Thesis, Universitat Politècnica de Catalunya, Barcelona, Spain.
- Dede B. (2016). *Building risk matrix by COSO and COBIT models in banking sector: An application for consumer loan process*. Unpublished PhD Thesis, Marmara University, Turkey.

- Dimon R. (2013). *Enterprise performance management done right: an operating system for your organization*. Wiley, USA.
- Franceschini F., Galetto M. & Maisano, D. (2007). *Management by measurement: Designing key indicators and performance measurement systems*. Springer Berlin Heidelberg, USA.
- ISACA (2021). *Information Systems Audit and Control Association; COBIT 5: Control Objectives for Information and Related Technology Framework*, ISACA, USA.
- ITGI (2020). IT Governance Institute, IL, USA.
- ITIL (2021, February 6). Information Technology Infrastructure Library. Retrieved from <https://www.axelos.com/best-practice-solutions/itil/what-is-itil>.
- Khudyakova, T., Zhuravlyov, V., Varkova, N., Aliukov, S., Shmidt, S. & Zhuravlyov, N. (2020). Improving approaches to strategic enterprise management in the context of sustainable development. *Sustainability*, 12(20), 1-20.
- Kitchenham, B. (2004). Procedures for performing systematic reviews. Technical Report, Keele University.
- Koyuncu A. E (2016). *The roles of the COBIT processes over the reliability of financial reporting and a research in banking sector*. Unpublished PhD Thesis, Dumlupınar University, Turkey.
- NCC Report, (2005). *IT Governance: Developing a successful governance strategy*. The National Computing Centre.
- Okul Ş., Muratoğlu O., Aydın M.A & Bilge H.Ş. (2019). A review on cyber risk management. *Acta Infologica*, 3(1), 34-45.
- Parmenter D. (2015). *Key performance indicators developing, implementing, and using winning KPI*. John Wiley & Sons, Inc., USA.
- Petersen, K., Feldt, R., Mujtaba, S. & Mattsso M. (2008). *Systematic mapping studies in software engineering*. Proceedings of the 12th International Conference on Evaluation and Assessment in Software Engineering, 68-77.
- Rehm, S.-V. & Goel, L. J. (2015). Role of information systems in empowering innovation networks. *MIS Quarterly*, 14(3), 87-103.
- Samsonowa T. (2012). *Industrial research performance management key performance indicators in the ICT industry*. Springer-Verlag Berlin Heidelberg, UK.
- Trkman, P. (2010). The critical success factors of business process management. *International Journal of Information Management*, 30(2), 125-134.
- Uysal, M.P. (2012). An e-learning design model integrated with information technology governance processes. *E-Journal of New World Sciences, Academy-Education Sciences*, 7(1), 251-268.
- Vincent, N.E., Higgs, J.L., Pinsker, R.E. (2017). IT governance and the maturity of IT risk management practices. *Journal of Information Systems*, 31(1), 59-77.

Appendix-1

	Sales & Marketing		Finance & Accounting		Manufacturing & Production		Human Resources	
CSF	Partnership with vendor	Vendor support	Financial risk assessment	Financial risk management	Adequate resources	Design & engineering	Leadership	Education
	Internal customer satisfaction	External customer satisfaction	Financial risk planning	Responses to financial risk	Operations and maintenance	Research & development	Training	Employee empowerment
	Community partnership	Closer customer relationships	Financial risk analysis	Expected commercial value	Project management	Energy management	Employee satisfaction	Knowledge and skills
	Transaction efficiency	Number of automated processes	Return on investment	Financial indebtedness	Resource management	Resource utilization	Employee attendance	Employee involvement
	Client's objectives clarified	Profit per on-line customer	Info. Man. Costs	Supply chain management costs	Failure frequency	Mean time between repair	Employee productivity	Management performance
	Customer increasing rate	Customer retention rate	Growth rate of investment	Internal rate of return	Maintenance intervention	Mean time to problem solution	Employee stability	Manpower Utilization rate
KPI	Demand management process	Number of customer referrals	Financial audit	Net income	Quality of product service	Mean down time	Team performance	Manpower Efficiency
	Willingness to purchase	Order entry accuracy	Profitability	Net present value	Reduction in tech. losses	Equipment effectiveness	Productivity	Labor efficiency
	Increase market share	Average waiting time	Debt ratio	Capital	Efficiency	Acceptance rate	Employees trained	Human resources capital
	Market share	Purchasing cost	Inventory cost	Revenue	Operating time	Process stability	Investment for training	Employees certified
	Profit per customer	Sales performance	Turnover cost	Net benefit	Process improvement	Throughput rate	Leadership index	Qualification index
	Customer visibility	Product availability	Turnover increment	Financial indebtedness	Volume flexibility	Variance of control error	Professional training	Operating hours
	Order-to-cash	Customer lost	Liquidity	Cash flow	Availability	Energy output	Employment of labor	Labor flexibility
	Market position	Order frequency	Cost reduction	Financial benefits	Process time	Cycle time	Turnover rate	Employee stability
Response time	Customer loyalty index	Total revenue	Total costs	Energy use	Level of safety	Recruiting costs	Analysis of abstain	

Two Level Kazakh Morphology*

İki Düzeyli Kazak Morfolojisi

Züleyha Yiner¹ , Atakan Kurt² 



*This submission appeared as part of Züleyha YINER's doctoral thesis, titled "Kazakça Gramer ve Semantik Analiz ve Kazakça-Türkçe Otomatik Çeviri Sistemi" with advisor Doç. Dr. Atakan KURT

¹(Res. Asst.), Siirt University, Faculty of Engineering, Department of Computer Engineering, Siirt, Turkey

²(Assoc. Prof.), Istanbul University-Cerrahpasa, Faculty of Engineering, Department of Computer Engineering, Istanbul, Turkey

ORCID: Z.Y. 0000-0001-7017-6114;
A.K. 0000-0002-9549-8475

Corresponding author:
Züleyha YİNER
Siirt University, Faculty of Engineering,
Department of Computer Engineering, Siirt,
Turkey
E-mail address: zulehayiner@siirt.edu.tr

Submitted: 24.12.2020
Revision Requested: 31.03.2021
Last Revision Received: 13.04.2021
Accepted: 26.04.2021
Published Online: 29.06.2021

Citation: Yiner, Z., & Kurt, A. (2021). Two level Kazakh morphology. *Acta Infologica*, 5(1), 79-98.
<https://doi.org/10.26650/acin.842758>

ABSTRACT

We present a comprehensive two level morphological analysis of contemporary Kazakh with implementation and a disambiguation test data set on the Nuve Framework. Our study differs from the similar studies in a number of ways: (i) Our study covers both derivational and inflectional morphology to a greater extent (ii) Our implementation consisting of orthographic rules, morphotactics, a root lexicon of roughly 24 thousand roots, a lexicon of roughly 150 suffixes is open source which can be downloaded, reviewed and tested. (ii) Roughly 10 thousand manually disambiguated parses are available as a morphological disambiguation data set. (iii) It is easily extensible meaning it can be modified or extended with new rules without any programming. (iv) we are able to tackle emerging problems quickly and easily since Nuve is maintained by our study group. (v) Our implementation can handle separately written morphemes or digraphs etc. directly. (vi) We also have a Turkish morphological parser/generator in Nuve for morphology based machine translation between Turkish and other Turkic languages since these closely related languages have a lot in common from lexical, morphological, and syntactic aspects.

Keywords: Kazakh Morphology, Natural Language Processing, Computational Linguistics

ÖZ

Bu çalışmada Çağdaş Kazakça'nın iki düzeyli kapsamlı bir morfolojisini sunulmuştur. Çalışma Nuve Çatısı üzerinde gerçekleştirilmiş ve belirsizlik giderme veri seti ile test edilmiştir. Çalışmamız benzerlerinden bir kaç yönden farklılık göstermektedir:(i) Çalışmamız hem yapım hem çekim morfolojisini benzerlerinden daha geniş olarak ele almaktadır. (ii) İki-düzeyli yazım kuralları, ek dizilim kuralları, yaklaşık 24 bin kelimeleik sözlük ve yaklaşık 150 adetlik ek sözlüğünden oluşan gerçeklememiz açık kaynak kodlu olarak paylaşımaya açılmıştır. Üçüncü taraflarca indirilebilir, gözden geçirilebilir ve test edilebilir. (iii) Gerçeklememiz var olan kuralların değiştirilmesi veya yenilerinin eklenmesiyle kolayca genişletilebilir bir yapıdadır. Programlama gerektirmez. (iv) Nuve Çatısı çalışma grubumuz tarafından geliştirildiği için ortaya çıkan yeni problemleri kolay ve hızlı bir şekilde çözebilmekteyiz. (v) Gerçeklememiz ayrı yazılan ekler, iki sembolden meydana gelen harfler gibi durumları kolayca ele alabilmektedir. (vi) Nuve Türkçenin iki düzeyli morfolojisini de içermektedir. Bu sayede kelime hazinesi, kelime yapısı ve cümle yapısı yönlerinden büyük benzerlikler içeren Türki dillerle Türkçe arasında morfoloji tabanlı makina çeviri yapılabilir.

Anahtar kelimeler: Kazak Morfolojisi, Doğal Dil İşleme, Hesaplamalı Dilbilim

1. INTRODUCTION

In this paper, a Kazakh two-level morphological description is given in detail. This description is implemented using an open-source morphological analyzer and machine translation system, Nuve. In this description a root word lexicon consisting of more than twenty thousand is used.

A number of morphophonemic processes take place during derivation or inflection of a word in Kazakh similar to those in Turkish including vowel drops, consonant drops, vowel changes as in vowel harmony rules, consonant changes as in consonant harmony rules, etc. These rules express the conditions in which these modifications occur. 22 two-level orthographic rules are written to describe Kazakh's morphophonemic processes in the system. Morphotactics is a sequence of rules that orders suffixes in order to generate a logical and meaningful word. For nominal and verbal paradigms, morphotactics is encoded as FSA.

Kazakh, belongs to the Turkic languages family, is spoken mainly in Kazakhstan as official language, and the other neighboring republics of Kyrgyzstan, China, Tajikistan, Uzbekistan, Pakistan, Russia, and Turkey. Also, it is spoken by more than 16 million people. Kazakh is an agglutinative language like other Turkic languages producing affixations of derivational and inflectional morphemes to root words. A number of morphophonemic rules help us to modify the surface realization of morphological constructions. Vowels in the affixed morphemes have to agree with the preceding vowel in certain aspects to achieve vowel harmony. Under certain circumstances, vowels in the root or affixes are deleted. Similar to vowel harmony, consonants in the root words or in the affixed morphemes experience certain modifications, and sometimes may be deleted. For example, the word аталарымыздан (atalarımızdan), can be broken into morphemes as follows:

ата+лар+ымыз+дан

where + indicates morpheme boundaries. This word can be translated into English as “from our ancestors”.

2. BACKGROUND and RELATED WORK

Morphotactics are the rules governing the order of morpheme affixations and are expressed as a finite state machine (FSA) using root words and suffixes. A two-level morphological model has been applied to a number of languages such as those by Oflazer (1994) for Turkish, by Alam (1983) for Japanese, by Antworth (1990) for English, by Kim et al. (1994) for Korean, Turkmen by (Tantuğ et al., 2006; Shylov, 2010), by (Görmez et al., 2011; Yiner et al., 2016) for Kyrgyz, by (Orhun et al., 2009; Keskin, 2012; Ablimit et al., 2016; Abdukerim, 2019), Crimean Tatar by (Altıntaş, 2001; Şanlı, 2018), Qazan Tatar by (Gökğöz, 2011) and so on.

Zafer et al. (2011) gives a brief two-level morphological description for Kazakh. In this description, the Kazakh phonological system is described with 27 two level rules which describe the transformation between the lexical level and the surface level of a word which is written in Latin (not original as in Cyrillic). And finite state machines to define nominal and verbal morphotactics of Kazakh. Both orthographic rules and the finite state morphotactics are implemented on a language independent framework which is Dilmaç (Shylov, 2010). Makhambetov et al. (2014) uses a data-driven approach to do morphological analyzing and labels morphemes as transition labels testing on the Kazakh National Corpus (Makhambetov et al., 2013). Their morphological analyzer model has two steps which are segmenting the given word and ranking each candidate segmentation using HMM and Markov chain rules. They do not consider compound words and some phonological rules which are important issues in a language.

Kessikbayeva & Cicekli (2014) gives a rule-based morphological analyzer using Xerox tools for the Kazakh language. This study does not work widely due to the lack of tools for analyzing Kazakh. Also, in this study a Latin transcription of Cyril text is used. Another lack of this study is not being able to handle separate written conjugations that is compound verbs. They used 1000 words randomly selected from web for testing and get nearly 96% correct analyses. Then, in the extended version of this study (Kessikbayeva & Cicekli, 2016), they use Foma which is an open source environment to implement a rule based morphological analyzer. They get approximately 99% correct analyses on test corpora which is nearly 15000

words written in Cyrillic. They use over 57 main alternation rules, including exception rules for each case separately, to define language grammar on Foma. This study is very important in terms of scope compared to the previous studies of the Kazakh language.

Bekmanova et al. (2017) proposes a uniform morphological analyzer both Kazakh (in Latin) and Turkish. But they give just morphological features of Kazakh based on (Eryiğit & Adalı, 2004), no more information how their morphological analyzer works. In (Washington, Salimzyanov & Tyers, 2014), an open-source finite-state transducer is proposed to get a morphological analysis of three Turkic languages; Kazakh, Tatar, and Kumyk with a limited root/stem lexicon.

3. TWO LEVEL DESCRIPTION OF KAZAKH ORTHOGRAPHY

Two-level morphology of (Koskeniemi,1983; Karttunen,1983) is one of the practical models in computational linguistics for morphological analysis of languages. In this model, for all languages rules and lexicons are combined with a parser for analyzing any language. Publicly available tools like PC-KIMMO (Antworth,1990) can be used to implement a two-level morphology. In the two-level morphology approach, both orthographic rules are defined using two-level rules and derivational/inflectional morphotactics are defined using FSAs. A word has two different representations or forms in this model: lexical and surface forms. The lexical form is a word structure or the representation of a word-formation, whereas the surface form is the written form of the word in the text generated by affixing suffixes by the morphology as given in the lexical form. The transformations from lexical to surface forms are defined with two-level orthographic rules as follows in (Oflazer, 1994).

The Kazakh alphabet has 23 consonants, 9 vowels (Biray, Ayan & Ercilasun, 2015). Table 1, Table 2, Table 3 shows Kazakh alphabet, consonants and vowels respectively.

Table 1

Kazakh Alphabet

Cyrillic	Latin	Cyrillic	Latin	Cyrillic	Latin	Cyrillic	Latin	Cyrillic	Latin	Cyrillic	Latin
а А	a	е Е	e	к К	q	п П	p	ф Ф	f	ь Ь	‘
ә Ә	ä	ё Ё	yo	л Л	l	р Р	r	х Х	x	ы Ы	ı
б Б	b	ж Ж	j	м М	m	с С	s	һ Һ	h	і І	i
в В	v	з З	z	н Н	n	т Т	t	ц Ц	c	ь Ь	
г Г	g	и И	ı	ң Ң	ñ	у У	w	ч Ч	ç	ә Ә	é
ғ Ғ	ğ	й Й	y	о О	o	ұ Ұ	u	ш Ш	ş	ю Ю	yu
д Д	d	к К	k	ө Ö	ö	ү Ү	ü	щ Щ	şş	я Я	ya

Table 2

Kazakh Vowels

	Unrounded	Rounded
Front	e,ə(e)	o(ö)
Back	a(a),ə(ä)	ı(u)

Table 3

Kazakh Consonants

		Labial	Labi-al-Dental	Dental	Palato-Al-veolar	Fronto-Palatal	Velar	Glottal Stop
stops	plosive	voiced	б(b)	д(d)	ц(c)	г(g)		
		voiceless	п(p)	т(t)	ч(ç)	к(k)		
continuant	trill	voiced	м(m),y(w)	н(n)		л(l),p(r)	һ(ñ)	
	fricative	voiced		в(v)	з(z)	ж(j)	ғ(ğ),й(y)	
		voiceless		ф(f)	с(s)	ш(ş)	х(x)	h(h)

In order to specify the two level morphology of Kazakh a subset of letters -called meta-letters – is used when expressing orthographic rules. These meta-letters are given below with their Cyrillic transliterations:

$$\begin{aligned}
 C &= \{b, \zeta, d, f, g, \check{g}, h, x, j, k, q, l, m, n, \check{n}, p, r, s, \check{s}, t, v, y, z\} \\
 &= \{\text{б, ч, д, ф, г, ф, х, ж, к, к, л, м, н, н, п, р, с, ш, т, в, й, з}\} \\
 C_{ts} &= \{f, x, h, k, q, p, s, \check{s}, t, \zeta\} = \{\text{ф, х, х, к, к, п, с, ш, т, ч}\} \\
 C_t &= \{b, d, g, \check{g}, j, l, m, n, \check{n}, r, v, y, z\} = \{\text{б, д, г, ф, ж, л, м, н, н, р, в, й, з}\} \\
 V &= \{a, e, i, i, o, \ddot{o}, u, \ddot{u}, \ddot{a}\} = \{a, e, ы, i, o, \text{э}, \text{ү}, \text{у}, \text{ә}\} \\
 V_b &= \{a, i, o, u\} = \{a, ы, o, \text{ү}\} \\
 V_f &= \{e, i, \ddot{o}, \ddot{u}\} = \{e, и, \text{ө}, \text{ү}\} \\
 I &= \{i, i\} = \{\text{ы, и}\} \\
 A &= \{a, e\} = \{a, e\} \\
 U &= \{u, \ddot{u}\} = \{\text{ү}, \text{у}\} \\
 K &= \{q, k\} = \{\text{к, к}\} \\
 N &= \{n, d, t\} = \{\text{н, д, т}\} \\
 M &= \{m, b, p\} = \{\text{м, б, п}\} \\
 G &= \{g, k, q, \check{g}\} = \{\text{г, к, к, ф}\} \\
 D &= \{d, t\} = \{\text{д, т}\} \\
 L &= \{l, d, t\} = \{\text{л, д, т}\}
 \end{aligned}$$

Above C, V, V_b, V_f represent the set of consonants, vowels, back and front vowels. We use the following notation in the lexical level of the two-level orthographic rules such as * for zero or more occurrence of the preceding letter, | for or (used to represent alternatives), ? for zero or one occurrence of the preceding letter, 0 for nothing on the surface (represents the dropped symbols), @ for the end of word (no more morpheme) and + for representing morpheme boundary (the beginning or end of a morpheme), not to be confused with regular expression + sign.

In the examples below, the first column is written in Cyrillic script and the second column is written in Latin script. In the third column (rightmost), we show the formulation in lexical form and the meanings of the root/words. The first line contains the lexical form of a word, and the second line the intermediate form. We don't show the surface form since it is almost identical to the intermediate form (the 0 place holders in the intermediate level are automatically deleted on the surface form).

1. These three rules express the vowel harmony in Kazakh. Vowels in the affixed suffixes to a word are resolved with respect to the previous vowel in the word. Back vowels are followed by vowels, and front vowels are followed by front vowels in a word according to the vowel harmony in Kazakh in a similar fashion to Turkish vowel harmony. An A on the lexical level in a suffix appended to a root is realized as a or e on the surface if the previous vowel in the root is (or is realized as) a back or front vowel respectively. In other words, the lexical A occurring after a back vowel in a word turns into the back vowel a on the surface:

$$A:a \Rightarrow V:V_b C^* + C^* _ _$$

$$A:e \Rightarrow V:V_f C^* + C^* _ _$$

ата+LAp ата0лар	ата+LAr ata0lar	N(ancestor) + NNI_PLU Ancestors
жер+DA жер0де	jer+DA jer0de	N(ground) + NNI_LOC on the ground

In some suffixes like Continuous tense suffix +A, verbal adjective suffix +ATIn or verbal adverb suffix +A, this rule does not provide the right changes for meta-letter +A. Thus, for these special cases, this rule works as below.

$$A:a \Rightarrow V:V_b C + C^* _ _$$

$$A:e \Rightarrow V:V_f C + C^* _ _$$

$$A:i \Rightarrow V + _ _$$

тұт+A+MIn тұт0а0мын	tut+A+MIn tut0a0mın	V(hold)+CONT+1PS (I will hold)
оқы+ATIn оқы0йтын	oqi+AtIn oqi0ytın	V(read)+JVD_ATIN reading (smt)
кел+ATIn кел0етін	kel+AtIn kel0etin	V(come)+JVD_ATIN coming

2. The lexical I in a suffix appended to a root is realized as ı or i on the surface if the previous vowel in the root is (or is realized as) a back or front vowel respectively:

$$I:ы \Rightarrow V:V_b C^* + C^* _ _$$

$$I:i \Rightarrow V:V_f C^* + C^* _ _$$

бала+NIң бала0ның	bala+NIn bala0nın	N(kid) + NNI_OBJ kid's
көз+LIK көз0дік	köz+LIK köz0dik	N(eye)+ NND_LIK Eyeglasses

3. This rule describes how the lexical U is resolved on the surface according to the Kazakh harmony rule. When a root is affixed with a suffix containing lexical U, lexical U is realized as u on the surface if the last vowel in the previous morpheme is a back vowel (V_b), otherwise, it is realized as ü on the surface.

$$U:ү \Rightarrow V:V_b C^* + C^* _ _$$

$$U:у \Rightarrow V:V_f C^* + C^* _ _$$

жый+Uв жый0ұв	jiy+Uv jiy0uv	V(collect)+ JVD_GAN Collecting
жет+Uв жет0ұв	jet+Uv jet0üv	V(reach)+ JVD_GAN reaching

4. When a suffix starting with lexical L is affixed to a root, L on the lexical level is realized as l on the surface if the last letter of the root is a vowel. Lexical L is realized as d or t if the last letter of the root is one of C_t or C_{ts} respectively:

L:l \Leftrightarrow V + ___

L:d \Leftrightarrow C_t + ___

L:t \Leftrightarrow C_{ts} + ___

ақыл+LI ақыл0ды	aqıl+LI aqıl0dı	N(intelligence) + JND_LI Intelligent
пайда+LI пайда0лы	payda+LI payda0lı	N(benefit) + JND_LI Beneficial
таш+LIк таш0тык	taş+LIк taş0tık	N(stone) + NND_LIK stone-filled

5. When a suffix starting with lexical N is affixed to a root, N on the lexical level is realized as n on the surface if the last letter of the root is a vowel. Lexical N is realized as d or t if the last letter of the root is one of C_t or C_{ts} respectively:

N:n \Leftrightarrow V + ___

N:d \Leftrightarrow C_t + ___

N:t \Leftrightarrow C_{ts} + ___

бала+NI бала0ны	bala+NI bala0nı	N(kid) + NNI_OBJ the kid
көз+NI көз0дү	köz+NI köz+dü	N(eye) + NNI_OBJ (was the) eye
иш+NIн иш0тин	iş+NIн iş0tin	N(work) + NNI_GEN of work

6. This rule which is quite similar to the previous rule deals with d and t transformations on the surface. The D on the lexical level maps to d on the surface, if a suffix starting with lexical D is affixed to a root ending with a vowel or a C_t consonant, otherwise D is resolved as a t on the surface:

D:d \Leftrightarrow [V| C_t] + ___

D:t \Leftrightarrow C_{ts} + ___

иш+DA иш0те	iş+DA iş0te	N(work) + NNI_LOC at work
жыл+DA жыл0да	jıl+DA jıl0da	N(year) + NNI_LOC in a year

7. This rule which is quite similar to the previous rule deals with b and p transformations on the surface. The B on the lexical level maps to b on the surface, if a suffix starting with lexical B is affixed to a root ending with a vowel or a C_t consonant, otherwise B is resolved as a p on the surface:

B:б \Leftrightarrow [V| C_t] + ___

B:п \Leftrightarrow C_{ts} + ___

көр+GAn+BIз
көр0ген0біз

kör+GAn+BIz
kör0gen0biz

V(see) + PAST+1PP
I saw

8. This rule deals with the realization of lexical M on the surface. The lexical M corresponds to the surface p when a suffix starting with lexical M is affixed to a root ending with a voiceless consonant C_{ts} . M is realized as b, if the root ends with z or c. Otherwise, M is resolved as m on the surface:

M:м \Leftrightarrow [V| C_t -{z,c}] + ___

M:б \Leftrightarrow {z,c} + ___

M:п \Leftrightarrow C_{ts} + ___

көр+MAcTAn
көр0местен

kör+MASTAN
kör0mesten

V(see)+ AVD_MASTAN
without seeing

жүз+MAЙ
жүз0бей

jüz+MAy
jüz0bey

V(swim)+ AVD_MAY
without swimming

кес+MAЙInшA
кес0пейінше

kes+MAyInşA
kes0peyinşe

V(cut)+ AVD_MAYINSA
without cutting

This orthographic rule of meta-letter M does not give the correct result on Question suffix +MA. For this suffix, orthographic rule works as below:

M:м \Leftrightarrow [V| C_t -{м,н,н,з}] + ___

M:б \Leftrightarrow {м,н,н,з} + ___

M:п \Leftrightarrow C_{ts} + ___

бала+MA
бала0ма

bala+MA
bala0ma

N(kid)+ QUES
is it kid?

ертең+MA
ертең0бе

erteñ+MA
erteñ0be

Adv(tomorrow)+QUES
is it tomorrow?

жоқ+MA
жоқ0па

jok+MA
jok0pa

Adj(absent)+QUES
is it absent?

In some suffixes that comes after verbal stem like noun derivation from verb suffix +MA, lexical M is realized as m, b or p again but in different circumstances:

M:м <=> { л,м,р,в,й,н,ң } + ___

M:б <=> [V|C_t-{л,м,р,в,й,н,ң}] + ___

M:п <=> C_{ts} + ___

жаз+МАстАН жаз0бастан	jaz+MAstAn jaz0bastan	V(write)+AVD_MASTAN without writing
қара+МАс қара0мас	qara+MAS qara0mas	V(look)+AVD_MAS does not look
бат+МАК бат0пақ	bat+MAK bat0paq	V(dive)+NVD_MAK swamp

9. This rule defines the transformation of lexical G at the beginning of some suffixes to ğ, g, k, q letters on the surface. G at the beginning of a suffix on the lexical level is realized as k and q on the surface when the last vowel of the affixed root is a front vowel V_f or back vowel V_b respectively and the root ends with a voiceless consonant C_{ts}. When the affixed root ends with back V_b or front V_f vowel and an optional voiced consonant C_v, then lexical G is resolved as ğ or g respectively:

G:ғ <=> [V:V_b|V:V_bC_t*] + ___

G:г <=> [V:V_f|V:V_fC_t*] + ___

G:к <=> V:V_fC_{ts}* + ___

G:қ <=> V:V_bC_{ts}* + ___

ал+ГАН ал0ған	al+GAN al0ğan	V(take)+ JVD_GAN Taker
көр+ГАН көр0ген	kör+GAN kör0gen	V(see)+ JVD_GAN Sighted
өт+ГАН+МІН өт0кен0мін	öt+GAN+MIN öt0ken0min	V(pass)+ VVI_TPASTGAN + VVI_PERS1s1 I passed
шап+ГАЛІ шап0қалы	şap+GALI şap0qalı	V(run)+ AVD_GALI running

10. When a root ending with the k consonant is affixed a suffix starting with a vowel, k at the lexical level is realized as g on the surface.

к:г => ___ + VC?

ек+Іл+ГАН ег0іл0ген	ek+Il+GAN eg0il0gen	V(sow) + VVI_PASSIL+ JVD_GAN sown
------------------------	------------------------	--------------------------------------

11. The p sound at the end of a root is realized as b on the surface when a suffix starting with a vowel is affixed to it.

п:б => ___ + VC?

қалып+cІ қалыб00ы	qalıp+sI qalıb00ı	N(cast) + NNI_POSS3s cast of
----------------------	----------------------	---------------------------------

12. When a root ending with the q consonant is affixed a suffix starting with a vowel, q at the lexical level is realized as ğ on the surface.

$q:F \Rightarrow _ + VC?$

соқ+Іп
соғ0ып

soq+ Ip
soğ0ıp

V(hit) + AVD_IP
by hitting

13. A syllable contains exactly one vowel in Kazakh and two consecutive vowels can't occur in a Kazakh word. This rule is based on this fact- states that when two consecutive vowels occur during affixation of suffixes, the one at the beginning of suffix is dropped. More precisely, when a root ending with a vowel takes a suffix starting with a vowel, then that vowel in the suffix is deleted on the surface.

$V:0 \Rightarrow V + _$

арба+Ім
арба00м

арба+Ім
arba00m

N(car) + NNI_POSS1s
my car

14. As an exception to the previous rule, in certain stems especially in the names of organs in the head such as (ауыз /mouth), (мойын /neck), (мұрын/nose), the last vowel in the stem is dropped instead of the vowel in the suffix. So when such a root takes a suffix starting with a vowel, then the last vowel in the stem is deleted on the surface.

$V:0 \Rightarrow _ C^* + V$

мойын+Ім
мойн0ым

moyın+Ім
moy0n0ım

N(neck) + NNI_POSS1s
My neck

15. When a suffix starting with a c (s) is affixed to a root ending with a consonant, the c at the lexical level in the beginning of the suffix drops on the surface.

$c:0 \Leftrightarrow C^* + _ I$

үй+cI
үй00ү

üy+sI
üy00ü

N(home) + NNI_POSS3s
it's home

16. When a stem takes a suffix containing K meta-letter, K meta-letter seems as k or q in surface form according to the last vowels of a stem.

$K:k \Rightarrow V:V_b C^* + C^* _$

$K:k \Rightarrow V:V_f C^* + C^* _$

жыл+LIK
жыл0дық

jıl+LIK
jıl0dıк

N(year) + NND_LIK
Annual

17. This rule is written for ablative case suffix +Dаh when coming after 1PS, 2PS, 3PS, and 3PP possessive suffixes. Under this circumstance, lexical D is realized as h on surface form.

$D:h \Leftrightarrow I_m | I_n | c | LArI + _ Ah$

көз+cI+Dаh
көз00i0нeн

köz+sI+Dаh
köz00i0nen

N(eye)+3PS-POSS+ABL
From his/her eye

18. Aorist tense suffix +Ap is realized as +c when it comes after negative suffix.

Ap:c <=> mA + __

бер+МА+Ар бер0ме00с	ber+MA+Ar ber0me00c	V(give)+AORST+3PS It does not give
------------------------	------------------------	---------------------------------------

19. One type of Optative Moods in Kazakh is done with +GI suffix followed by a person suffix and auxiliary word _келеді. If this auxiliary word _келеді followed by past tense suffix _еді, _келеді will be realized as _келетін in surface form.

ал+GI+Im+_келеді+_еді ал0ғы00м00келетін00еді	al+GI+Im+_keledi+_edi al0ġı00m00keletin00edi	V(take)+OPT+1PS+AUX+NARR I felt like taking
---	---	--

20. The lexical G in dative case suffix +GA is deleted on the surface if it comes after a personal suffix other than 3PS and 3PP. After 3PS and 3PP possessive suffixes, the lexical G in +GA is realized as н on the surface form.

G:0 <=> Ім|Ің|LArІң|LArІңІз + __A

G:н <=> cI|LArI + __A

жол+LAr+Ің+GA жол0дар0ың00а	jol+Lar+Iñ+GA jol0dar0ıñ00a	N(path)+PLU+2PS+DAT To your paths (singular)
бала+cI+GA бала0сы0на	bala+sı+GA bala0sı0na	N(child)+3PS-POSS+DAT To his/her child

This rule works differently while the dative case suffix +GA appended to 1st, 2nd and 3rd singular person pronouns, +GA is seen as +GAн (+GAн). This rule works as below:

мен+GA ма00ған	men+GA ma00ġan	P(i)+DAT To me
-------------------	-------------------	-------------------

21. Person pronouns can take case suffixes like nouns. This rule is written for transformation of 1 and 2 singular person pronouns мен (men), сен (sen) when take case suffixes such locative, genitive, ablative and the others. If 1 and 2 singular person pronouns take the Dative case suffix +GA, they are realized as маған and саған in surface form respectively (this transformation includes an orthographic rule for +GA suffix.). If they take the instrumental case suffix +мен (+men), i vowel epenthesis will occur between pronouns and suffix and they will be realized as менімен (menimen), сенімен (senimen).

сен+GA са00ған	sen+GA sa00ġan	P(you)+DAT To you
сен+мен сен0імен	sen+men sen0imen	P(you)+INS with you

22. The last sound of ol 3rd singular person pronoun drops whenever it takes one of genitive, accusative, dative, locative, ablative case suffixes. If this person pronoun takes instrumental case suffix, л at the end of pronoun is realized as ны in surface form.

ол+NI о00ны	ol+NI o00nı	P(it)+ACC it (object pronoun)
ол+мен оны0мен	ol+men onı0men	P(it)+INS with it

4. KAZAKH MORPHOTACTICS

Morphotactics in computational linguistics is a term for describing how the words in a natural language can be generated or parsed as a sequence of morphemes by affixing suffixes to roots in certain orders as defined by the morphology of the

language itself. The ordering of morphemes a nominal or verbal root takes in Kazakh is well defined and is strictly obeyed during word-formation.

The complete description of morphotactics of a language is usually given as a finite state machine (FSA). The states and the directed edges in an FSA represent the (subtypes of) Part-Of-Speech (POS or word class) of words and the affixed morphemes respectively. The two initial states are usually labelled as nominal root and verbal root since these are basic word classes. Nominal root includes nouns, adjectives, adverbs, etc. in our model. Final states represent the classes of words that take no more morphemes. The other states are intermediate states where generation/parsing can either continue or stop. Parsing or generation stops at the final states. The directed edges in an FSA are labelled with the lexical forms of morphemes (or sometimes with 0 meaning a transfer without a morpheme) a word takes while going from the source to the target state.

Such an FSA is used in creating morphological parsers or generators. Parsers usually have four main components:

1. Orthographic rules for modelling the phonological processes during affixation
2. An FSA for representing the ordering of morphemes during word formations
3. A root lexicon of nouns, verbs, adjectives, adverbs, etc.
4. A suffix lexicon for nominal/verbal derivations and conjugations

We used the Nuve Framework to implement the Kazakh morphology. Our implementation consists of the above components each specified in a separate file in the system. Thus we are able to parse (i.e. get the lexical form of a word from the surface form) and to generate (i.e. get the surface form of a word from its lexical form). In this study our aim was to generate all the words to avoid under generation at the expense over generation (generating invalid words).

4.1. Nominal Morphotactics

In this section, we present nominal morphotactics in the form of an FSA. Nominal morphotactics consist of nominal derivations and nominal conjugations. Nominal conjugation can be given as a simple formula called basic nominal model:

Nominal + Plural + Possessive + Case + Relative

The basic nominal model above represents the ordering of morphemes for nominal conjugation in the general case which is approximate and used for discussing morphology only. Comprehensive and precise nominal models are specified using FSAs given in Fig. 1. The parsing of an ordinary nominal word starts at Nominal root state and follows through Plural state with +LAp suffix, Possessive state with one of +Im, +H, +cI, +LApIH, +LApIHIZ, +LApI suffixes, Case state with one of +NIH, +NI, +GA, +DA, +DAn, +Men, +Menen suffixes, Relative state with +GI suffix. Since it is possible to treat a nominal as nominal verb (as in the case of a simple sentence with the nominal as the subject and “to be” as the verb), some nominal words will continue through the nominal verb (past and past perfect tense, conditional and imperfect moods), negative and question states shown in the lower part of the nominal FSA. The lexical form of кiтаптарымдағы (one which is in my books) is given below as an example basic nominal model:

Кiтап	+LAp	+Im	+DA	+GI
Кiтап	Отар	Оым	Ода	Оғы
Book	Plural	Poss.1SP	Case	Relative

Nominal derivation which creates new stems is provided through following suffixes in Kazakh shown in Fig. 1: noun to noun (+Ab, +Dac, +LIK, +Lip, +AK, +cIz, +шIK, +шIl, +шAK, +шAн, +шA, +шI, +тай, +LI, +ғылт, noun to adjective (+LI, +ғылт, +қылт, +DAGAn, +ншI, +Nан, adjective to noun (+шIK), noun to verb (+cIH, +cI, +pA, +GAp, +LAT, +LAc, +LAn, +LA, +IK, +I, +DA, +Aй, +Ap, +Al, +A), verb to noun (+MA, +Im, +Ic, +GI, +IK, +AK, +IHdI, +Gш, +ш, +In, +MAK, +GIH, +MAJ, +шAK) verb to adverb (+II, +MAй, +GANdA, +GANшA, +GAlI, +MAcтAn, +MAйIHшA, +A), adjective to adverb (+DAGAn).

4.2. Verbal Morphotactics

In this section, we describe verbal morphotactics in the form of an FSA. Verbal morphotactics consists of verbal derivations and verbal conjugations similar to nominal expressed as simple formulas called basic verbal model:

Verb + Negative_MA + Tense + Tense | Mood + Person

Verb + Tense + Negative_jok/emes + Person

There are two ways of creating negatives in verbal model one with Negative_MA another with Negative_jok/emes. көрмегенбіз (kormegenbiz, we did not see or we had not seen) and алмақ емеспін (almaq emespin, I will not take) are given as verbal model examples:

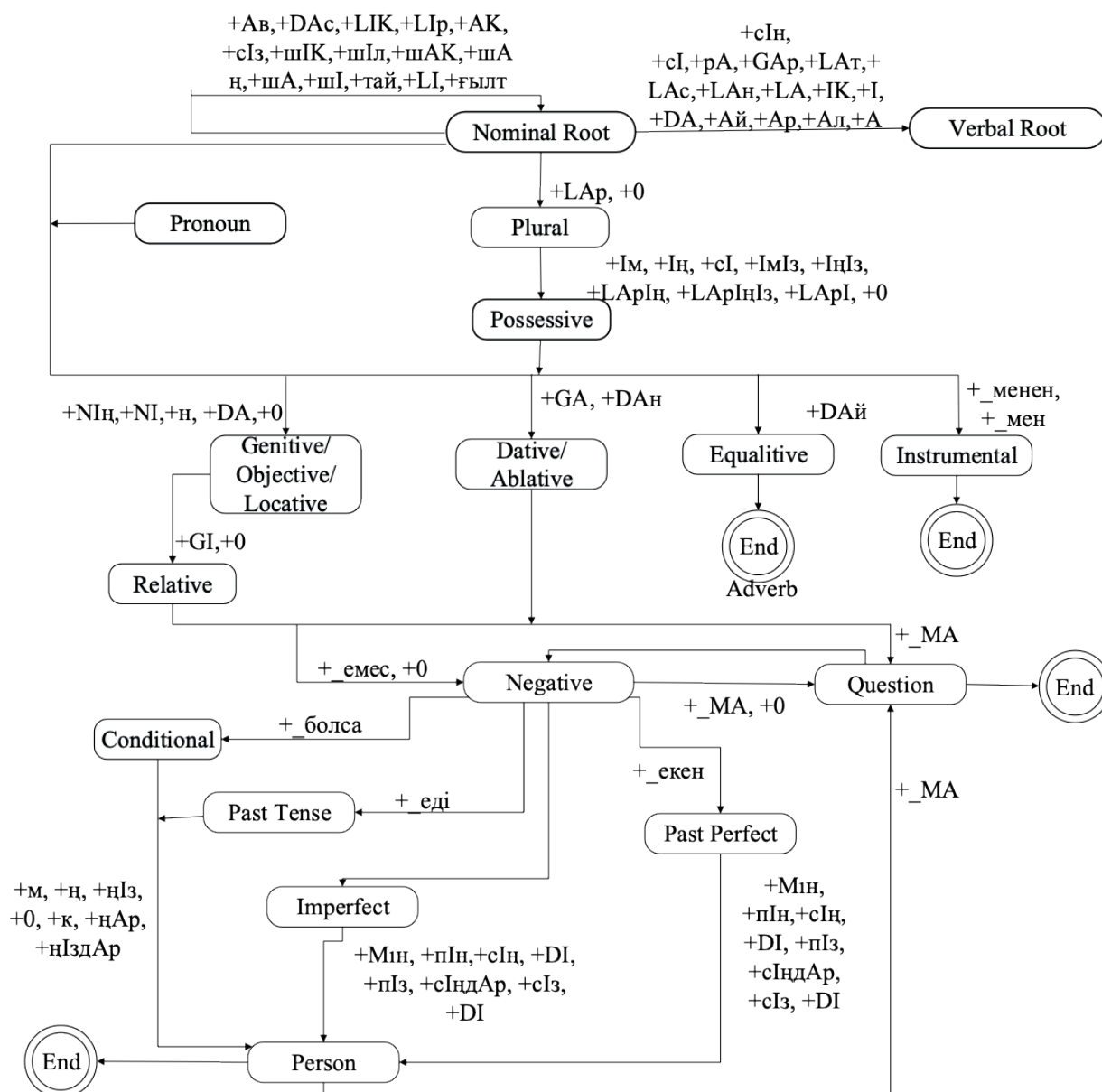


Figure 1. Nominal morphotactics

көр	+mA	+GAH	+BIз
көр	0ме	0геH	0бiз
verb	Negative	Past Tense	1.Prsn Plural
ал	+MAK	+_емес	+MIH
ал	0мақ	00емес	0пiH
Verb	Future Tense	Negative	1.Prsn Sing

Verbal paradigm is described by using FSA in Fig. 2 and Fig. 3. It is given as two parts for easy reading. Voice suffixes always come before tense suffixes. But it is not obligatory for the verbal root to take voice suffixes, it can take tense suffixes directly. Person suffixes come after simple tense suffixes. If a verbal stem takes a narrative tense suffix, person suffixes will come after narrative suffixes. There is not a specific order for negative verbal inflection. In some tenses, a negative suffix comes before tense suffixes while others are taken as the last suffix before person suffixes. ‘+0’ means transition without a suffix (empty transition).

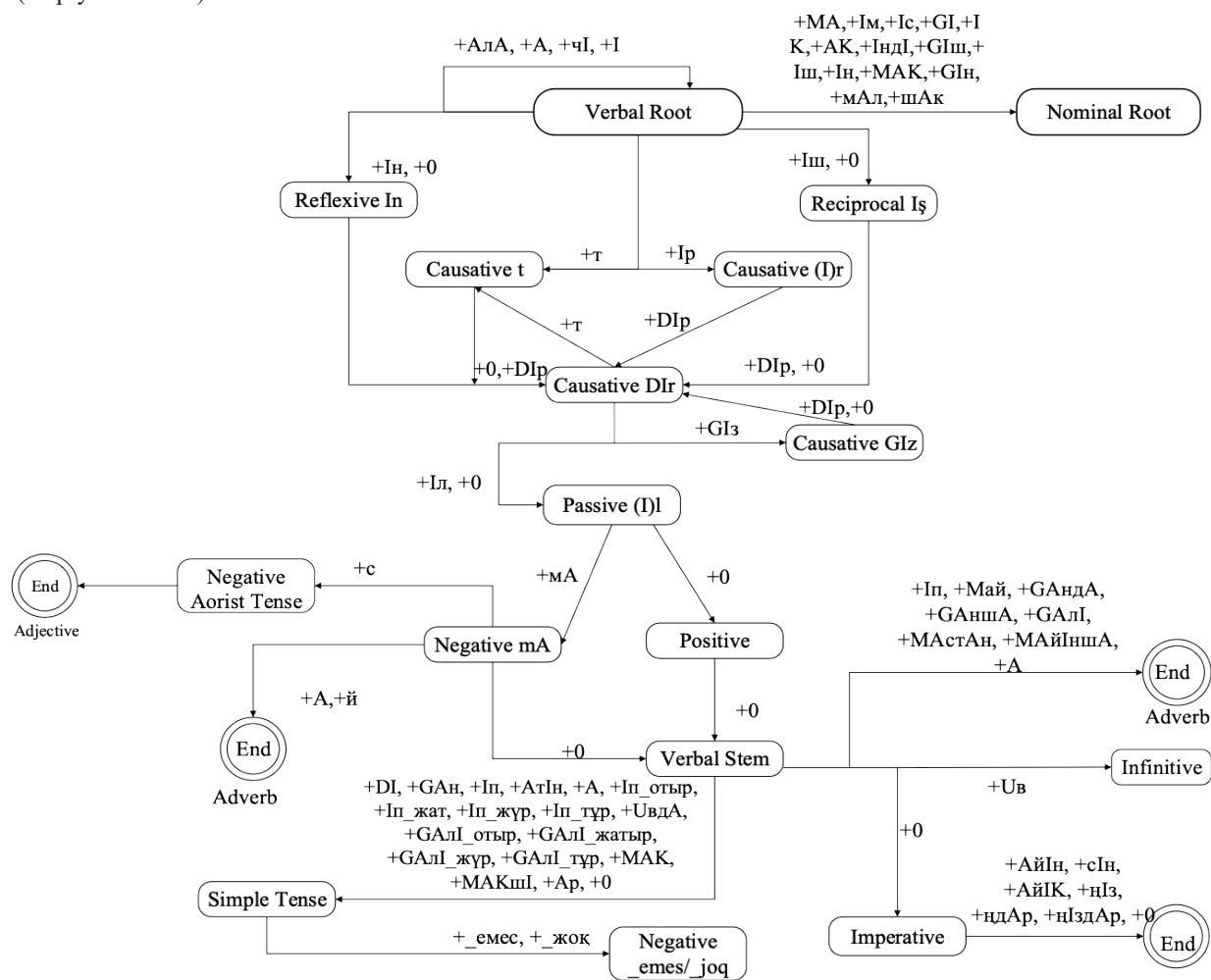


Figure 2. Verbal morphotactics first part

5. IMPLEMENTATION

An extensive description of Kazakh morphology is given above. We used Nuve for the implementation. Nuve [24] is an open-source project for morphological parsing and machine translation. Nuve was primarily developed for Turkic languages and includes an implementation of two level Turkish morphology.

The implementation in Nuve requires four distinct components each provided as a formatted text or XML file: root lexicon (TXT), suffix lexicon (TXT), orthographic rules (XML), morphotactics (XML).

Root Lexicon: There are approximately 24000 entries in the root lexicon. A few entries for the root lexicon are given in Table 5 below. Each root entry is specified using surfaces (different surface forms of the root), lex, active, id (POS: Noun, Adjectives, Verbs, Pronouns, Adverbs, Exclamations, Conjunctions, etc.), flags, rules (specific orthographic rules to apply for this root entry) attributes in this text file in the CSV format. Lex and flags attributes are used for special cases.

Table 4
Root lexicon

root	surfaces	Id	flags	rules
абажур		NOUN	noun	
абайла		VERB	verb	
абақ	абар	NOUN	noun	MUTATION_qg
бар		VERB	verb, cnt	
кел		VERB	verb, cnt	

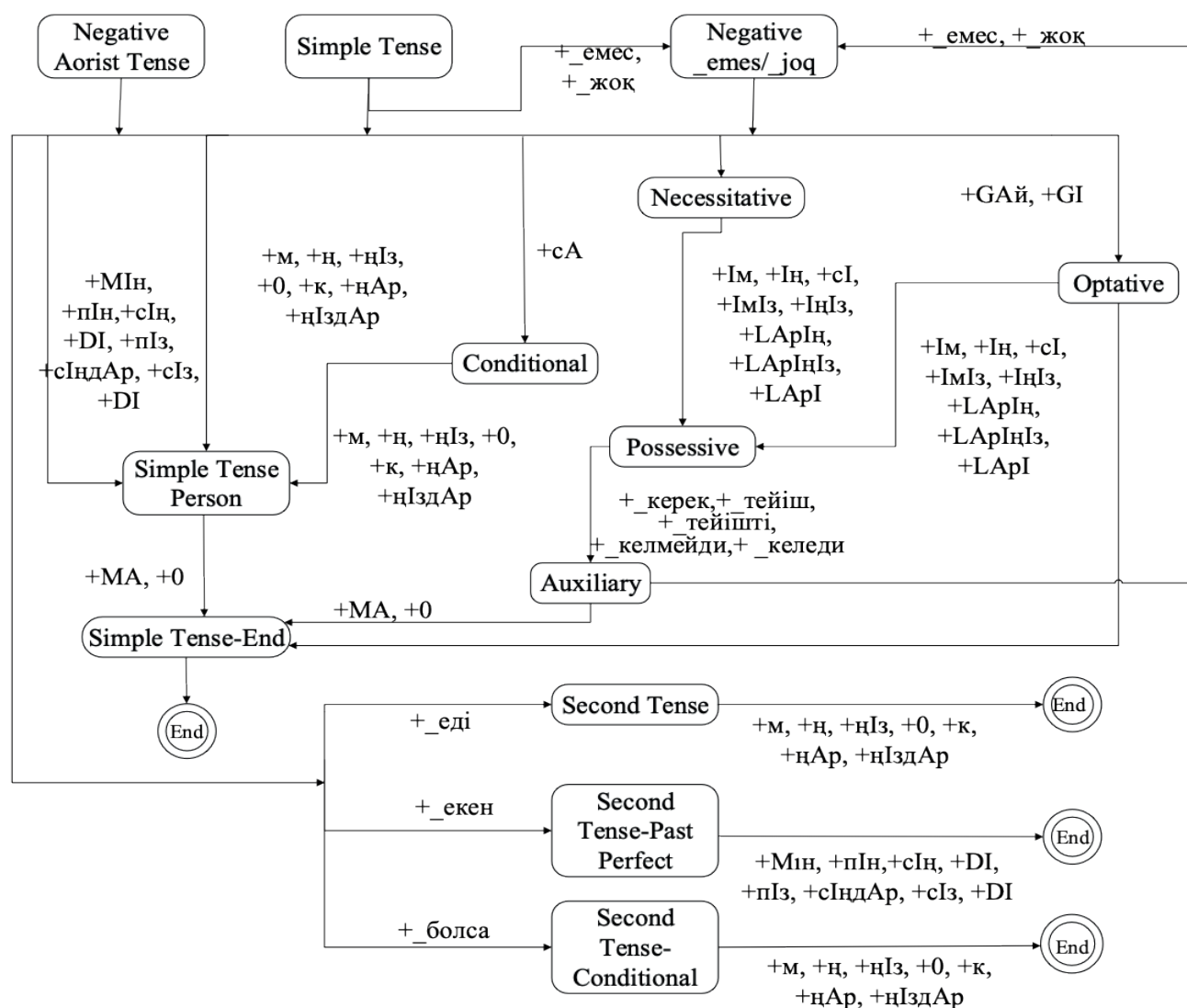


Figure 3. Verbal morphotactics second part

Suffix lexicon: Derivational and inflectional suffixes in the suffix lexicon are defined using id (a descriptive label), lexical (lexical form of the entry), type (suffix type: I: inflectional, D: derivational), suffix group (suffix category), rules (orthographic rules specifically applied to this entry), surfaces (distinct surface forms of the entry) attributes in this text file in the CSV format.

Table 5

Suffix Lexicon

id	lexical	grup	rules	surfaces
ni_plural_LAr	LAp		TRANSFORMATION_A, TRANSFORMATION_L	лар,лер,дар,дер,тер,тар
ni_possesive_1sg_(I)m	Im	noun-possesive	TRANSFORMATION_I, VOWEL_DROP	м,ым,ім
vi_tense_fut_GAll_otr	ГАлI_отыр	verb-tense	TRANSFORMATION_I, REMOVE_UNDERSCORE, TRANSFORMATION_G, TRANSFORMATION_A	ғалы_отыр, гелі_отыр, қалы_отыр, келі_отыр

Morphotactics file: This file contains suffix groups and the FSA in XML format as shown in Fig. 4. Suffixes are grouped into suffix groups for ease of use. The nodes and edges of the FSA are specified using source and targetGroup elements in the file.

```

<morphology lang="Kz-KZ">
  <suffixGroup name="NOUN_POSSESSIVE">
    <suffix>NI_POSSESSIVE_1sg_(I)m</suffix>
    <suffix>NI_POSSESSIVE_2sg_(I)ñ</suffix>
    ...
  </suffixGroup>
  ...
  <graph>
    <source id="NOUN">
      <targetGroup id="NOUN_POSSESSIVE" />
      <targetGroup id="NOUN_CASE" />
      <targetGroup id="ND_NOUN" />
      <targetGroup id="ND_VERB" />
      ...
    </source>
    ...
  </graph>
</morphology>

```

Figure 4. Morphotactics XML file

Orthography file: Two-level orthographic rules are encoded in XML in this file shown in Fig. 5. The conditions and transformations of the rules are defined using the transformation and conditions elements respectively. Complex logical conditions can be formed using and, or or operands. Symbols or morphemes in the words can be accessed using previous and next attribute values. Symbols can be deleted, inserted or replaced with action attribute of the transformation element.

```

<rule id="TRANSFORMATION_A" phase="2">
  <description>Rule for meta-letter A changing into (a or e) </description>
  <transformation morpheme="This" action="Replace" operandOne="A" operandTwo="a" flag="all">
    <conditions flag="Or">
      <condition morpheme="Previous" operator="LastVowelEquals" operand="аәёуыюя" />
    </conditions>
  </transformation>
  <transformation morpheme="This" action="Replace" operandOne="A" operandTwo="e" flag="all">
  </rule>

```

Figure 5. Orthography XML file

6. DATA SET AND TEST

We tested two level description of Kazakh morphology on data set from Universal Dependency¹. The text of UD Kazakh treebank (Makazhanov et al., 2015; Tyers et al., 2015) is collected from various sources such as Wikipedia, some folk tales, sentences from the UDHR, news and phrasebook sentences. The tokenization and morphological processing in the Kazakh UD treebank follow the principles of Turkic lexica in Apertium. Originally, the treebank is randomly split into training (80%), testing (10%), and development (10%) sets. We gave data set to Nuve as input for morphological analysis. The data set consists of over 1K sentences and more than 10K tokens. Over 8K tokens have only one morphological analysis. More than 1500 tokens have two or more morphological analyses. While finding more than one morphological analysis for a word is considered an advantage in terms of performance of a parser, it results in the problem of choosing right morphological parsing among many (morphological disambiguation). Morphological disambiguation is important for many NLP applications which depend on the accurate parsing of words, such as syntactic parsing, word sense disambiguation, spelling correction and machine translation and so on. We plan to use this data set for an experimental study on the morphological disambiguation of Kazakh in the future.

7. SPECIAL CASES AND EXAMPLES

While defining the morphology of the language, some special cases are encountered. In this section, these special cases are presented.

1. There is a confusion about the noun to adjective +LI suffix and the accusative +NI suffix both applied to nouns. For example, құрметті адам (kurmet+LI) kurmetti adam, *respectful man*, and көрсеткен құрметті (kurmet+NI) жақшы көрді(көр+DI) *körsetken kurmetti jaqsı kördi, s/he liked the respect shown to her/him*. Our implementation is able to parse both words.

2. Sometimes the negative +MA suffix and the verb to noun derivation +MA suffix may be confused since both are affixed to verbs and have the same surface. For example, аспа ac+MA, aspa, *hanging (nominal) or do not hang (verb)*. However, this ambiguity can be resolved at the morphological disambiguation phase.

3. After possessive suffixes (except 1st and 2nd plural person), ablative case suffix +DAn (+DAn) becomes +nAn (+nAn). For example:

ата+(I)m+DAn, атамнан, *atamnan, from my father*

үй+(I)іміз+DAn, үйімізден, *üyimizden, from our home*

4. +_şe is used in the question form of Conditional Mood as a special case, in addition to the usual +MA question suffix. For example, жазсам ше or жазсам ба? *jazsam ше/be?, should I write*, келсем ше? or келсем бе? *kelsem ше/be?, should I come?*

5. In Kazakh some morphemes are written separately such as the question suffix +MA, or the continuous tense suffixes +Ip отыр, +Ip тұр, +Ip жүр, +Ip жат. Our implementation is able to parse these separately written suffixes without difficulty as shown below whereas other implementations either parse these suffixes as separate words or can't parse at all.

келіп отырсың, кел(VERB)+Ip отыр+2.PrsnSing

6. In the continuous tense, jatır auxiliary verb requires an extra +A suffix only if it is appended to kel- and bar- verbs. This problem is handled with the *cnt* flag attribute in the root lexicon in the implementation. кел+е жатырмын, *I am coming* and бар+а жатырсыздар, *I am going/arriving*.

7. In adjective clauses made with verbal adjective +GAn suffix, the possessive meaning is provided by a pronoun before the clause. For example, мен жазған хат, *men jazğan hat, the letter I wrote*, сен күткен арба, *sen kütken arba, the car you*

¹ https://universaldependencies.org/treebanks/kk_ktb/index.html

waited for. The person suffix normally attached to the verb, in this case, is replaced with a pronoun preceding the verb. These types of adjective clauses are too complex to be properly parsed by any parser.

барма	бар+МА	V(бар) +NEG+2PS (Imperative) <i>do not go</i>
	бар+МА	V(бар) + VND(MA) <i>going</i>
үйлерің	үй+LAp+Ің	N(үй) + PLU+2PS-POSS <i>your homes</i>
	үй+LApІң	N(үй) + 2PP-POSS <i>your home (polite)</i>
жылы	жылы	Adj(жылы) <i>warm</i>
	жылы	Adv(жылы) <i>warmly</i>
	жыл+сі	N(жыл) + 3PS-POSS <i>it's year</i>
жүз	жүз	Adj(жүз) <i>hundred</i>
	жүз	N(жүз) <i>face</i>
	жүз	N(жүз) <i>tribe</i>
	жүз	N(жүз) <i>tribe</i>
басты	бас+LI	N(бас) +NND(LI) <i>having (a) head</i>
	бас+DI	V(бас)+ PAST+ 3PS <i>(s/he) stepped on</i>

8. CONCLUSION

We have provided a comprehensive two level description of Kazakh morphology with an implementation on Nuve Framework consisting of 22 orthographic rules for the definition of morphophonemic processes, the morphotactics for nominal and verbal derivations and inflections using a large root lexicon and a suffix lexicon in which special or exceptional cases are manually annotated for ease of use. Our study covers both derivational and inflectional morphology to a greater extend (ii) Our implementation consisting of orthographic rules, morphotactics, a root lexicon of roughly 24 thousand roots, a lexicon of roughly 150 suffixes is open source and free (https://nuvestudio.com/downloads/kz_KZ/kz_KZ.zip) which can be downloaded, reviewed and tested. (ii) Roughly 10 thousand manually disambiguated word parses are available as a morphological disambiguation data set. (iii) Our implementation is easily extensible meaning it can be modified or extended with new rules without any programming, because all resources are encoded as XML/CSV files and ready to use as soon as they are uploaded into the system. (iv) We are able tackle emerging problems quickly and easily, since Nuve is maintained by our study group. The original version of Nuve was developed in .NET. The current version is in PHP. (v) Our implementation can handle problems such as separately written morphemes or digraphs etc. directly. No indirect methods are necessary to handle these cases. (vi) We also have a Turkish morphological parser/generator in Nuve for morphology based machine translation between Turkish and other Turkic languages (which is the next part of our project) since these closely related languages have a lot in common from lexical, morphological, and syntactic aspects. (vii) Finally, we have Kazakh morphological parser and generator available online to anyone who wishes to use.

Peer-review: Externally peer-reviewed.

Conflict of Interest: The authors have no conflict of interest to declare.

Grant Support: The authors declared that this study has received no financial support.

Author Contributions: Conception/Design of Study- Z.Y., A.K.; Data Acquisition- Z.Y., A.K.; Data Analysis/Interpretation- Z.Y., A.K.; Drafting Manuscript- Z.Y., A.K.; Critical Revision of Manuscript- Z.Y., A.K.; Final Approval and Accountability- Z.Y., A.K.

Hakem Değerlendirmesi: Dış bağımsız.

Çıkar Çatışması: Yazarlar çıkar çatışması bildirmemiştir.

Finansal Destek: Yazarlar bu çalışma için finansal destek almadığını beyan etmiştir.

Yazar Katkıları: Çalışma Konsepti/Tasarım- Z.Y., A.K.; Veri Toplama- Z.Y., A.K.; Veri Analizi/Yorumlama- Z.Y., A.K.; Yazı Taslağı- Z.Y., A.K.; İçeriğin Eleştirel İncelemesi- Z.Y., A.K.; Son Onay ve Sorumluluk- Z.Y., A.K.

References/Kaynaklar

- Abdukerim, G., Tursun, E., Yang, Y., & Li, X. (2019). Uyghur morphological analysis using joint conditional random fields: Based on small scaled corpus. *Discrete & Continuous Dynamical Systems-S*, 12(4&5), 823.
- Ablimit, M., Kawahara, T., Pattar, A., & Hamdulla, A. (2016). Stem-affix based Uyghur morphological analyzer. *International Journal of Future Generation Communication and Networking*, 9(2), 59-72.
- Alam, Y. S. (1983). A two-level morphological analysis of Japanese. *In Texas Linguistic Forum*, 22, 229-252.
- Altıntaş, K., & Cicekli, I. (2001). A morphological analyser for Crimean Tatar. *Proceedings of the 10th Turkish Symposium on Artificial Intelligence and Neural Networks (TAINN'2001)*, North Cyprus, 180-189.
- Antworth, E. L. (1990). PC-KIMMO: A two-level processor for morphological analysis. *Summer Institute of Linguistics, International Academic Bookstore*, Dallas, Texas.
- Bekmanova, G., Sharipbay, A., Altenbek, G., Adali, E., Zhetkenbay, L., Kamanur, U., & Zulkhazhav, A. (2017). A uniform morphological analyzer for the Kazakh and Turkish languages. *Proceedings of the Sixth International Conference on Analysis of Images, Social Networks and Texts*, Moscow, Russia, 20–30.
- Biray N., Ayan E., Ercilasun G. K., (2015). *Çağdaş Kazak Türkçesi Ses-Şekil- Cümle Bilgisi- Metinler* (2nd ed.). Istanbul, Turkey: Bilge Kültür Sanat.
- Eryiğit, G., & Adalı, E. (2004, February). An affix stripping morphological analyzer for Turkish. *Proceedings of the IASTED International Conference Artificial Intelligence and Applications*, Innsbruck, Austria, 299–304.
- Gökgöz, E., Kurt, A., Kulamshae, K., & Kara, M. (2011, May). Two-level Qazan Tatar morphology. *Proceedings of the 1st International Conference on Foreign Language Teaching and Applied Linguistics (FLTAL'11)*, Sarajevo, Bosnia and Herzegovina, 428-432.
- Görmez, Z., Ünlü B. S., Kurt, A., Kulamshae, K., & Kara, M. (2011). An overview of two-level finite state Kyrgyz morphology. *Proceedings of the 2. International Symposium on Computing in Science & Engineering (ISCSE)*, Aydin, Turkey, 48-52
- Karttunen, L. (1983, December). KIMMO: a general morphological processor. *In Texas Linguistic Forum*, 22, 163-186.
- Keskin, R. (2012). *Two Level Uyghur Morphology and Uyghur Turkish Machine Translation*. (Master's Thesis). Fatih University the Graduate Institute of Sciences and Engineering, Istanbul.
- Kessikbayeva, G., & Cicekli, I. (2014, June). Rule-based morphological analyzer of Kazakh language. *In Proceedings of the 2014 Joint Meeting of SIGMORPHON and SIGFSM*, Baltimore, Maryland, 46-54.
- Kessikbayeva, G., & Cicekli, I. (2016). A rule based morphological analyzer and a morphological disambiguator for Kazakh language. *Linguistics and Literature Studies*, 4(1), 96-104
- Kim, D. B., Lee, S. J., Choi, K. S., & Kim, G. C. (1994, August). A two-level morphological analysis of Korean. *In Proceedings of the 15th Conference on Computational Linguistics, Vol 1*, 535-539.
- Koskenniemi, K. (1983, August). Two-level model for morphological analysis. *In Proceedings of the Eighth International Joint Conference on Artificial Intelligence*, Karlsruhe, Germany, 683-685.
- Makazhanov, A., Sultangazina, A., Makhambetov, O., & Yessenbayev, Z. (2015). Syntactic annotation of Kazakh: following the universal dependencies guidelines. A report. *In Proceedings of the International Conference Turkic Languages Processing- TurkLang-2015*, Kazan, Tatarstan, 338-350.
- Makhambetov, O., Makazhanov, A., Yessenbayev, Z., Sabyrgaliyev, I., & Sharafudinov, A. (2014). Towards a data-driven morphological analysis of Kazakh language. *Türkiye Bilişim Vakfı Bilgisayar Bilimleri ve Mühendisliği Dergisi*, 7(2), 31-36.
- Makhambetov, O., Makazhanov, A., Yessenbayev, Z., Matkarimov, B., Sabyrgaliyev, I., & Sharafudinov, A. (2013, October). Assembling the Kazakh language corpus. *In Proceedings of the 2013 Conference on Empirical Methods in Natural Language Processing*, Seattle, Washington, USA, 1022-1031.
- Oflazer, K. (1994). Two-level description of Turkish morphology. *Literary and linguistic computing*, 9(2), 137-148.
- Orhun, M., Tantug, A. C., & Adali, E. (2009). Rule based analysis of the Uyghur nouns. *International Journal on Asian Language Processing*, 19 (1), 33-43.
- Shylov, M. (2010). *Two level Turkmen morphology and a Turkmen Turkish machine translation*, (Master's Thesis). Fatih University the Graduate Institute of Sciences and Engineering, Istanbul.

- Şanlı, T. (2018). *Kırım Tatarcası'nın biçimbilimsel çözümlemesi ve Kırım Tatarcası-Türkçe biçimbilimsel makina çevirisi Sistemi*. (Master's Thesis). Istanbul University Institute of Graduate Studies in Science and Engineering, Istanbul.
- Tantuğ, A. C., Adalı, E., & Oflazer, K. (2006, August). Computer analysis of the Turkmen language morphology. *Proceedings of the 5th International Conference on Natural Language Processing*, Turku, Finland, 186-193.
- Tyers, F. M., & Washington, J. (2015). Towards a free/open-source Universal Dependency Treebank for Kazakh. In *Proceedings of the International Conference Turkic Languages Processing, TurkLang-2015*, Kazan, Tatarstan, 276-289 .
- Washington, J., Salimzyanov, I., & Tyers, F. M. (2014, May). Finite-state morphological transducers for three Kypchak languages. *Proceedings of the Ninth International Conference on Language Resources and Evaluation, LREC*, Reykjavik, Iceland ,3378-3385.
- Yiner, Z., Kurt, A., Kulamshae, K., & Zafer, H. R. (2016, May). Kyrgyz orthography and morphotactics with implementation in NUVE. *Proceedings of International Conference on Engineering and Natural Sciences*, Sarajevo, Bosnia and Herzegovina, 1650-1658.
- Zafer, H. R., Tilki, B., Kurt, A., & Kara, M. (2011, May). Two-level description of Kazakh morphology. *Proceedings of the 1st International Conference on Foreign Language Teaching and Applied Linguistics (FLTAL'11)*, Sarajevo, Bosnia and Herzegovina, 560-564.
- Zafer, H. R., "Nuve: A natural language processing library for Turkish in C#". [Online]. Available: <https://github.com/hrzafer/nuve>. (05.12.2020).

Audits in a Textile Company and Advantages

Bir Tekstil İşletmesinde Denetimler ve Avantajları

Mine Seçkin¹ , Feden Koç² , Ahmet Çağdaş Seçkin³ 



ABSTRACT

Textile companies must develop strategies to survive in today's competitive and volatile world and flexibly adapt these strategies to today's conditions. In this competitive environment, audits conducted by an internal or external auditor, especially at the request of the customer, enable companies to differentiate themselves from other companies. Audits performed by companies can be classified in various ways. In this study, audit types are classified into internal and external audits depending on the type of auditor. If the audit is conducted by a person working in a company with a contract of pay, it is an internal audit, if not, it is an external audit. The principles and benefits of the audit to the company are explained in the study. In this study, a simulation was prepared to investigate the effect of audit frequency, internal interaction and number of employees in the company on the improvement of the company. In the prepared simulation, relevant changes were processed according to internal parameters on daily basis. As a result, it was found that with sufficient recovery of all parameters within the company, the highest improvement success is achieved when the audit frequency is conducted at least every 3 months. Future studies, information on how the developing information technologies and artificial intelligence provide convenience and perspective for audits are presented.

Keywords: Audit, GMP audit, Quality management system, Environment management system, Remote auditing, Online auditing

ÖZ

Tekstil fabrikaları günümüzün rekabetçi ve değişken dünyasında ayakta kalabilmek için stratejiler geliştirmeli ve bu stratejileri esnek bir şekilde günümüz koşullarına uyarlamalıdır. Bu rekabet ortamında özellikle müşteri isteği üzerine iç veya dış denetçi tarafından yapılan denetimler işletmelerin diğer işletmelerden bir adım öne çıkmasını sağlamaktadır. Şirketlerin aldığı denetimler çeşitli şekillerde sınıflandırılabilir. Bu çalışmada denetim türleri denetçi türüne göre iç ve dış denetim olarak sınıflandırılmıştır. Denetim, bordro sözleşmesi olan bir şirkette çalışan kişi tarafından yapılıyorsa, iç denetim, yoksa dış denetimdir. İşletmeye yönelik denetimin ilkeleri ve avantajları çalışmada açıklanmıştır. Bu çalışmada denetim sıklığı işletme içi etkileşim ve işletmedeki çalışan sayısının işletmenin iyileştirilmesindeki etkisini incelemek amacıyla bir simülasyon hazırlanmıştır. Hazırlanan simülasyonda işletme içi parametrelere göre ilgili değişimler günlük bazda işlenmiştir. Sonuçta eğer işletme içi tüm parametreler yeterli bir biçimde toparlanırsa denetim sıklığının en fazla 3 ayda bir yapılması durumunda en yüksek iyileşme başarımları elde edildiği görülmüştür. Gelecek dönem çalışmaları, gelişen bilişim teknolojileri yapay zekanın denetçiye denetimler için nasıl kolaylık ve bakış açısı sunduğu konusunda bilgiler sunulmuştur.

Anahtar kelimeler: Denetim, GMP denetimi, Kalite yönetim sistemi, Çevre yönetim sistemi, Uzaktan denetim, Online denetim

¹Balta Group, Uşak, Turkey

²(Lecturer), Usak University, Karahalli Vocational School, Department of Office Management and Administrator Assistantship, Uşak, Turkey

³(Asst. Prof.), Adnan Menderes University, Faculty of Engineering, Computer Engineering, Aydın, Turkey

ORCID: M.S. 0000-0002-9564-1534
F.K. 0000-0003-4413-5188
A.Ç.S. 0000-0002-9849-3338

Corresponding author:
Mine SEÇKİN
Balta Group, Uşak, Turkey
E-mail address: mine1seckin@gmail.com

Submitted: 24.10.2020
Revision Requested: 07.01.2021
Last Revision Received: 25.01.2021
Accepted: 02.02.2021
Published Online: 31.05.2021

Citation: Seckin, M., Koc, F., Seckin, A. C. (2021). Audits in a textile company and advantages. *Acta Infologica*, 5(1), 99-116. <https://doi.org/10.26650/acin.815934>

1. INTRODUCTION

In today's world, the relationship between the customer and the producer is not only supply and demand, but also the two elements develop each other according to constantly changing needs. For the development of companies and product quality, every step from the source to the final product must be monitored and controlled. These monitoring and inspections can be carried out by machines (Erdoğan, 2019) and/or people. Control by machines is currently only at the data collection level. The main workforce under human control. A Human audit is divided into two types as internal and external. These control types are explained in detail in the second part.

Today, the phenomenon of globalization and the changes in the social sense that are reflected in business life and in the product and consumer-oriented marketing periods of the old days are being replaced by the period of social responsibility. It was during this period that the need for consumer decisions and needs and the interests of the economy and the interests of society emerged. The relationships that companies establish with their social stakeholders are important, and they must take into account both the interests of social stakeholders and their interests. Otherwise, the company's reputation, and thus the desired profit, will decline in the future. All segments of the company, such as employees, shareholders, society, suppliers and customers, can be listed as the company's social stakeholders. The concept of social responsibility is based on the consideration of the interests of these social stakeholders and their interests in the company's decisions. The changing world order has led to a change in the wishes of consumers, employees and all other social stakeholders.

Consumers expect companies to produce high-quality products and show interest in social issues. While carrying out these production activities, customers do not want to pollute the environment. They also want to contribute to the environment (Delmas, 2001), produce recyclable products. We see that consumer demands have changed in Maslow's hierarchy of needs (Kula ve Çakar, 2015). While the society that initially focused on consumption with emotional elements in the early days then took into account the rational elements in consumption, today's consumer is increasingly realizing the consumer activity by attaching importance to spiritual and moral values (Bozkurt, 2013).

Corporate social responsibility in a textile company must first begin at the workplace. Conditions must be created that should be present in the workplace and employees must be valued. Then the company should be aware of its responsibility under market conditions. Responsibility towards consumers should be known and there should be good relations with suppliers. Finally, the company should pay attention to environmental and social issues. The company should minimize damage to the environment and increase its contribution, be committed to human rights and be intertwined with society.

In the literature , it has been observed that it is beneficial to use big data analysis techniques in accounting and financial auditing, in the measurement of financial crises in enterprises, in detecting financial frauds, in stock market predictions, but the use of the technology is a little behind (Kılıç, 2019). In the development of information technologies, it has been observed that software programs and processes that take days can be completed in a much shorter time. Problems with mathematical calculations that require a lot of effort , accumulated paperwork or notebooks could be avoided (Aslan ve Kızıl, 2019).

The emergence of COVID-19, in addition to the existing audits that are already required by regulations, laws or urgent needs, and the worldwide travel restrictions, has led to a focus on alternatives that can be used immediately instead of the traditional face-to-face audits. Remote auditing can be a good saving alternative, especially given that most companies limit their travel to business-critical functions and many countries around the world are temporarily closing their borders. Unlike on-site inspections of records, remote audits do not allow simultaneous investigations. During the remote audit, the auditor should take notes and write down questions to be asked during remote interviews. In addition, a video telephone conference call may be organized in which the responsible supervisor and the person responsible for the implementation of the revised program can participate so that questions can be asked simultaneously during the remote review. In this way, documents can be exchanged and reviewed in real-time, and the feasibility of audits should be improved by asking questions and receiving answers at the same time. In the last part of this article, a simulation has been prepared for the audit. For companies, the effects of parameters such as audit frequency, internal communication and the number of personnel in the business have been simulated. The study consists of 9 parts. Introduction in part 1, Audit definition and historical development in part 2,

Audit purpose and process in part 3, Principles and characteristics of audit in part 4, Audit types in part 5, Points to be considered by auditors in Covid-19 process in part 6, Audit simulation part 7, Audit simulation results and discussion in part 8, Conclusion and future work in part 9.

2. Audit

The origins of auditing are rooted in history, and when people start living in society, control proves to be an indispensable phenomenon. An audit is an examination of the activities and operations of an organization to determine whether they are in accordance with the objectives and rules set out in advance. It is a process in which findings are made and reports are written as a result of the audit. Audit in general; setting standards to ensure that the results of the activity are as relevant as possible to the plans, comparing the results with these standards and identifying corrective action where practices are separate from the plan (Bozkurt, 2013).

Audit, which is one of management's functions, has important functions for the success of organizations. It follows the planning, organizational, steering and coordination functions of management. Auditing provides an assessment of the results of combining these four functions and reveals new strategies, policies and even a new management philosophy. In this way, it is an important tool that provides feedback on management activity and contributes to the dynamic structure of the existing organizational structure. In this respect, auditing is an activity that is the responsibility of management. Without a functioning audit mechanism, organizations rarely perform their full duties. Therefore, the audit is a must. It comes from necessity.

In order for the audit, which is the responsibility of management to add value to management, it should be expected to operate in accordance with established standards and criteria. The audit is part of the management process. Without the audit, the management process is incomplete. The audit is an important requirement for institutions. If the audit is carried out correctly, it is an important function for institutions, and if it is not carried out correctly, it can cause considerable damage.

Information of the approach of audit, the purpose of the audit and the development of the parties to the audit in history are presented in Table 1 below from the pre-industrial revolution to post 1990.

Table 1

Development of Audit in History

Period	Approach Based on Auditing	The Purpose of The Audit	Related Party in Audit
Before the industrial revolution	Auditing of all documents and records	Detecting mistakes and tricks	Operating owners
The 1900s	Auditing of all documents and records	Detecting mistakes and tricks	Partners and credit institutions
1900-1930	Auditing of all documents and records /audit through sampling	To determine the accuracy of financial statements	Partners, credit institutions and the government
1930-1960	Auditing financial data by sampling	Providing positive and negative opinions about the accuracy of financial statements	Shareholders, credit institutions, government, unions, consumers and other related segments
1960-1990	Audit through sampling /The development of statistical sampling	To determine the accuracy of financial statements	Partners, credit institutions, government, unions, consumers and other related segments
After 1990	Risk-oriented audit Auditing of data processing systems	To determine the accuracy of financial statements	Partners, credit institutions, government, unions, consumers and other related segments

Reference: (International Standard On Auditing, adapted from 2009a).

When the above table which presents the development of the audit in history is examined, it is seen that the main purpose of the audit is to reveal errors and tricks based on all documents and records regarding businesses and institutions. Also, in these years auditing activities were made on behalf of business owners, partners and credit institutions. In the period from the 1900s to the 1930s, the review of all documents and records was replaced by auditing financial data. In this period as in the previous period, it was aimed to prove the accuracy of the information presented in financial statements and its annexes and importance to create an audit report including the opinions of the auditors in this regard was gained. These audit reports

created by the auditors provide information for shareholders, credit institutions, government, labour unions, consumers and other interested parties. In the period between 1960-1990 audits provided a sampling of statistically selected noun needed here from all data of businesses and institutions. In the post-1990 period, the adoption of risk-based auditing and information processing systems can be observed. Despite this difference in audit understanding in both periods, in relevant periods the main purpose of the audit is to determine the accuracy of the information presented in the financial statements and its annexes and present the results of the audit activity to business-related interest groups.

In addition to this information presented in the table above, Independent Audit Regulation dated 26.12.2012 with Turkey Auditing Standards including information systems auditing, it was decided to comply with international standards in the field of independent auditing (Independent Auditing Regulation, 2012). Therefore, it is possible to say that the importance of information systems auditing has increased in order to determine the accuracy of financial statements in accordance with international auditing standards since 2012, when international auditing principles were adopted in auditing. In addition, the parties in the information age where instant possibility access to today's information includes not only business interest groups but also other users of information outside the country.

3. Aim of Audit and Audit Process

It is concluded that the purpose of the audit is necessary to find what is better and more useful than the current situation, to eliminate the differences in implementation, to detect the negative aspects and to make the right decisions (Kebeli, 2012; Bozkurt, 2013).

The objectives of the audit are: to ensure continuity, prevention, and limitation, correction, reforming according to the recommendations made as a result of the audit, to ensure efficiency and effectiveness, to make the various decisions, to seek better and more useful than the current situation within the audited unit, to determine the differences between the application, to take measures and to reveal the negativity in practice (Kebeli, 2012; Bozkurt, 2013). Given the objectives, the purpose of the audit is to provide a generally functional and accountable management approach.

In line with the objectives described above, the process that the inspection should follow can be compared to a building, as shown visually in Figure 1 below. A solid structure must have a solid foundation. Establishing a solid structure in auditing is possible with the determination of audit principles by the audit team on a solid basis. Customer acceptance procedures are like a gate of the building in this structure. At the beginning of the audit process, the enterprise that will be subject to audit first is required to collect the necessary information about the business and assessment of the admissibility of the audit in terms of risk. After that, the customer acceptance process studies are started by signing an audit contract between the parties. In the first stage of the audit process, it is proposed to know the business better by planning. In the following process, evidence collection techniques for business and business operations are applied within the established planning. After the evaluation of collected evidence about the business and completing the identified shortcomings the stage of reporting the results is passed. In the reporting stage, which constitutes the last stage, the auditor's report is prepared based on the evidence found. The audit process in the company ends with the completion of the quality control procedures, which can be likened to the roof of the building (Şavlı, 2012).

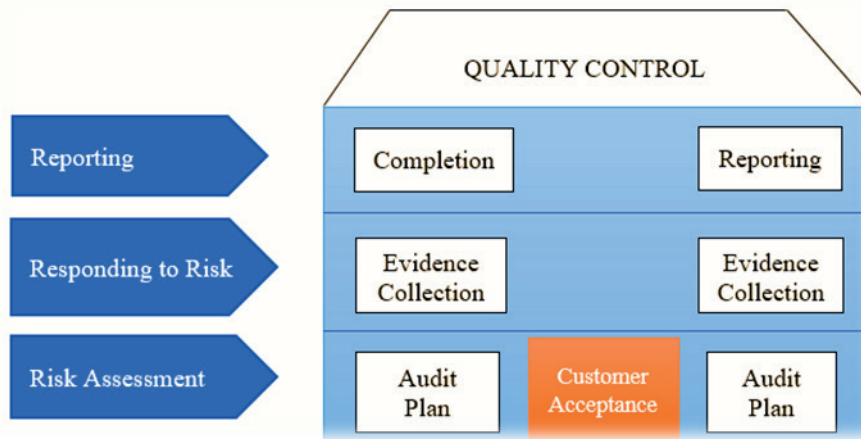


Figure 1. Audit Process (Şavlı, 2012).

4. Principles and Characteristics of Audit

In order for the audit to perform the functions expected of it and make a positive contribution to the functioning of the system, the organization and operational view of the audit, which we can express as audit management, should be parallel to the functions to which it is loaded (Kuluçlu, 2006).

In order to achieve the expected benefit, the audit must have the following qualifications:

- The objectives of the audit should be clearly defined,
- There/It Must be a reliable and valid audit system (Akyel ve Köse, 2010).

Audit and audit reports should be appropriate for the period. When planning the audit period, the planning should be such as to ensure that the personnel involved in the audit work to success with maximum efficiency in achieving the audit objectives and targets. The duration of the audit should be kept as short as possible, but it should include sufficient time to achieve the audit objectives. In addition, appropriate time should be scheduled for the audited entities.

Audit reports must be written within a short time after the audit to ensure that the non-conformities identified in the reports are corrected by management and that the proposals are implemented. The audit should be flexible and should not restrict the alternative solutions proposed by management. The audit process and the report produced at the end of the audit should be understandable. If the audited service does not have a good understanding of the audit process, it cannot assist the auditors at the expected level in assessing management risks and the effectiveness of controls to prevent those risks.

The fact that the audit reports are not understandable leads to the risk that the recommendations are not implemented due to management's lack of understanding of the findings and recommendations. The audit should have criteria. The absence of criteria can lead to an inaccurate assessment of risk controls. The audit should be strategic. Instead of focusing on strategic areas, it is a misunderstanding of the audit to remain in arbitrary areas based on random assessments or without criteria. This leads to the idea that the audit is an unnecessary activity rather than the expected benefit. The auditor should be constructive and anticipate corrective action. The audit should not only uncover and criticize nonconformities, but also analyze the causes of errors and provide proactive solutions to prevent recurrence.

With the audit;

- It is used to determine the degree of success of organizations in system applications.

- It ensures that non-conformities cannot be detected by the organization itself.
- Corrective and preventive activities of the organization should be strengthened and maintained more effectively.
- The status of the process is analyzed and the points to be optimized are determined.
- Process and organizational improvements are provided with the provided improvement suggestions.
- As a result of all improvement activities, the efficiency of the organization is increased, costs are reduced, alternative costs are determined, and better products/services start to be offered.
- As a result of all these improvements, the organization gains prestige and trust in the market.

5. Audit Types

The types of audits can be classified according to various criteria. Within the study, audits are divided into two parts. These are internal audits and external audits.

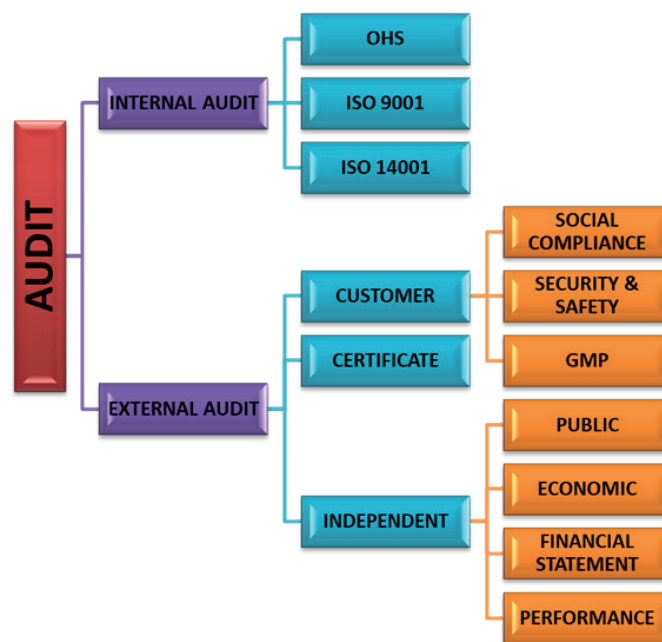


Figure 2. Audit types

5.1. Internal Audit

Internal audit is the examination of all business activities which are in the permanent status of the company or which offer a continuous status as internal auditor from outside the company. In other words, it is a type of audit carried out by the people who work in a company with a wage contract. Internal auditors are those who work within the organizational structure of the company and generally present the information they receive to the management of the company. The main tasks of internal auditors are to monitor the internal control structure of the institutions and organizations to which they are affiliated and to make proposals. The internal audit is an important tool for companies to examine the effectiveness of their activities and to show the extent to which the entity has achieved its objective. Internal auditors carry out their activities within the scope of their responsibilities as defined by them under the Board of Directors. In addition to this information about internal audit and internal auditors, it can be said that the number of internal auditors is also important to ensuring its effectiveness in the internal audit. However, when table 2, in which the number of cadres allocated for internal auditing in public administrations was presented in our country in 2018, is examined, although the number of staff assigned to internal auditors is 2.054, only 880 of this number are employed and the other 1,174 staff are not employed. Therefore, the occupancy rate of internal auditors is only 43.03 percent. This rate in the number of internal auditors employed also negatively affects the effectiveness of internal audit in our country (Ministry of Treasury and Finance, 2021).

Table 2

Internal Auditor Staff in Republic Administration

Total Cadre	Full Cadre	Empty Cadre
2.054	880	1.174

Reference: (Ministry of Treasury and Finance, Number of Full-Empty Internal Auditors Cadre, 2021).

The duties of the internal auditor are listed as follows (Regulation on Working Procedures and Principles of Internal Auditors);

- “To evaluate the management and control structures of public administrations on the basis of an objective risk analysis.
- To conduct investigations and make recommendations with a view to effective, economical and efficient use of resources.
- To verify compliance with laws by expenditure.
- To monitor and evaluate the consistency of the administration’s expenditure, decisions and savings on financial transactions with the objectives and policies, development plan, programs, strategic plans and performance programs.
- To audit the system of financial management and control processes and make suggestions on these matters.
- To make suggestions for improvement in the context of the audit results and follow up on them.
- To inform the highest superior of the respective administration if during the audit or after the audit a situation arises that requires investigation.
- To verify the accuracy of the information produced by the public administration.
- Assist in establishing performance indicators when deemed necessary by the most senior manager and evaluate the applicability of the established performance indicators.
- To inform the most senior manager of the findings of criminal situations “.

5.1.1. ISO 9001 Internal Audit

Considering the fact that consumers have quality, that their expectations of goods and services are increasing, and that they are difficult and not easy to meet, it can be said that quality plays a strategic role in achieving competitive advantage. In this respect it is a well-known fact that quality is one of the sines qua non of companies today. The literature shows that internal audits contribute to the revitalization of corporate performance (Alič ve Rusjan, 2010).

ISO 9001 Quality management System The aim is to record every process of the organizations published by the World Standards Organization in the production or service sector and to ensure customer satisfaction at all levels of management on the basis of continuous improvement. ISO 9001 internal audits are reviewed and reinforced by an internal auditor who has ISO 9001 basic training and a certificate as an internal auditor and knows whether or not he meets the requirements of the quality management systems of companies. They prepare and implement the studies that prove the existence of the substances of ISO 9001 standard in the company.

5.1.2. ISO 14001 Internal Audit

ISO 14001 Environmental Management System is the management system standard published by the international standards organization which identifies and describes the studies to be done to minimize the damages to the environment as a result of the interaction of the enterprises with the environment during their processes (Jiang ve Bansal, 2003). Because ISO 14001 Environmental Management System is a management system standard, a company can be established and applied alone, and because it has common clauses with ISO 9001 Quality Management System standard, it can be installed and implemented in integration with management system standards such as iso 9001 standard and OHSAS 18001 standard. Businesses that establish and implement the ISO 14001 Environmental Management System should certify this standard with ISO 14001 certification. ISO 14001 internal audits are checked and reinforced by an internal auditor who has received ISO 14001 basic

training and internal auditor certificate and checks whether or not the company meets the quality management system rules. ISO 14001 standard in the enterprise shows that the existence of substances in the establishment of the documentation set up and implement the application.

5.1.3. OHSAS 18001 Internal Audit

OHSAS 18001 is an international standard that determines the requirements for the occupational health and safety management system (Fernández-Muñiz, Montes-Peón ve Vázquez-Ordás, 2012). OHSAS 18001 has been developed in compliance with ISO 9001 (QMS) and ISO 14001 (EMS) standards to enable organizations to integrate quality, environmental and occupational health and safety management systems. Compliance with national health and safety standards is mandatory for all organizations. Occupational Health and Safety Management System is a systematic and scientific way of identifying hazards and risks and taking precautions against these hazards and risks to protect them from accidents and other effects that may cause health damage due to various reasons and to provide a better working environment during the realization of work at workplaces. The company is constantly monitored by the occupational safety advisor in terms of occupational safety who reports to senior management.

5.2. External Audit

This is an audit conducted by audit staff and institutions ,which are not related to the functional and organizational hierarchy of the audited company and which is independent of it, to express opinions on the legality, regularity of the operations of the related company and efficiency, frugality and efficiency of management. The external audit may be the customer audit with the customer request, International Organization for Standardization (ISO) 9001 and 14001 and others certificate audits and public audits.

5.2.1. Customer Audit

Customer audit activities, which are performed by an auditor on behalf of the audit firm whether or not for certification purposes, at the request of customers. The audit is carried out by an impartial audit company to determine the system implementation level of the organization, nonconformities, the effectiveness of corrective and preventive activities, the status of its processes, and the points that need to be improved. In particular, audits made by customers of a third party; Social compliance audits, security and safety audits, and Good manufacturing process (GMP) audits.

A social compliance audit is carried out under the name of social responsibility audit or social compliance audit to determine and improve the working conditions of the manufacturers in the supply chains of big brands. The globalizing world and rising consumer awareness are based on the determination and evaluation of large companies' suppliers. International Labor Organization, local and national units are published in the official newspaper. Laws, regulations, the rules and regulations specified in the regulation are based on scoring the degree of compliance. Local statutory requirements are audits that are assessed and reported in line with customer expectations. The Code of Ethics includes controls to ensure assurance to entities requesting audit and having commercial relations with the audited entity.

In social compliance audits, the auditor checks the code of conduct of the company. These are;

- Forced labor,
- Payments,
- Working hours,
- Child labor,
- Discrimination,
- Employee Awareness,
- Prevention of Harassment and Abuse

- Rest Days and Holidays
- Labor Discipline

Occupational Health and Safety: It considers human beings as the most valuable asset in all branches of activity and considers it a primary goal to minimize any loss of life and property by providing a healthier and safer work environment. It organizes pieces of training to ensure that all employees are aware of their responsibilities and provides the necessary personal protective equipment for occupational health and safety. Occupational Health and Safety aim to comply with local laws and regulations within the scope of which all personnel can work under the Occupational Health and Safety policy. Occupational health and safety risks, environmental measurements, periodic checks, electrical internal installation compliance reports, and electrical ground measurement and controls should be identified. It aims to keep these risks under control by reducing them to acceptable levels.

An evaluation score is given to the company after the record controls. After the audit, the company is given a CAPA (Corrective and Preventive) report in which nonconformities are written. The audited company sends a report to the impartial audit firm with a description of the dates and activities for these nonconformities. Finally, the report is presented to the customer by the audit firm.

GMP is a quality system that guides the manufacturers in terms of the conditions under which GMP standards should be produced in many sectors with good manufacturing practices. The use of the GMP standard is used in particular in the United States as cGMP. The letter “c” used at the beginning of the GMP standard stands for “Current”, meaning “dynamic. The GMP standard is not a one-off system. It is a production technique that is continuously controlled and kept under control, correcting the systemic or operational deficiencies that arise during these controls and developing based on the innovations brought by today’s technology and sectoral studies. Topics covered in the GMP audit are Management commitment and continual improvement, Risk management, Quality management, Site and facilities management, Product control, and Personnel training. (Figure 3.)



Figure 3. Topics examined in GMP audit

Particular points a GMP auditor pays attention to; how and where the records are kept, training procedure, whether adequate training is provided, and the calibration of test devices.

5.2.2. Certification Audit

Certification audit is the process of issuing the certificate which is valid for 3 years as a result of the certificate renewal audit after the expiry of the ISO 9001 and 14001 certificates issued for 3 years, provided that the enterprises carry out an audit

every year in accordance with the accreditation rules. Obtaining an ISO certificate is not a document nor is it sold on the shelf. First of all, a company should establish the documentation showing the existence of the certificate standard in its business as ISO 9001,14001. Upon this, the company must determine an accredited organization and apply for ISO 9001 Certification and remove any major or minor errors in the system that occur as a result of this certification audit and as a result, obtain ISO Certificate from the ISO Management system certification.

5.3. Independent Audit

An independent audit is an audit performed by independent auditors who are not affiliated with the institution it audits. The concept of independent audit has been included in Law No. 6102 which entered into force on 01.07.2012 for the first time in our commercial law. Independent audit means the audit and evaluation of the financial statements and other financial information following the financial reporting standards by applying the required independent audit techniques to obtain sufficient and appropriate independent audit evidence, to provide reasonable assurance and to be evaluated and linked to the report. Thus, a complete standard will be provided in the financial statements of the companies and the comprehensibility of the financial statements will be increased. An independent audit is carried out by external auditors who are not loyal to the institution they audit.

The following are the basic characteristics that the independent auditor should have (Cömert ve Uzun, 2014);

- Impartiality, honesty and authority to the extent that the duty requires,
- Work independently and have reporting responsibilities,
- Evaluate internal controls, risk management processes, and corporate governance processes as they relate to the company, and perform the audit of the financial statements,
- Comply with the requirements of laws and standards,
- Protect the interests of all stakeholders and interest groups associated with the Company,
- Serve the public interest (Cömert ve Uzun, 2014).

5.3.1. Public audit

These are the audits carried out by the inspectors of public institutions to ensure public order and protection of public rights. Public audits are carried out by auditors affiliated to the ministry and other public organizations based on the provisions of the law (Korkmaz, 2007). Some ombudsmen performing ombudsman activities can be listed as follows (Independent audit under international audit standards, 2009b).

- Ministry of Finance auditors,
- Ministry of Finance account experts,
- Capital Markets Board (CMB) auditors,
- Banking Regulation and Supervision Agency (BRSA) auditors,
- Social Security Institution (SGK) inspectors.

5.3.1. Economical audit

Financial statements are a type of audit in which compliance and performance audits are conducted and reported together. Within the scope of the economic audit; on the one hand, it is determined whether the activities of the organization are carried out under the legislation in force, and on the other hand, the compliance of the financial statements with the current financial legislation and generally accepted accounting standards are made through a detailed examination, taking into consideration all operational functions of the resources allocated to the organization, efficiency, and is used following the principles of frustration measures and makes recommendations in this direction.

5.3.2. Financial Statements audit

The department that is directly related to the activities and results of an enterprise is not only the owner or partners of that enterprise. The source of information about the activities and results of the enterprises are the accounting records and documents of that enterprise and the financial statements prepared based on these enterprises. The purpose of the audit of financial statements is to determine the irregularities that would prevent the realization of the objectives expected in these statements or in other words, to determine the reliability of these statements by measuring the degree of compliance with the principles, accounting procedures and principles, laws and all other predetermined criteria (Dalak, 2000). This audit is performed by independent auditors and ombudsmen. In this context, the general elements of the financial statements audit are presented in Table 2 below.

Table 2

Elements of the Financial Statement Audit

Aim	Examination of the accuracy and reliability of the information presented in the financial statements
Content	financial statements and annexes
Criterion	Generally accepted accounting principles and other regulations affecting financial information
Reference:	(Independent Audit Under International Audit Standards, 2009b).

5.3.3. Performance audit

This audit comprehends all levels of management in terms of their frugality, efficiency, and effectiveness in their programming, implementation, and monitoring stages. It covers the review of the implementation of procedures and methods associated with these activities to assess the efficiency and effectiveness of an organization's activities (Kubali, 1999 ; Öztürk, 2014).

The following are the purposes of performance auditing:

Kubali et.al explained that the primary purpose of performance auditing is to provide the state with information on the use of resources and to make recommendations to managers by providing impartial and independent information. The second purpose of performance auditing is to encourage the audited entity to take the necessary steps to improve its systems and controls to achieve maximum efficiency with the resources spent. The main objectives of performance auditing announced at the 12th Congress of SAI (Supreme Audit Institutions) in Sydney in April 1986 are as follows;

- Improving public administration,
- Providing reliable information to the legislature and the public on the results of the administration's work,
- Encouraging the administration to evaluate its own performance,
- Ensuring that the chain of accountability is extended (Kubali, 1999).

6. Points to Be Consider By Auditors in the Covid -19 Period

During the Covid 19 pandemic, auditors need to access more information and documents about the company and its environment to assess the impact of the 2020 financial statements on the economy. In this context, the following are the items that auditors should focus more on during the Covid-19 process (Demir, 2020).

6.1. Risk Assessment and the determination of Material Misstatement Risks by knowing the company and its environment "ISA 315 regarding the determinations to be given to the Assessed Risks and the valuation"

The Independent Auditing Standard (ISA) 315 governs the auditor's responsibility to identify and evaluate the risks of material misstatement in the financial statements in order to familiarize the auditor with the entity and its environment, including the entity's internal control (Independent Auditing Standard-315).

As part of this auditing standard, the auditor should assess how financial and operational actions should be taken as a result of epidemics such as Covid-19. In this context, the auditor should also evaluate a possible disruption in the companies' relationships with customers and suppliers and the risks that may be caused by a disruption (Türmob, 2020b).

6.2. “ISA 500 Independent Audit Evidence” at Independent Audit Evidence

In disclosing audit evidence, ISA 500 addresses the auditor’s responsibility to plan and perform audit procedures to reach appropriate conclusions. This auditing standard is also applicable to any audit evidence obtained during the audit (Independent Auditing Standard-500).

During the Covid 19 epidemic process, constraints such as curfew and transportation should be considered by the auditor, and in this direction, data sharing should be enabled by organizing meetings in a digital environment (Türmob, 2020b).

6.3 Regarding independent audit evidence, “ISA 540 Independent Audit of Accounting Estimates and relevant disclosures, including those with fair value”

ISA 540 sets out the auditor’s responsibilities in relation to estimates and disclosures in the audit of financial statements. In addition, there are provisions and explanations within this standard for assessing misstatements of accounting estimates and indicators of potential management bias (Independent Auditing Standard-540).

Since the Covid 19 epidemic process creates a great predictive difficulty worldwide, it is of great importance that the auditor uses his professional judgement during the epidemic process by reviewing all accounting estimates of the company (Türmob, 2020a).

6.4 Regarding the Use of Third Party Work “570 Business Continuity”

ISA 570 sets out the auditor’s responsibilities in relation to the continuity of the entity and the impact of going concern on the auditor’s report (Independent Auditing Standard-570). In this context, the financial statements are prepared on a going concern basis. Therefore, unless the manager intends to cease the economic activity of the entity, the manager is responsible for acting on the going concern basis. However, financial statements prepared for specific purposes may or may not be prepared on a going concern basis of accounting. Where it is appropriate to apply the going concern principle, assets and liabilities are arranged on the basis that the entity will meet its assets and liabilities from which it will benefit (Independent Auditing Standard-570).

It is understood that Covid-19 may have a significant destructive impact on the financial structures of companies. Therefore, companies may have difficulty meeting their debt obligations. At this point, management should classify short-term and long-term financial liabilities and determine which liabilities have been affected by the epidemic process and how this effect affects the continuity of the company (PWC, 2019). However, the impact of sudden fluctuations in exchange rates during the epidemic process on the financial statements and business continuity should be evaluated. If the auditor finds any uncertainty about the continuity of the business during this process, he should not refrain from issuing a qualified opinion (Türmob, 2020a).

6.5 Findings of Independent Audit and Reporting “Responsibilities of 720 Auditors in Relation to Other Information.”

ISA 720 addresses auditor’s responsibilities in relation to information other than the auditor’s opinion included in an entity’s annual report. In this context, companies’ annual activity reports may consist of a single document or a combination of several documents that serve the same purpose (Independent Auditing Standard-720). Under this auditing standard, a material misstatement of financial statements is considered to exist when other information is materially inconsistent with the financial statements or with information obtained by the auditor during the audit process. Therefore, in such cases, when the reliability of the auditor’s opinion could be affected, the auditor is required to consider other information (Independent Auditing Standard-720). During the Covid-19 outbreak, the auditor should also assess whether the information in the annual report prepared by management is consistent with the data presented in the companies’ financial statements, whether the annual report provides explanatory information to the parties that require it, and whether the relevant report fully presents the objectives and risks of the company’s future (Türmob, 2020a).

7. Audit Simulation

Process changes in industrial facilities will cause both time and cost. These processes are first tested with simulations or small-scale applications. In this part of the paper, the effect of the number of employees in a company, the frequency of audit, the rate of interaction between employees on the company's recovery time, the number of employees fired and the number of employees recovered is examined. The simulation is prepared in Python programming language with the `simpy` library (Matloff, 2008). The simulation works on a daily basis. Audit frequency is the period of audit of the company by internal or external auditors. These parameter values were chosen as 4, 8, 13, 26 and 52 weeks. The number of employees working in the company was chosen as 50 and 100. Interaction between employees refers to how much continuous employees in the company interact with each other on a daily basis. This value states that when employees don't conform, others see and are affected by this nonconformity and spread the nonconformity. The spread of nonconformity is defined as a random process and the probability of spreading nonconformity when two employees interact is assumed to be 50% in the simulation. In the simulation, each employee is run as one thread. The probability of self-nonconformity of employees was set as 10%. Nonconformist employees are detected and warned by the person responsible for the audit. The probability of an unsuitable worker overlooked from the audit was set to 10%. The probability of recovery for the warned employees was assumed to be 0.8. If the employee has not recovered after 2 inspections, they will be fired. Some screenshots taken during the simulation are presented in Figure-4.

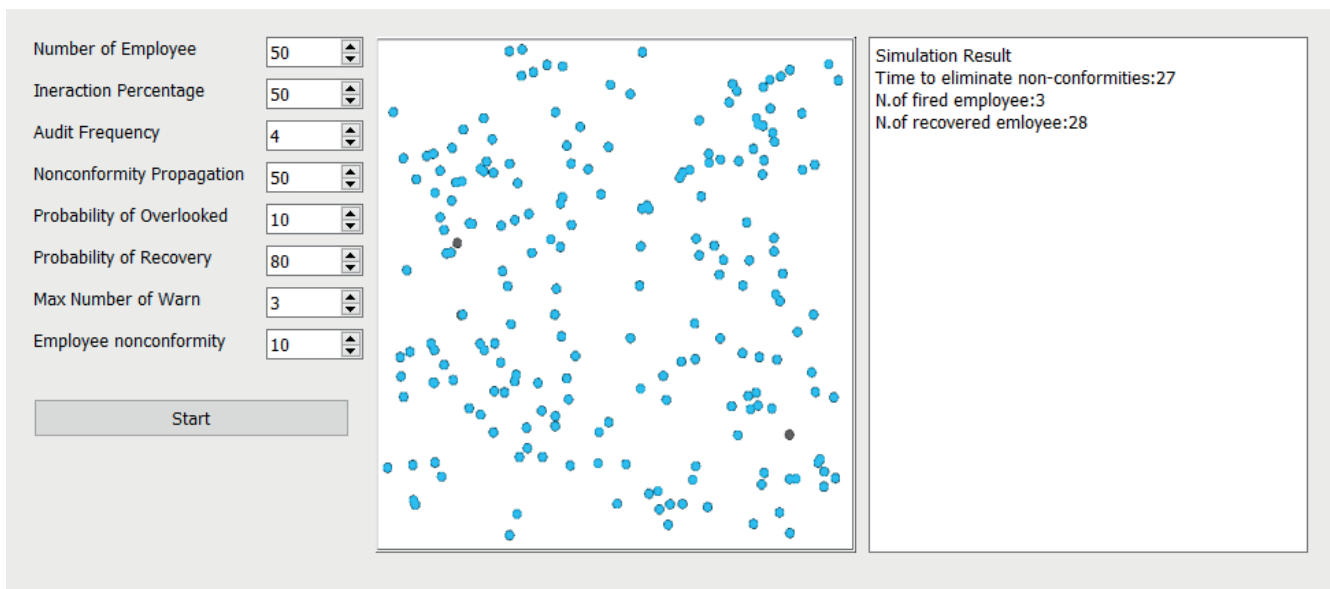


Figure 4. Audit Simulation User Interface

8. Audit Simulation Results and Discussion

The first parameter examined was the duration of elimination of nonconformities in the firm by changing the audit frequency in the simulation. The second parameter studied in the simulation was the impact of audit frequency on the employees in the company. The number of employees who did not cover their compliance was determined by changing the audit frequency. Finally, the simulation changed the interactions of the employees in the company and examined the effect of non-conformities on closing time. The number of employees recovering is shown in Figure 5. Changing the frequency of the audit or changing the interaction during the simulation generally did not significantly change the number of employees who recover. It can be seen that the standard deviation is only larger when the audits are tightened.

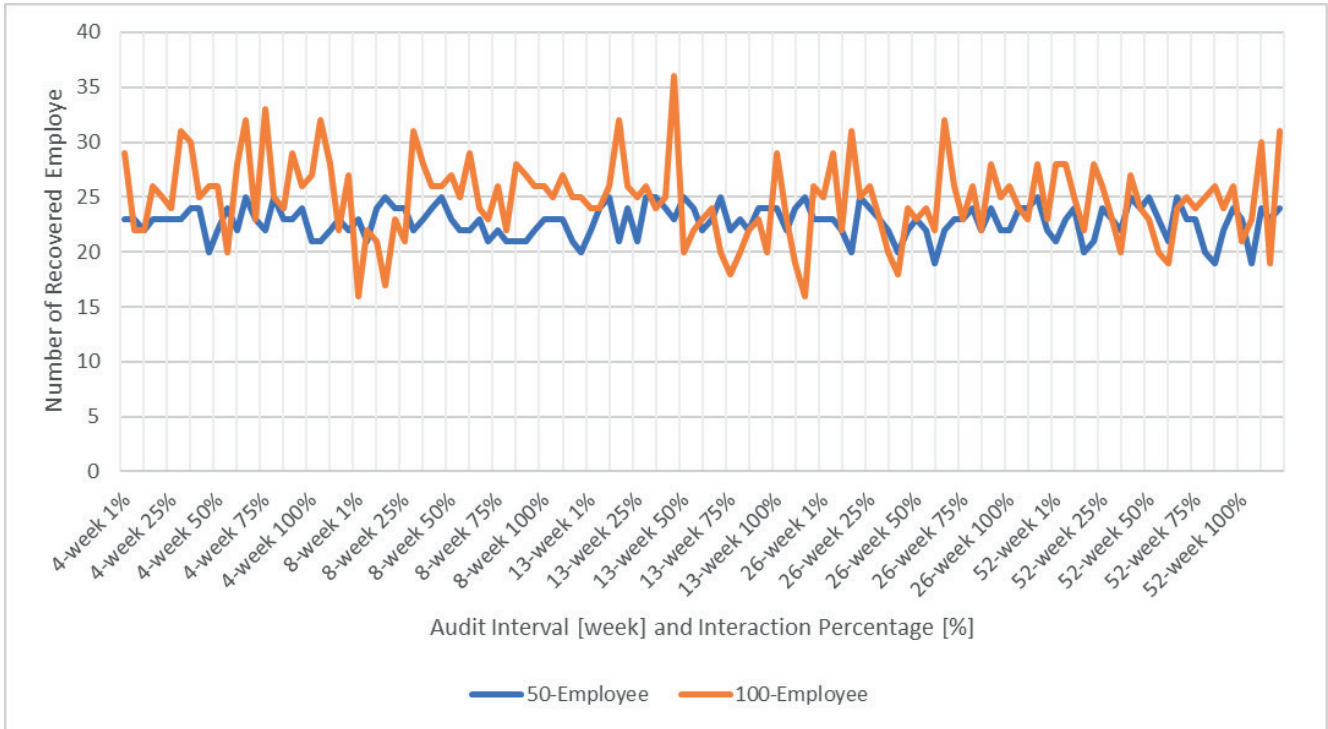


Figure 5. Audit Simulation on Number of Recovered Employee

Figure 6 shows the change in the number of fired employees. Increasing the frequency of audit or increasing the interaction does not lead to a significant change in the number of fired employees. It can only be said that as the number of employees increases, the number of firings also increases.

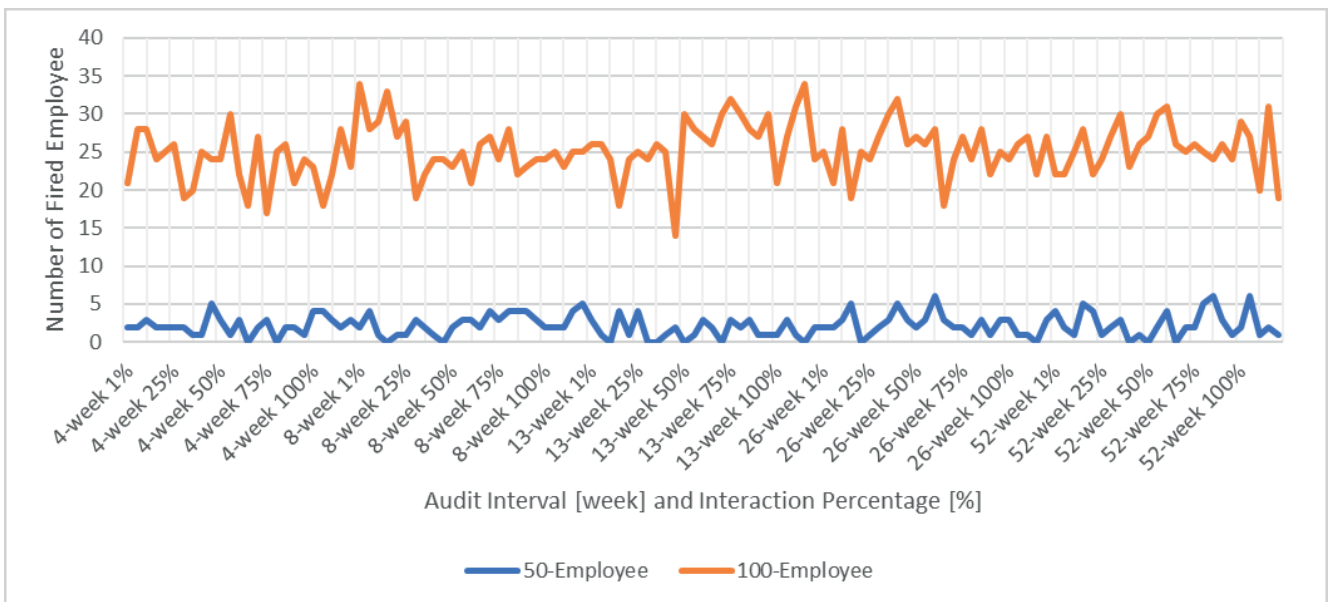


Figure 6. Audit Simulation on Number of Fired Employee

Figure 7 shows the recovery times of the company. As can be seen, recovery from nonconformities is delayed as interaction within the company increases. The company's recovery time suggests that the interaction should be lower and the frequency of audits higher. In 13-week audits in companies with fewer employees, it was observed that even with 75% interaction, the nonconformities in the company can be eliminated. However, irrespective of the number of employees, similar behavior is

observed when audits are conducted semi-annually or once a year. The graph shows that if audits are conducted at least every 3 months, success can be achieved even in companies with high employee interaction.

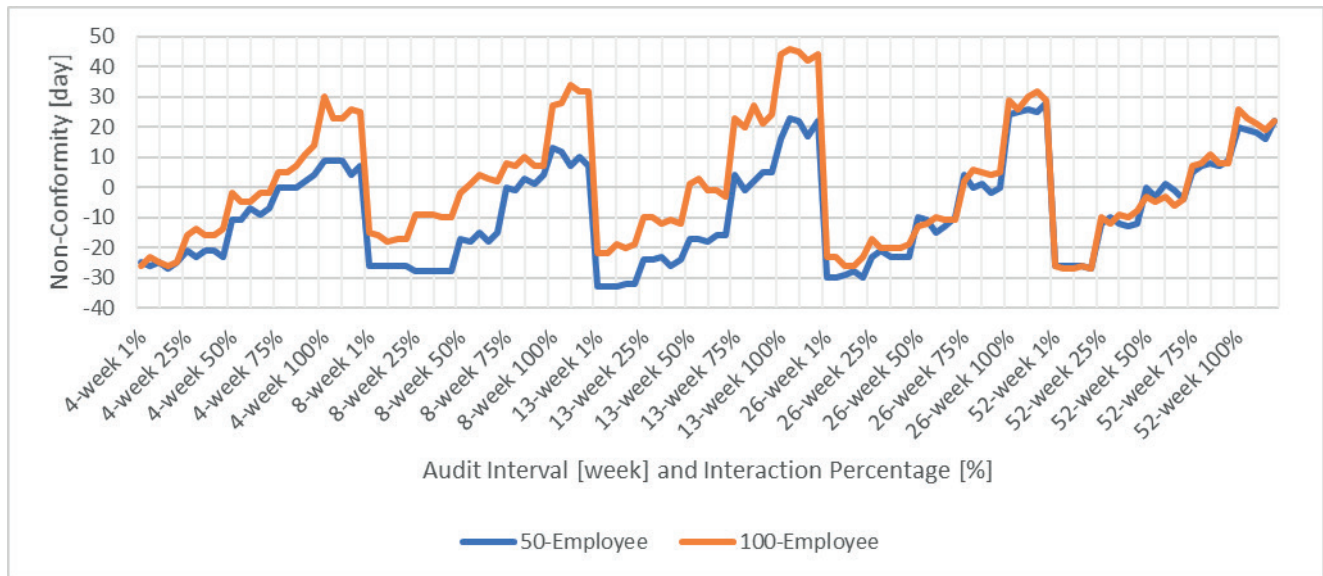


Figure 7. Audit Simulation on Elimination of Non-conformity

9. Conclusion and Future Works

Textile companies need to develop competitive strategies and flexibly adapt these strategies in line with today's conditions to survive in today's competitive and volatile world. The concept of the audit is evolving in many countries and the focus of the audit approach is changing. The audit approach in Turkey is following a parallel course with global developments. The importance and benefits of auditing in a textile company are becoming more and more noticeable. The audit, which may be encountered in different ways in terms of its name, structure, and location, has an important function to achieve the goals and objectives determined by the management. Changes in the audit approach to better fulfill this function are closely followed in many countries and Turkey. In this study, a simulation was prepared to investigate the effect of audit frequency, internal interaction and number of employees in the company on the improvement of the company. In the prepared simulation, relevant changes were processed according to internal parameters on a daily basis. As a result, it was found that with sufficient recovery of all parameters within the company, the highest improvement success is achieved when the audit frequency is conducted at least every 3 months.

In a textile company, data storage and processing are carried out through computers by greatly reducing the labor force benefiting from automation and digitalization. With technological advances, it is tried to facilitate information control at the optimum level in the companies. Various artificial intelligence applications are available in textile enterprises (Shamey ve Hussain, 2003; Karthikeyan ve Sztandera, 2010; Seçkin, Seçkin ve Coşkun, 2019). Data processing in the computer environment has given the information using a great opportunity to diversify the information. In this case, the data traffic is more diversified and concentrated. The audit of this big data is becoming more difficult than before. At this point, the auditor focuses more on analytical audit methods than before. As an auditor, if we look at the process of digitization and automation from another perspective, this can be a great advantage for an auditor. Since the records and documents are in the digital environment and the algorithms are on the system, an audit can be performed with the support of artificial intelligence and can increase the level of assurance (Mert ve Tas, 2019). Cognitive technologies can be considered to reduce or possibly eliminate the need for auditors. Cognitive technologies have indeed enabled the automation of manual tasks such as stock counting or drafting communications for decades. However, this will allow the auditor to focus more on focusing on risky areas than on habitually focused areas. However, cognitive technologies can not only automate routine tasks but can also enhance the professional judgment of the auditor by modeling thought processes that can be compared to the initial results.

With the developing information technologies, recently especially due to COVID 19, remote auditing can be a good saving alternative, especially given that most companies limit their travel to business-critical functions and many countries around the world are temporarily closing their borders. A video-telephone conference call may be organized in which the responsible supervisor and the person responsible for the implementation of the revised program can participate, so that questions can be asked simultaneously during the remote review. In this way, documents can be exchanged and reviewed in real time, and the feasibility of audits should be improved by asking questions and receiving answers at the same time.

Peer-review: Externally peer-reviewed.

Conflict of Interest: The authors have no conflict of interest to declare.

Grant Support: The authors declared that this study has received no financial support.

Author Contributions: Conception/Design of Study- M.S., F.K., A.Ç.S.; Data Acquisition- M.S., F.K., A.Ç.S.; Data Analysis/Interpretation- M.S., F.K., A.Ç.S.; Drafting Manuscript- M.S., F.K., A.Ç.S.; Critical Revision of Manuscript- M.S., F.K., A.Ç.S.; Final Approval and Accountability- M.S., F.K., A.Ç.S.

Hakem Değerlendirmesi: Dış bağımsız.

Çıkar Çatışması: Yazarlar çıkar çatışması bildirmemiştir.

Finansal Destek: Yazarlar bu çalışma için finansal destek almadığını beyan etmiştir.

Yazar Katkıları: Çalışma Konsepti/Tasarım- M.S., F.K., A.Ç.S.; Veri Toplama-M.S., F.K., A.Ç.S.; Veri Analizi/Yorumlama-M.S., F.K., A.Ç.S.; Yazı Taslağı- M.S., F.K., A.Ç.S.; İçeriğin Eleştirel İncelemesi-M.S., F.K., A.Ç.S.; Son Onay ve Sorumluluk- M.S., F.K., A.Ç.S.

References/Kaynaklar

- Akyel, R. ve Köse, H. Ö. (2010). Kamu Yönetiminde Etkinlik Arayışı: Etkin Kamu Yönetimi İçin Etkin Denetimin Gerekliliği [Search for Event in Public Administration: for Effective Public Administration The Requirement of Effective Audit]. *Türk İdare Dergisi*, (466), 9-45.
- Aliç, M. ve Rusjan, B. (2010). Contribution of the ISO 9001 internal audit to business performance. *International Journal of Quality & Reliability Management*, 27(8), 916-937.
- Aslan, T. & Kızıl, C., (2019). *İktisadi ve Sosyal Yönleriyle İnovasyon*, İKSAD Publishing House. Chapter 5 İnovasyonun Muhasebe ve Denetime Etkileri". s. 159-196.
- Bozkurt, P. (2013). Denetim Kavramı ve Denetim Anlayışındaki Gelişmeler [Audit Concept and Supervision Understanding Improvements]. *Denetim*, (12), 56-62.
- Cömert, N. ve Uzun, A. K., (2014). Bağımsız Denetçiler İçin İtibar Yönetimi [The Significance of Advertisement on The Reputation Concept and Advertisement Influence on Branding World]. *İstanbul Aydın Üniversitesi Dergisi*, 9(2), 1-23.
- Dalak, G. (2000). Denetim ve kalite denetimi. *Muğla Üniversitesi Sosyal Bilimler Enstitüsü Dergisi*, (1), 65-79.
- Delmas, M. (2001). Stakeholders and competitive advantage: The case of ISO 14001. *Production and Operations Management*, 10(3), 343-358.
- Demir, Z. (2020). Covid-19'un TFRS Kapsamındaki Finansal Raporlamaya Olan Etkileri ve TMS Kapsamında Denetçinin Dikkate Alması Gereken Hususların Değerlendirilmesi [Evaluation of The Effects of Covid-19 on Financial Reporting Under The Scope of TFRS and The Issues That Must be Considered By The Auditor in The Scope of TMS]. *Journal of Accounting and Auditing Overview*, 21(61), 255-278.
- Erdoğan, M. (2019). Denetim 4.0 ve Ötesi. [Audit 4.0 and Beyond]. *Muhasebe ve Vergi Uygulamaları Dergisi (MUVU)/Journal of Accounting & Taxation Studies (JATS)*, 12(3).
- Fernández-Muñoz, B., Montes-Peón, J. M. ve Vázquez-Ordás, C. J. (2012). Safety climate in OHSAS 18001-certified organisations: Antecedents and consequences of safety behaviour. *Accident Analysis & Prevention*, 45, 745-758.
- Jiang, R. J. ve Bansal, P. (2003). Seeing the need for ISO 14001. *Journal of Management Studies*, 40(4), 1047-1067.
- Karthikeyan, B. ve Sztandera, L. M. (2010). Analysis of tactile perceptions of textile materials using artificial intelligence techniques: Part 1: forward engineering. *International Journal of Clothing Science and Technology*, 22(2/3), 187-201.
- Kebeli, A. (2012). Kamu İç Denetiminde Kurumsal Performans Denetimi. [The Place of Performance Auditing in The Public Internal Auditing]. *Journal of Denetim*, (20), 5-34.
- KGK Independent Auditing Standard-315, (2020, 1 January). "Identifying and Assessing Material Misstatements by Knowing the Business and its Environment". Retrieved from <https://kgk.gov.tr/Portalv2Uploads/files/Duyurular/v2/BDS/bdsyeni25.12.2017/BDS%20315-Site.pdf>
- KGK Independent Auditing Standard-570, (2020, 28 December). "Business Continuity", Retrieved from <https://www.kgk.gov.tr/Portalv2Uploads/files/Duyurular/v2/BDS/BDS%20570.pdf>
- KGK (Public Oversight Authority), (2020, 28 December). Independent Auditing Evidence-720, "Responsibilities of the Independent Auditor Regarding Other Information". Retrieved from <https://kgk.gov.tr/Portalv2Uploads/files/Duyurular/v2/BDS/BDS%20720.pdf>
- KGK Independent Auditing Standard -500, (2020, 26 December). "Independent Auditing Evidence". Retrieved from https://www.kgk.gov.tr/Portalv2Uploads/files/Duyurular/v2/BDS/BDS_500_kurumsitesi.pdf
- Kiliç, Ö. Ü. B. İ. (2019). Big Data Analysis Techniques and Technologies in The Field of Accounting, Finance and Auditing International Congress of Management Economy and Policy-ICOMEPEP'2019 Autumn, Proceedings Book, p. 498 https://www.researchgate.net/profile/Oezlem-Erguet/publication/338922004_ICOMEPEP%2019_AUTUMN_PROCEEDINGS_BOOK/links/5e3300db299bf1c9b9ff2471/ICOMEPEP19-AUTUMN-PROCEEDINGS-BOOK.pdf#page=508

- Korkmaz, U. (2007). Kamuda İç Denetim [Internal Audit in The Public]. *Bütçe Dünyası Dergisi*, 2(25), 4-15.
- Kubali, D. (1999). Performans denetimi [Performance Audit]. *Amme İdaresi Dergisi*, 32(1), 31-62.
- Kula, S. ve Çakar, B. (2015). "Maslow İhtiyaçlar Hiyerarşisi Bağlamında Toplumda Bireylerin Güvenlik Algısı ve Yaşam Doyumu Arasındaki İlişki". *Bartın Üniversitesi İİBF Dergisi*, 6(12), 191-210.
- Matloff, N. (2008)." Introduction to discrete-event simulation and the simpy language. Davis, CA. Dept of Computer Science". University of California at Davis. Retrieved on August, 2(2009), 1-33.
- Mert, H. ve Tas, O. (2019). "An application of artificial intelligence on auditing". *Pressacademia*, 9(9), 65-68. doi:10.17261/Pressacademia.2019.1067
- Ministry of Treasury and Finance, (2021, 13 January). "Number of Full-Empty Internal Auditors Cadre", <https://www.hmb.gov.tr/duyuru/idareler-itibariyle-guncel-dolu-bos-ic-denetci-kadro-sayilari> adresinden erişildi.
- Kuluçlu, E. (2006). Yönetimin Denetiminden Denetimin Yönetimine [Management Control to The Management of The Audit]. *Sayıştay Dergisi*, (63), 3-35.
- KGK Independent Auditing Standard-540, (2020, 28 December). "Independent Audit of Accounting Estimates and Related Disclosures". Retrieved from https://www.kgk.gov.tr/Portalv2Uploads/files/Duyurular/v2/BDS/BDS%20540_%20kurum%20sitesi.pdf
- Öztürk, Y. (2014). Performans Denetimi Ve Türkiyede Uygulanabilirliği . Selçuk Üniversitesi Sosyal Bilimler Meslek Yüksekokulu Dergisi , 7 (1-2) , 153-170 .
- PWC. (2019). " International Financial Reporting Standards Bulletin, The Effect of Corona Virus on Accounting Practices", p. 17. Retrieved from <https://www.pwc.com.tr/tr/Hizmetlerimiz/denetim/ufrs-tfrs-raporlamasi/finansal-raporlama-standarti-tfrs-mart-2020.pdf>
- Seçkin, M., Seçkin, A. Ç. ve Coşkun, A. (2019). Production fault simulation and forecasting from time series data with machine learning in glove textile industry. *Journal of Engineered Fibers and Fabrics*, 14, 155892501988346. doi:10.1177/1558925019883462
- Shamey, R. ve Hussain, T. (2003). Artificial intelligence in the colour and textile industry. *Review of Progress in Coloration and Related Topics*, 33, 33-45.
- Şavlı, T. (2011). Independent Audit Process within the Scope of International Audit Standards. International Accounting Symposium, Number. 10, Publication of Istanbul Chamber of Certified Public Accountants (İSMMMO) Yayını, ss: 7-8. Retrieved from http://archive.ismmmo.org.tr/docs/SEMPOZYUMLAR/SEMPOZYUM_10/sunumlar/TUBASAVLI.pdf
- T.C. Presidential Legislation Information System, (2021, Ocak 22). Independent Audit Regulation, Official Gazette Number: 28509, Official Gazette Date: 26.12.2012, Retrieved from <https://www.mevzuat.gov.tr/mevzuat?MevzuatNo=16907&MevzuatTur=7&MevzuatTertip=5>
- TÜRMOB, (2020b). The Impact of the Corona Epidemic on the Turkish Economy, Evaluations and Recommendations Report. TÜRMOB Publications-503, p. 126.
- TÜRMOB (Türkiye Serbest Muhasebeci Mali Müşavirler ve Yeminli Mali müşavirler Odaları Birliği), (2020a). COVID-19: Third Party Risks and Management. TÜRMOB Publications, p. 2.

Süperpiksel Küme Bölgeleri Tabanlı Aktif Çevrit Ve GrabCut Sinerjisini Kullanarak İnsan Kulağının Otomatik Bölütlenmesi*

Automatic Segmentation of the Human Ear Using Active Contour and GrabCut Synergy Based on the Superpixel Cluster Regions

Bahadır Karasulu¹ 



* Bu çalışma 8th International Conference on Future Learning and Informatics: Data Revolution (FL 2020) konferansında sözlü olarak sunulmuştur.

¹(Doç. Dr.), Çanakkale Onsekiz Mart Üniversitesi, Mühendislik Fakültesi, Bilgisayar Mühendisliği Bölümü, Çanakkale, Türkiye

ORCID: B.K. 0000-0001-8524-874X

Corresponding author:

Bahadır KARASULU
Çanakkale Onsekiz Mart Üniversitesi,
Mühendislik Fakültesi, Bilgisayar Mühendisliği
Bölümü, Çanakkale, Türkiye
E-mail address: bahadirkarasulu@comu.edu.tr

Submitted: 20.01.2021

Revision Requested: 03.03.2021

Last Revision Received: 16.03.2021

Accepted: 01.04.2021

Published Online: 12.05.2021

Citation: Karasulu, B. (2021). Süperpiksel küme bölgeleri tabanlı aktif çevrit ve grabcut sinerjisini kullanarak insan kulağının otomatik bölütlenmesi. *Acta Infologica*, 5(1), 117-128.
<https://doi.org/10.26650/acin.865375>

ÖZ

Kulak bölgesi bireyin yaşına bağlı olarak fizyolojik bakımdan çok az değişikliğe maruz kalan değerli biyometrik bilgi içeren bir insan vücut bölgesidir. Biyometrik bilgi elde etmede kulak bölgesinin kullanımıyla ilgili çeşitli yöntemlerde kulak bölgesinin elle, yarı otomatik veya tam otomatik olarak bölütlenmesi önemli bir araştırma alanıdır. Çalışmamızda, insan kulak bölgesinin görüntüden tam otomatik olarak bölütlenerek arka plandan ayrıştırılması için süperpiksel küme bölgeleri, jeodezik bilgiye dayanan aktif çevrit tespiti ve çizge kesme yoluyla ön plan ayrıştırma işlemleri uygulayan bir yaklaşım sunulmaktadır. Çalışmamızdaki bu yaklaşım sayesinde kulak ön plan maskesi programatik ve tam otomatik biçimde kulak görüntüsünden oluşturulmaktadır. Kulak görüntüleri veri kümesi ile yapılan deneylerde uzman tarafından işaretlenen referans kulak bölgesi maskesi otomatik olarak oluşturulan ön plan maskesi ile karşılaştırılmıştır. Jaccard endeksi ölçütüne dayalı benzerlik oranları (birleşim kesişimi) dikkate alındığında yüksek başarımlı değerleri elde edilmiştir. Yaklaşımımız bu veri kümesindeki görüntüler için %84 ilâ %92 aralığında oldukça iyi başarımlı değerlerine sahiptir. Çalışmamızda, önerilen sinerjik yaklaşımın başarımlı hem niteliksel hem de niceliksel olarak deneysel sonuçlarla ortaya konulmaktadır.

Anahtar kelimeler: Süperpiksel, Aktif çevrit, Çizge kesme, Görüntü bölütleme

ABSTRACT

The ear region is a region of the human body region containing valuable biometric information that is subjected to a few physiological changes depending on the individual's age. Manual, semi-automatic, or fully automatic segmentation of the ear region in various methods related to the use of the ear region in obtaining biometric information is an important area of research. In our study, we present an approach that applies superpixel cluster regions, active contour detection based on geodesic information, and foreground separation by graph cutting, to segregate the human ear region from the image by fully automatic segmentation from the background. Thanks to this approach in our study, the ear foreground mask is created programmatically and fully automatically from the ear image. In the experiments with the ear images data set, the reference ear mask marked by the expert was compared with the automatically created foreground mask. It has been obtained high performance values were obtained, considering the similarity rates (i.e., intersection over union) based on the Jaccard index metric. Our approach has quite good performance values (in the range of 84% to 92%) for the images in this dataset. In our study, the success of the proposed synergistic approach is demonstrated both qualitatively and quantitatively with experimental results.

Keywords: Superpixel, Active contour, Graph cut, Image segmentation

1. GİRİŞ

E-Sağlık ve turizm alanlarında bireylerin doğru olarak tanınabilmesi, kimlik tanımlama (identification), yetkilendirme (authentication) ve yetki geçirme (validation) konularında birçok çalışma mevcuttur. Bu alandaki çalışmalar özellikle bireylerin günlük yaşantılarını aksatmamayı amaçlamaktadır. Temassız olarak bireyin kim olduğunun tespit edilmesi ve bireyin belirli süre zarfında temassız olarak takip edilmesine dayanan sistemlere olan ihtiyaç gün geçtikçe artmaktadır. Biyometrik; genel olarak ölçülebilen, fiziksel karakteristiğe sahip pratik uygulamaya yönelik, ayrıştırıcı, evrensel, kalıcılığı olan ölçüm ve ölçütleri içermektedir. Bu bakış açısıyla tanımlama ve yetkilendirme sistemlerinde; ses kayıtları, parmak izleri, imzalar, el avuç içi geometrisi, göz iris, yüz görüntüleri, klavye tuş bilgileri gibi bilgiler biyometrik alanında üzerinde çalışılan uygun veriler olarak kullanılmaktadır (Tomczyk ve Szczepaniak, 2019). Biyometrik yetkilendirme sistemleri, bireylerin fizyolojik ve davranışsal ölçümlerini alarak bunları bireyin tanımlayıcı (kimlik) bilgisini onaylamada kullanılmaktadır. Bu bilgilere dayanan karar verme, çıkarsama ve akıl yürütme işlemleri ile etiketleme (tagging), sınıflandırma (classification) ve tanıma (recognition) sonuçlarına yapılan deneylerle ulaşılmaktadır (Schuckers, 2010). Bu bilimsel bir yoldan günlük hayattaki güvenliğin (sağlık, turizm, askeri, finansal, eğitim ve benzeri) ölçülebilir, tekrar edilebilir ve sürdürülebilir bir hale gelmesi anlamına gelmektedir.

Bilgi modeli oluşturulurken kaynak, bu kaynaktan elde edilen özellikler ve bunların değerleri göz önüne alınmaktadır. Özellikle yüz görüntüleri son yıllarda derin öğrenme ve makine öğrenmesi tekniklerindeki gelişmelerle birlikte oldukça önem kazanmıştır. Fakat yüz bölgesinde yaşa bağlı oluşan değişimler, sakal, gözlük, şapka (gölge düşüren başlıklar) ve benzeri tespit ve tanımayı güçleştirici, örtücü olan veya diğer bireylerden ayrıştırıcılığı bozan öğeler bulunduğu durumlarda biyometrik bilginin elde edilme kalitesi düşmekte veya imkânsız hale gelebilmektedir. Bu açıdan yaş ile değişimle de olsa fizyolojik yapısındaki oranların görece sabit kaldığı bilinen ve başka nesnelere daha az örtüşmeye maruz kalan fizyolojik bir vücut bölgesinin işlenmesi biyometrik açıdan önem kazanmıştır.

İnsan kulağı içeren görüntülerin elde edilmesi diğer fizyolojik bölgelere göre nispeten daha kolaydır. İnsan kulağının anatomik bazı özellikleri yaşa bağlı olarak değişim gösterse de ayrıştırıcı olan iç ve dış kıvrımlarının geometrik oranları büyük çapta sabit kalmakta, yüz ifadesindeki değişimden etkilenmemektedir. Ayrıca, kulak geometrisinden kulağı ifade eden kenarların (edge) elde edilmesine dayalı yöntemler ile güçlü bir tanımlamanın yapılabilmesi mümkün olurken, kulağı kısmen örten saç, küpe, takı ve kulaklıkların görüntüden temizlenmesi de söz konusu olmaktadır. Özellikle, bu kulak kıvrımlarının belirleyici köşe noktalarına bazı görsel işaret noktalarının (landmark) yerleştirilmesi yoluyla (Cintas, Delriux, Navarro, Quinto-Sánchez, Pazas ve Gonzalez-Josè, 2019), iki boyutlu (2D) veya üç boyutlu (3D) kulak görüntülerinde değişik ışıklandırma, kamera bakış açısı ve bireyin kafasının konumu gibi çeşitli durumlar bilgi modelinin oluşturulmasında olumlu ya da olumsuz etki edebilmektedir. Bu işaret noktaları elle (manuel) veya otomatik olarak kulak üzerine yerleştirilmekte, bu sayede kulak bölütleme (ear segmentation) ve kulak tespiti (ear detection) gerçekleştirilebilmektedir. Ayrıca, kulak üzerine kulağı kapsayan çevrit (contour) yerleştirilmesi yoluyla kenar tabanlı bölütleme (edge-based segmentation) yapılabilmektedir. Özellikle literatürdeki çalışmalarda insan cilt rengi tespiti yoluyla ve kenar eğim şiddetinin (magnitude) artırılması yoluyla (Chen ve Bhanu, 2005) çevrit eşleme yapılabilmekte, bu sayede kulak tanıma (ear recognition) sonucu elde edilebilmektedir. Joshi ve Chauhan (2011) çalışmasında kulak tespiti için iki farklı yaklaşım önerilmiştir. Buna göre; ilk yaklaşım kenar tespitine dayanmakta, ikinci yaklaşımda ise şablon eşleme (template matching) ile kulak tespiti yapılmaktadır. Yapılan tespitlerin doğruluğu destek vektör makineleri (support vector machine) aracılığıyla kontrol edilmiştir. Şablon eşlemede ise öklid mesafesi kullanılmıştır. Literatüre daha geniş açıdan bakan Emeršič ve ark. (2017) çalışmasında ve Pflug ve Busch (2012) çalışmasında da belirtildiği üzere; kulağın biçimine dayanan geometrik teknikler, kulağın genelini ele alarak kulak temsiline (representation) dayanan bütüncül (holistic) teknikler, yerel özniteliklerin elde edilmesiyle kulak ilgi noktalarının (point of interest) belirlenmesine dayanan yerel teknikler ve melez yaklaşımlar mevcuttur. Özellikle bütüncül tekniklerle kulak biçimini ortaya koymada kenar tespiti, çeşitli filtreler ve kuvvet alanı dönüşümlerinin de kullandığı görülmektedir. Bu sayede kulak tespiti ve kulak tanıma yüksek doğrulukta yapılabilmektedir. Khatib ve ark. (2014) çalışmasında, renkli görüntülerin bölütlenmesinde farklı renk uzayı modelleri için çizge kesme (graph cut) tekniği ile otomatik yoldan görüntüden nesnelere bölütlenmesi üzerine çalışılmıştır. Bu sayede çeşitli doğal görüntüler üzerinden kıyaslama yapılarak çizge

kesmenin otomatik görüntü bölütleme için başarımının oldukça yüksek olduğu kanıtlanmıştır. Bu yaklaşımın renkli kulak görüntülerinde kullanılabilmesi fikrine dayanarak çalışmamızda görüntüden nesne (kulak bölgesi) bölütlemeye dair bir sinerjik yaklaşım oluşturularak literatüre katkı olarak sunulmaktadır. Buna göre renk uzayı modelleri aracılığıyla görüntüdeki çeşitli küme bölgelerinin geometrik yapısı belirlenerek bütüncül bir teknikle aktif çevritin uygulanabildiği ve çizge kesmeye dayanarak oluşturulan bu sinerji sayesinde otomatik olarak kulak bölgesinin bölütlenmesi gerçekleştirilebilmektedir.

Çalışmamızda makine öğrenmesi kapsamında basit doğrusal iteratif kümeleme (simple linear iterative clustering, SLIC) yöntemi yoluyla süperpikseller ile oluşturulan kümelerin ifade ettiği bölgeler kullanılmaktadır. Bu yolla insan kulağının verilen görüntü içerisinde bulunduğu konumunun en uygun dış çevritini tespit edecek morfolojik jeodezik aktif çevrit (geodesic active contour, GAC) yöntemi uygulanabilmektedir. Bu sayede bu çevrit ile oluşturulan maske görüntüsüne bağlı olarak çizge tabanlı (GrabCut) biçimdeki iki seviyeli bölütleme (bi-level segmentation) yapılmaktadır. Bu bölütleme otomatik olarak jeodezik terim kullanan aktif çevrit ile bulunan kulak dış ve iç bölgesine dair bilgiyle oluşturulan maskenin GrabCut yönteminde kullanılarak çizge kesme yaklaşımıyla arka plan (background) ve ön plan (foreground) olarak görüntüden kulak nesnesinin otomatik bölütlenmesi şeklinde gerçekleşmektedir (Achanta, Shaji, Smith, Lucchi, Fua ve Süsstrunk, 2012; Banerjee ve Mery, 2016; Márquez-Neila, Baumela ve Alvarez, 2014). Çalışmadaki deneyler, Esther Gonzalez-Sánchez tarafından oluşturulan AMIDB (Mathematical Analysis of Images, AMI) isimli iki boyutlu (2D) insan kulak görüntü veri kümesi üzerinde gerçekleştirilmiştir (Gonzalez-Sánchez, 2008).

Literatürdeki çeşitli kulak görüntü veritabanları ve veri kümelerine bakıldığında; Tomczyk ve Szczepaniak (2019) çalışmasında belirtildiği gibi farklı veritabanları veya veri kümeleri için birbirinden farklı nesnel ölçümler ve ölçütler (metric) kullanılabilir. Bazen başarımlar değerlendirilmede tek pozdaki tek açıdan çekilmiş görüntü kullanılırken, sadece sol kulak, sadece sağ kulak veya her ikisi birden kullanılabilir. Bu durum çalışmaların birbirleriyle kıyaslanabilirliğini olumsuz yönde etkileyebilmektedir. Çalışmamızda AMIDB veritabanındaki yedi farklı poz açısından (etiket) ve her iki kulak birden değerlendirmeye alınmış, böylece veritabanı sayesinde daha tutarlı bir deney ortamı oluşturulmuştur.

Kulak tespitini görüntü içerisindeki nesne veya nesnelerin kenar tespitine dayanan kulak konumlandırma (ear localization) doğruluğu veya tespit doğruluğu ile değerlendiren birçok çalışmada cilt renginin ve bölgesinin de dikkate alındığı görülmektedir. Bu tarz kulak tespiti çalışmalarında AWE veri kümesi (Emeršič, Štruc ve Peer, 2017), UND (University of Notre Dame, 2015) ve USTB (University of Science and Technology Beijing, 2002) veritabanları gibi çeşitli veri kümeleri kullanılmıştır (Emeršič, Gabriel, Štruc ve Peer, 2015). Bunlar oluşturulurken; koşullu bir biçimde (sabit noktada duran bireylerin) veya kısıtlı bir ortamda (kapalı iç mekân) görüntüler elde edilebildiği gibi Internet ortamından ve dış mekân dahil değişik ortamlardan (*in-the-wild*) görüntüler de elde edilebilmektedir. Koşullu ve genellikle kapalı iç mekânda kısıtlar altında oluşturulan veritabanlarındaki biyometrik değerlendirme sonuçları kıyaslama (benchmarking) yapmaya daha uygun bir ölçüm sınaması sağlamaktadır (Wang, Yang ve Zhu, 2021). Çalışmalarda görüntü işleme, makine öğrenmesi veya derin öğrenme yöntemleri ile yapılan deneylerdeki başarımlar ölçümlerinde bilişim teorisinden (information theory) bilgi elde etme (information retrieval) alanındaki birçok nesnel ölçüt kullanılarak başarımlar değerlendirme sonuçları elde edilmektedir (El-Naggar, Abaza ve Bourlai, 2018). Genellikle, cilt rengi tespit edilerek kulak bölgesi içeren görüntüden cilt bölgesinin ayrıştırılmasında çeşitli renk uzayları kullanılmaktadır. Bu sayede hem gürültü olarak görüntü bölütlemeyi etkileyecek ışıklandırma koşuluna bağlı piksel parlaklık yoğunluk farklılıklarının önüne geçilmeye hem de kulak dış çevritinin daha baskın olarak görüntüden elde edilmesine dair bir ön işlem (preprocessing) gerçekleştirilmektedir (Jacob ve Raju, 2011).

Bu makale aşağıdaki şekilde organize edilmiştir. İkinci bölümünde uygulanan yöntemlerin detaylarına yer verilmektedir. Böylece verilen görüntüye ön işlem yapılması ve iteratif kümeleme yöntemiyle süperpiksel tabanlı bölgelerin oluşturulması anlatılmaktadır. Bu bölgeler temel alınarak oluşturulan kulak bölgesine dair aday çevritler içerisinde en uygun olarak en dıştaki çevritin morfolojik jeodezik aktif çevrit yöntemiyle belirlenerek çizge kesmeye dayanan iki seviyeli bölütlemeye kullanılmak üzere maske görüntüsünün oluşturulması incelenmektedir. Bu yolla çalışmadaki tam otomatik bölütlemeye dair detaylara yer verilmektedir. Üçüncü bölümde, deneysel sonuçlar ve tartışmaya yer verilerek AMIDB veri kümesi üzerinden yapılan deneylerden elde edilen görsel nitel ve nesnel nicel sonuçlar sunulurken, bunun yanı sıra yaklaşımımızdaki yöntemlere dair parametrik değerlere de yer verilmektedir. Dördüncü bölüm olan sonuç bölümünde ise yapılan çalışmaya dair elde edilen bilimsel bulgulara ve sonuçlara dayanan tartışmaya yer verilerek, alanla ilgili çeşitli önerilere değinilmektedir.

2. YÖNTEM

Kulak bölgesinin bölütlenmesi konusu, öncelikle renkli görüntülerden özellikle günlük hayattan koşullu veya koşulsuz olarak alınarak oluşturulan bir görüntü veri kümesinin toparlanması, bu veri kümesinin üzerinde çeşitli yöntemlerin denenmesi ve bunlar arasında başarımları kıyaslaması yapılabilmesine olanak tanınmaktadır. Farklı renk uzaylarında görüntü bölütleme yapılmasında sıklıkla kümeleme (clustering), eşikleme (thresholding), bölge tabanlı (region-based) veya kenar tabanlı (edge-based) bölütleme yaklaşımları kullanılmaktadır. İlgiilenen nesne (Object of Interest, OOI) içeren bölgenin arka plandan ayrılarak ön planda etiketlenmesi en sık uygulanan yaklaşımdır. Bu sayede iki seviyeli bölütleme (bi-level segmentation) işlemi heterojen bölgelerin kenarlarla ayrıştırıldığı, homojen veya benzer özellikteki bölgelerin ise belirli nesnelere olarak işaretlendiği bir süreç olarak karşımıza çıkmaktadır (Khattab ve ark., 2014).

Basit doğrusal iteratif kümeleme (SLIC) yöntemi k -ortalamlar (k -means) kümeleme yönteminin süperpiksel bölgeleri elde etmeye uyarlanmış biçimi olarak görülebilir. Süperpikseli oluşturan algoritmalar ya çizge (graph) tabanlı ya da eğimli çıkış (gradient-ascent) tabanlı olabilmektedir. Bu algoritmalar görüntü içerisindeki pikselleri anlamlı atomik bölgelere gruplandırmakta, böylece bölütleme işlemi daha hızlı ve kaliteli hale gelebilmektedir. Çizge tabanlı algoritmalar her bir piksele bir çizgenin bir düğümü olarak davranırken, eğimli çıkış algoritmaları piksellerin ilk kümesi ile başlayıp süperpiksellerin oluşturulduğu nihai biçime karşılık gelen yakınsama kriterine kadar iteratif olarak kümeleri oluşturmaktadır. Süperpiksellerin oluşturulmasında görüntüdeki nesne konumlarının kaybolmamasına dikkat edilmesi önemli bir konudur.

SLIC yöntemi, eşit büyüklükte süperpiksel üreten, k adet kümenin ilk merkez noktalarının belirlenmesinin ardından görüntüdeki pikselleri iteratif olarak arama alanı içerisinde kaldıkları en yakın küme merkezine ilişkilendiren bir algoritmaya sahiptir (Achanta ve ark., 2012). SLIC algoritmasında, ilgili piksel ile ilgili küme merkezi arası mesafe $labxy$ renk görüntü düzlemi uzayında beş boyutlu öklid uzaklığı olarak her adımda hesaplanır. Böylece piksellerin ilgili kümeye dahil olduğu bölge oluşturulur. Güncelleme adımına kadar hata miktarı hesaplanarak piksellerin en yakın küme merkezine olan ilişkisi kurulur. Böylece güncelleme adımında küme merkezinin yeni konumu hata miktarı değeri göz önüne alınarak belirlenir. Bu işlemlerin çizge tabanlı algoritma ve bazı diğer algoritmalara göre hesapsal karmaşıklığı oldukça düşüktür ve $O(N)$ olarak işlenmektedir.

Çevrit gelişim (evolution) yöntemleri özellikle iki seviyeli bölütlemelerde kullanılmaktadır. Jeodezik aktif çevrit (geodesic active contour, GAC) ve bunun bir çeşiti olan morfolojik GAC, görüntüde ilgiilenen nesne (ön plan) ve arka plan olarak iki bölgeyi ayrıştıran sınır çizgisi etrafına denk gelen bir çevriti bulmayı amaçlamaktadır. Bu yöntemin parlaklık (intensity) ve eğim (gradient) gibi öznitelikler kullanılarak bir gömülme (embedding) fonksiyonunun kısmi diferansiyel denklemlerinin çözümüne dayanan bir çalışma mekanizması vardır. Bu gömülme fonksiyonu sıfır seviyesindeki kümeyi çevrit olarak alır ve çeşitli nesne sınırlarını hesaplayarak her bir adımda çevriti günceller. Bu güncelleme esnasında görüntüdeki parlaklık ve eğim farklı bölgeler için hesaplanmakta fakat sınır bölgesinde genellikle hesaplama durdurulamamaktadır. Burada jeodezik terim sayesinde eğri gelişimi hesaplaması durdurularak uygun sınır bölgeleri oluşturulur (Banerjee ve Mery, 2016; Márquez-Neila ve ark., 2014).

Açma (open) ve kapama (close) gibi çeşitli matematiksel morfolojik işlemler GAC yöntemiyle enerji tabanlı bir biçimde uygulanmakta, minimum enerjili yüzey veya eğri hesabı yoluyla hesaplanan bir jeodezik terim elde edilmektedir. Bilgi içermeyen alanlara çevrit eğrisinin takılıp kalmaması için bir balon kuvveti terimi de morfolojik GAC yönteminde mevcuttur (Cohen, 1991). Márquez-Neila ve ark. (2014) çalışmasında yılan (snake) olarak ifade edilen bir eğrinin I görüntü içeriğini ifade edecek şekilde bir enerji fonksiyonu biçiminde tanımlanan çevrit ile oluşturulabildiği ifade edilmektedir. Buna dair Eşitlik (1) aşağıda verilmektedir. Burada z terimi yılan eğrisini, C terimi ise enerji tabanlı hesaplanan çevriti ifade etmektedir.

$$\begin{aligned} E(C) &= \int_0^{\text{uzunluk}(C)} g(I)(C(z)) dz \\ &= \int_0^1 g(I)(C(r)) |C_r| dr \quad (1) \end{aligned}$$

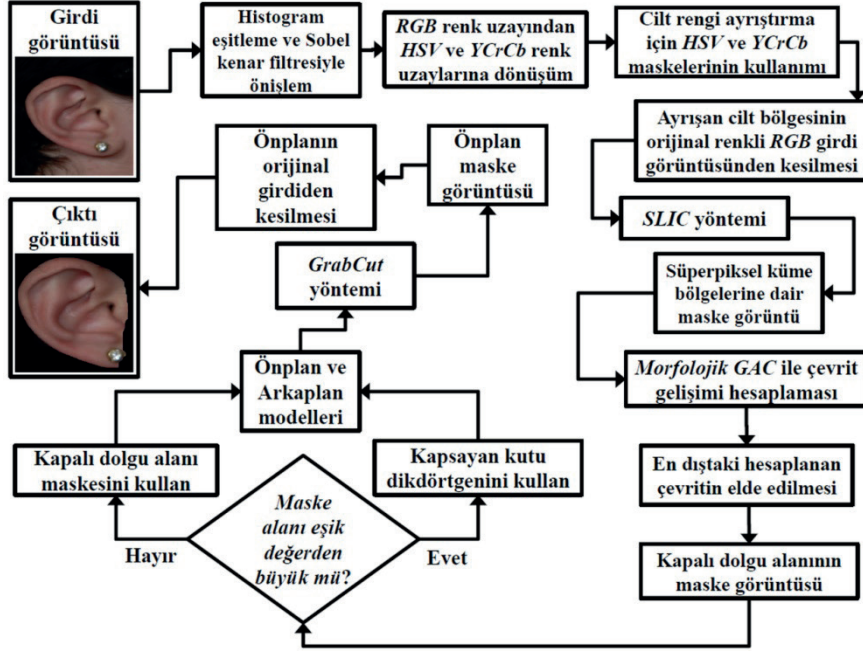
Burada, Eşitlik (1) 'deki öklid uzaklığı biçimindeki eğri uzunluğu parametresi olarak $dz = |C_r| dr$ parametresi kullanılmıştır. Böylece görüntüdeki kenarların belirlenmesinde Gauss filtresine dayanarak oluşturulan ve Eşitlik (1) eşitliğinde de verilen

terim $g(I): \mathbb{R}^d \rightarrow \mathbb{R}^+, x \rightarrow g(I)(x)$ olarak ifade edilmekte ve görüntü içerisinde ilgilenilen bölgenin seçilmesini sağlamaktadır. Aynı çalışmada morfolojik GAC kısmi türevli denklemlerinin morfolojik balon, çekim kuvveti ve morfolojik yumuşatma (smoothing) bileşenleri bu Eşitlik (1) eşitliğine eklenip çözümlenmiştir. Buna göre çalışmamızdaki morfolojik GAC yukarıda bahsedilen diğer temel parametrelerin yanı sıra dört temel parametreye bağlı çalışmaktadır. Durdurma kriterinin (stopping criterion) parametreleri olan alfa ve sigma parametreleri balon kuvveti terimi parametresi (genişleme kuvveti), bunların yanı sıra balonun ve yumuşatmanın eşikleme (thresholding) parametresi çalışmamızda görüntü içerisindeki kulak nesnesi için dış çevritin uygun biçimde gelişim hesaplamasında kullanılmaktadır. Bu parametrelerin teknik detayına Márquez-Neila ve ark. (2014) çalışmasında yer verilmiştir.

GrabCut yöntemi Gauss karışım modelleri (Gaussian mixture models, GMM) kullanarak renk uzayındaki parlaklıklara istinaden ön plan ve arka plan bölgelerini öğrenmektedir. Çizge kesme yaklaşımında görüntüdeki kenar ve bölgesel bilgilerle dayanarak bir enerji fonksiyonu oluşturulur ve iteratif olarak eniyilemesi (optimization) yapılır. Çizge kesme süreci öncelikle bir kullanıcının ilklendirme amacıyla yarı interaktif (etkileşimli) bir yoldan ilgilenilen nesnenin etrafına ilgilenilen nesneyi kapsayan bir dikdörtgen (kapsayan kutu) çizmesi ile başlar, ayrıca nesneye göre dışta kalan ve içte kalan bölgeleri belirler. Buna göre görüntü her bir pikselin etiketlenerek çizgede bir düğüme (*node*) atandığı ağırlıklandırılmış bir çizge olarak modellenmiş olur. Kaynak (*source*) ve bitiş (*sink*) terminalleri olarak iki düğüm oluşturularak çizge tamamlanır.

Literatürde *min-cut* ismi verilen yöntem sayesinde bu yarı interaktif yaklaşıma dayanan çizge, kaynak ile bitiş terminallerinin ayırdığı iki farklı düğüm kümesine bölünür. Bu işlem kenar (edge) olarak ifade edilecek olan kısmın komşu pikseller arasındaki hesaba dayanan ağırlıkların atanması sayesinde çizge kesme olarak yapılır (Banerjee ve Mery, 2016). Bu yolla bu dikdörtgen içerisinde kalan her bir etiketlenmemiş piksel ön plana veya arka plana ait olarak etiketlenerek yakınsama sağlanana kadar iteratif olarak süreç devam eder. Tushar (2018) çalışmasında Hue-Saturation-Value (*HSV*) renk uzayında GrabCut yöntemini kullanarak otomatik bölütleme yoluyla cilt lezyonu ön plan görüntüsü elde etmeyi başarmıştır. Önişlemeden geçirilen görüntülerde ayrıştırılmak istenen bölge ile uzman tarafından işaretlenmiş referans bölge (ground-truth) benzerlik oranı *Jaccard* endeksi değeri olarak ölçülmüştür.

Çalışmamızda GrabCut yönteminde otomatik çizge kesme yapılması için gerekli olan arka plan ve ön plan modellerinin oluşturulmasında nesneyi çevreleyen kısıtlayıcı kapsayan kutu (*bounding box*) şeklindeki ilklendirme (initialize) dikdörtgeninin belirlenmesi yoluyla çizge kesme işlemi veya belirli bir eşik değerine göre kapalı dolgu alanı biçimindeki maske görüntünün (0 ve 1 değerli piksellerden oluşan siyah beyaz görüntü) elde edilmesinde morfolojik GAC yöntemiyle belirlenen insan kulağına dair en dıştaki çevritin kapsadığı bölgeden yararlanılmıştır. Buradaki kapsayan kutu, görüntü işleme alanında genellikle kullanılan *dışbükey zarf* (convex hull) yöntemi temel alınarak oluşturulmuştur. Buna göre insan kulağına dair en dıştaki çevritin azami uç değerli köşe noktalarına dokunan kenarlara sahip bir dikdörtgen belirlenmiştir. Bu çevriti kapsayacak biçimde kapsayan kutu otomatik olarak görüntüye yerleştirilmektedir. GrabCut yönteminin ilklendirme koşulu için belirli bir eşik değeri dikkate alınarak, bu maskeyle görüntüdeki kapalı dolgu alanının mı yoksa kısıtlayıcı kapsayan dikdörtgenin mi seçilerek işleme başlanacağı belirlenmektedir. Morfolojik GAC yöntemi ise verilen görüntünün SLIC yöntemiyle elde edilen süperpiksel küme bölgeleri olarak elde edilmiş kaba hatlı bölgeler üzerinden çalıştırılarak kulak nesnesini kapsayan uygun en dış çevriti ve buna dair kapalı dolgu alanını oluşturmaktadır. Maske ile ifade edilen nesneye dair bölgenin (kapalı dolgu alanı) dışarısında kalan alanların belirlenmesi ve bu yolla GrabCut yöntemiyle iki seviyeli bölütlemenin tam otomatik olarak yapılması da sağlanmıştır. Bu iki alternatife (kapalı dolgu alanı veya kapsayan kutu kullanımı) uygun ön plan ve arka plan modelleri otomatik olarak oluşturulup kapsama durumlarına göre GrabCut yöntemine gönderilmektedir. GrabCut yönteminin gerekli iterasyon kadar çalıştırılmasıyla eniyilenmiş ön plan maskesi oluşturulmaktadır. Aşağıdaki Şekil 1'de çalışmamızda önerilen yaklaşıma dair akış şeması verilmektedir.



Şekil 1. Önerilen yaklaşımın akış şeması.

Şekil 1'de görülebileceği gibi renkli *RGB* görüntü girdi olarak alınarak gerekli ön işlemlerden geçirildikten sonra sırasıyla SLIC yöntemine, oradan morfolojik GAC yöntemine geçirilmektedir. En son olarak GrabCut yöntemine bu iki yöntemin uygulanmasıyla elde edilen ilgili maske görüntüsü geçirilerek nihai aşamada ön plan olarak işaretlemesi yapılmaktadır. Bu sayede kulak bölgesinin arka plandan kesilmesiyle elde edilen nihai maske görüntüsü kullanılarak renkli görüntüden kesilen kulak bölgesini içeren çıktı görüntüsü elde edilmektedir. Yapılan deneylere ve yaklaşımımızda kullanılan yukarıda anılan yöntemlere dair çeşitli parametrik detaylara sonraki bölümde yer verilmektedir.

3. DENEYSEL SONUÇLAR VE TARTIŞMA

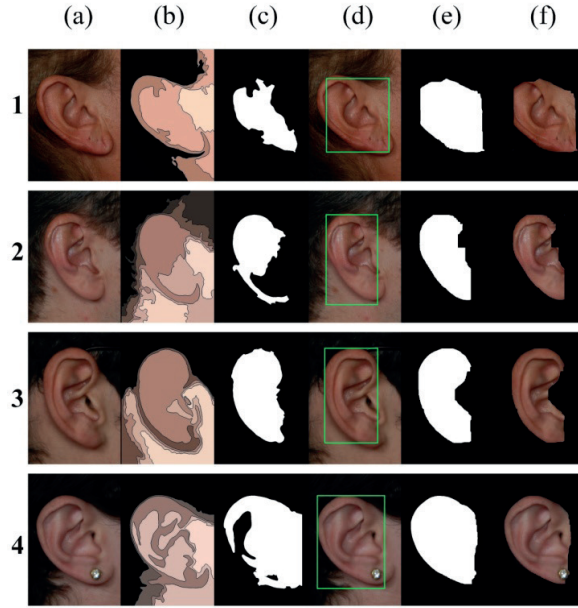
Çalışmamızdaki deneylerde Esther Gonzalez-Sánchez tarafından oluşturulan iki boyutlu (2D) insan kulağı görüntüleri veri kümesi AMIDB kullanılmıştır. İspanya'daki bir üniversitede yaşları 19 ilâ 65 arası değişen 100 farklı bireyden koşullu bir biçimde (sabit noktada duran bireylerin) altı farklı açıda sağ kulak ve sol kulak görüntüsü sabit ışık ve kamera uzaklığında kaydedilmiştir. Bu veri kümesinde görüntüler her birey için yedi farklı etiket ile sunulmuştur. Bunlar arka (*back*), ön (*front*), yukarı (*up*), aşağı (*down*), sol (*left*), sağ (*right*) ve yakın (*zoom*) olarak verilmiştir. AMIDB veri kümesinde her biri 492x702 piksel çözünürlükte olan *RGB* renk uzayında tanımlı 700 adet JPEG formatında görüntü yer almaktadır. Bu görüntüler ilgili deneye katılan ilgili bireyin tanımlayıcı numarasıyla ve çekim şartı (arka, ön, yukarı, aşağı, sol, sağ ve yakın) olarak etiketlenmiştir. Veri kümesi dört adet alt kümeye (subset) bölünmüş olarak ilgili Internet adresinde sunulmuştur (Gonzalez-Sánchez, 2008).

Çalışmamızda saç ve diğer kulak harici nesnelere tespit ve bölütlemeye etki etmemesi adına ön işleme aşamasında histogram eşitleme yoluyla ışıklandırma ve renk ton farklılıklarından kaynaklı bölütleme sorunlarını en aza indirecek bir yol izlenmiştir. Ayrıca, literatürde sıklıkla kullanılan Sobel kenar filtresi de görüntüye uygulanarak kulak çevrit tespitine yardımcı olacak cilt ilâ cilt olmayan bölgeler arasında eğimin oluşması sağlanmıştır. Ardından diğer renk dağılımlarından cilt rengini ayrıştırmak için *HSV* ve *YCrCb* renk uzaylarında uygun maskeler elde edilerek orijinal *RGB* renkli resimdeki cilt bölgesi elde edilmiştir. *YCrCb* renk uzayında parlaklık (*Y*), kırmızı renk (*Cr*) ve mavi renk (*Cb*) bileşenleri mevcuttur. Burada *HSV* renk uzayı kullanılması renk dağılımlarının tepe noktası (mode) tespitinde yararlı olmaktadır. Böylece bu renk uzayı modeli renkli bölgelerin görüntü içerisinde daha belirgin olarak ayrıştırılabilmesine yardımcı olmaktadır. *YCrCb* renk uzayındaki yapılan işlem sayesinde parlaklık (*Y*) bileşenin ayrıştırılmasıyla piksel bazlı parlaklığa bağlı yoğunlukların renk bileşenleri olan kırmızı renk (*Cr*) ve mavi renk (*Cb*) bileşenlerinin üzerindeki etkisi azaltılabilmektedir (Jacob ve Raju, 2011). Bu işlem cilt bölgesinin diğer bölgelerden ayrıştırılmasında oldukça iyi sonuçlar alınmasına da yol açmaktadır. Bu işlemin ardından

bu cilt bölgesinden oluşan ve orijinal renkli *RGB* görüntüden kesilerek elde edilen renkli görüntü doğrudan SLIC yöntemine geçirilmiştir. SLIC yöntemiyle toplam en fazla yedi küme merkezi olacak şekilde süperpiksel küme bölgeleri cilt bölgesini içeren bu görüntüden elde edilmiştir. SLIC çıktısı morfolojik GAC yöntemine aktarılarak çevrit gelişimi jeodezik terime uygun olarak ilerletilmiş elde edilen kulak nesnesi kapsayan en uygun dış çevritin içerisinde kalan kapalı eğri alanında dolgu yapılmıştır. Bu sayede maske görüntü olarak (siyah beyaz) bir görüntü elde edilmiştir. Buna göre görüntüde arka plan siyah renkli piksellerle ön plan ise beyaz renkli pikseller ile ifade edilmektedir.

Deneylerde SLIC yöntemi için literatürdeki Achanta ve ark. (2012) temel çalışmasında kullanılan parametrelerin belirtilen aralıkları göz önüne alınmıştır. Ayrıca, Walt ve ark. (2014) çalışmasında belirtildiği haliyle *Scikit-Image* görüntü işleme kütüphanesinin Python programlama dili paketinin görüntü bölütleme modülündeki SLIC yöntemiyle ilgili parametreleri dikkate alınmıştır. Çalışmamızda oluşturulan sinerjik yaklaşıma bu parametrelerin uygunluğu göz önünde bulundurularak deney parametrelerinin el ile belirlenen ilklendirilmiş (initialize) değerleri olacak şekilde; Gauss yumuşatması çekirdek büyüklüğü olarak *sigma* parametresi 4, renk yaklaşıklık ve uzamsal yaklaşıklık için *yoğunluk* (compactness) parametresi 4 ve azami küme sayısı 7 olarak belirlenmiştir. Buna göre SLIC ve sıradüzensel çizge birleştirme yapılması sayesinde süperpiksel bölgeleri oluşmaktadır. Bu küme bölgelerinden oluşan görüntü gri tonlamalı hale getirilerek morfolojik GAC yöntemine verilmektedir. Bu morfolojik GAC yöntemi için alfa parametre değeri 2300, sigma parametre değeri 5,58 alınmış, balon kuvveti terimi 1 alınarak, eşikleme değeri ise 0,41 olarak belirlenmiştir. Buradan elde edilen dış çevrite dair (kulağı kapsayan en dış çevrit) iç bölge kapalı dolgu alanı haline getirilmekte böylece siyah beyaz renkte maske görüntüsü oluşturulmaktadır. Morfolojik GAC ile oluşturulan maskenin kullanılmasıyla eğer belirli bir eşik değerinin altında kalıyorsa ön plan ve arka plan modelleri otomatik bir biçimde oluşturularak GrabCut yöntemi çalıştırılmaktadır. Bu eşik değeri dolgunun kapladığı alan büyüklüğünün görüntünün %50'lik kısmına denk gelip gelmediğini kontrol eder. Eğer belirlenen eşik değerinden bu dolgu alanı büyükse doğrudan ilklendirme dikdörtgeni kullanılarak otomatik bir biçimde ön plan ve arka plan modelleri oluşturulup GrabCut yöntemi çalıştırılmaktadır. Bu işlemde kulak nesnesinin içerisinde bulunduğu içte kalan bölge (ön plan modeli) için morfolojik GAC yönteminin yukarıda verilen parametrik değerlerle 355 iterasyon sonucunda ürettiği bölge kullanılmaktadır. GrabCut yöntemi iterasyon sayısı ise 5 olarak belirlenmiştir.

Orijinal girdi görüntüsü ve bu maske görüntüsü kullanarak GrabCut çizge kesme yöntemi çalıştırılmaktadır. GrabCut tarafından yapılan kesme işlemi sonucunda oluşturulan yeni maske görüntüsü sayesinde ön plan olarak işaretlenen kulak bölgesi arka plandan kesilerek renkli çıktı görüntüsü olarak elde edilmektedir. Aşağıdaki Şekil 2'de AMIDB veri kümesinden rasgele seçilen dört farklı görüntü için sırasıyla Şekil 2 (a) orijinal renkli görüntü, Şekil 2 (b) SLIC süperpiksel küme bölgeleri, Şekil 2 (c) morfolojik GAC en dış çevritiyle oluşturulan kapalı dolgu alanına uygun maske görüntü, Şekil 2 (d) GrabCut için nesneyi kapsayan kutu olarak ilk dikdörtgen bölge, Şekil 2 (e) GrabCut ile çizge kesme sonucu elde edilen yeni maske görüntüsü, Şekil 2 (f) renkli çıktı görüntüsü olarak verilmektedir. Şekil 2'deki satır numaraları ile belirtilen sırasıyla 1) "002_up_ear", 2) "020_front_ear", 3) "056_zoom_ear" ve 4) "090_zoom_ear" isimli görüntüler ve bunların yukarıda anılan yöntemlerle işlenmiş halleri verilmektedir.

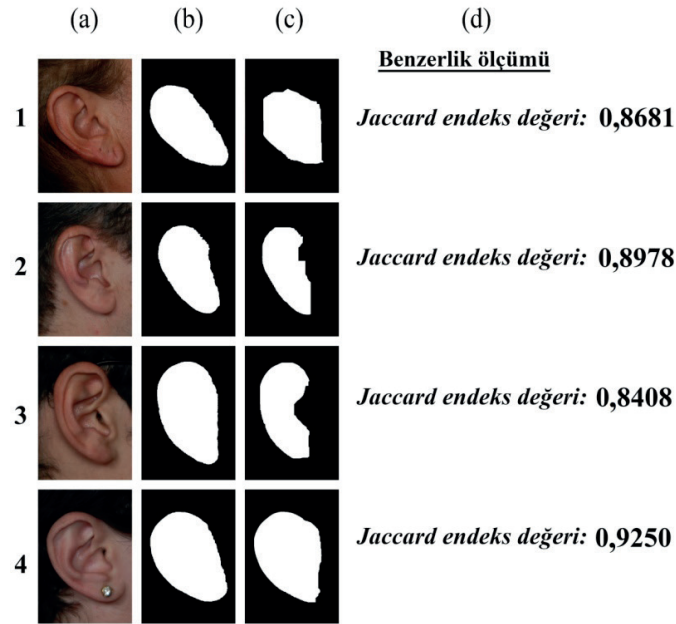


Şekil 2. AMIDB deneysel sonuçlar.

Deneylerde uzman tarafından işaretlenmiş referans (ground-truth) bölge maske görüntüsü ile programatik olarak tespit ve bölütlenmesi yapılmış bölgenin karşılaştırılmasında; birleşim kesişimi (Intersection over Union, *IoU*) olarak tarif edebileceğimiz *Jaccard* endeks değerleri kullanılmıştır. Bir sınıflandırmada doğru olarak sınıflandırılan veya yanlış olarak sınıflandırılan pikselleri ifade etmek için bilişim teorisinden bilgi elde etme (information retrieval) alanındaki doğru pozitif (true positive, *TP*), yanlış pozitif (false positive, *FP*), doğru negatif (true negative, *TN*) ve yanlış negatif (false negative, *FN*) ölçümleri kullanılmaktadır. Buna göre; *TP* değeri tespit edilmesi istenilen bölge piksellerinin kaç tanesinin doğru olarak programatik biçimde bu bölgede tespit edilmiş olduğunu gösterilmektedir. *FN* değeri tespit edilmesi istenilen bölge piksellerinin kaç tanesinin yanlış olarak programatik biçimde bu bölgenin dışarısında tespit edilmiş olduğunu gösterilmektedir. *FP* değeri ise tespit edilmesi istenilen bölgenin dışında kalan piksellerin kaç tanesinin yanlış olarak programatik biçimde bu tespit edilmek istenilen bölgenin içerisinde tespit edilmiş olduğunu göstermektedir. Bu ölçümleri içeren *Jaccard* benzerlik endeksi (katsayısı) ölçütü aşağıdaki Eşitlik (2) ile ifade edilmektedir (Tushar, 2018). Bu ölçüt ile [0,1] aralığında değerler elde edilmekle birlikte literatürdeki bazı çalışmalarda ise bu değerler yüzdelik orana dönüştürülerek de verilebilmektedir.

$$Jaccard = \frac{|TP|}{(|TP|+|FP|+|FN|)} \quad (2)$$

Elde edilen deneysel sonuçlarda büyük oranda kulak bölgesinin doğru konumunda ve tam olarak bölütlendiği görülmektedir. Şekil 3'deki satır numaraları ile belirtilen sırasıyla 1) "002_up_ear", 2) "020_front_ear", 3) "056_zoom_ear" ve 4) "090_zoom_ear" isimli görüntüler deneylerde kullanılan görüntüler arasından rasgele seçilmiştir. Bu dört görüntüye ait *Jaccard* benzerlik endeksi değerleri, aşağıdaki Şekil 3'te verilmektedir.



Şekil 3. Jaccard benzerlik endeks değerleri.

Şekil 3'te (a) orijinal görüntü, Şekil 3 (b) referans (*ground-truth*) görüntü, Şekil 3 (c) çalışmamızda önerilen yaklaşımın ürettiği çıktı maske görüntüsü ve Şekil 3 (d) ilgili görüntü için *Jaccard* endeks değeri verilmektedir. Referans (*ground-truth*) olarak uzman tarafından el ile işaretlenen arka planı (siyah) ve ön planı (beyaz) ifade eden görüntü kullanılmıştır. Bu referansla yapılan kıyaslamada nesneye karşılık gelen programatik olarak çalışmamızda üretilmiş çıktı maske görüntüsünün benzerlik oranı (örtüşümün birleşime oranı) veya birleşim kesişimi (Intersection over Union, *IoU*) olarak tarif edebileceğimiz *Jaccard* endeks değerlerine bakıldığında (bu değerler yüzdelik orana dönüştürüldüğünde) %84 ilâ %92 arasında bir birleşim kesişimi (benzerlik oranı) ile deneylerde referans veri ile karşılaştırıldığında oldukça iyi sonuçlar elde edildiği görülmektedir. AMIDB veri kümesinde sırasıyla *subset-1*, *subset-2*, *subset-3* ve *subset-4* isimli dört ayrı alt küme bulunmaktadır. Tüm bu alt kümelerdeki toplamda 700 adet görüntü üzerinden *ortalama Jaccard* endeks değeri (*IoU*) hesaplandığında standart sapma miktarı da göz önüne alınarak $86,32 \pm 0,6$ olduğu görülmektedir.

Literatürde AMIDB veritabanından farklı veritabanları için bu ölçütü veya farklı ölçütlerle deneysel sonuçlar elde eden çalışmalar da mevcuttur (Wang, Yang ve Zhu, 2021). Emeršič ve ark. (2015) çalışmasında belirli bir koşula bağlı olmayan ortamdan alınarak etiketlenmiş ağ kulak görüntüleriyle oluşturulan AWE veri kümesi (Emeršič, Štruc ve Peer, 2017) üzerinde piksel çapında kulak tespitine dayanan kodlayıcı kod çözücü (encoder-decoder) sinir ağı ile yapılan deneylerle bazı *Jaccard* endeksi (*IoU*) değerleri elde edilmiştir. Emeršič ve ark. (2015) çalışmasında AWE veri kümesi için ortalama *Jaccard* endeks değeri olarak $55,7 \pm 25,0$ elde edilmiştir. AMIDB veritabanı ile AWE veri kümesi arasındaki temel fark ise AWE veri kümesinin AMIDB veritabanının aksine koşula veya kısıta bağlı olmaksızın (*in-the-wild*) oluşturulmuş olmasıdır. Böyle görüntülerde insanın kafa bölgesinin yan profilden görünmesi veya yüz bölgesinin tamamen görünür olup olmaması kulak bölgesinin tespitini, bölütlenmesini veya tanınmasını daha farklı bir konu başlığı altında incelenebilecek hale getirmektedir (Wang, Yang ve Zhu, 2021).

Birçok çalışmada kulak bölütleme; kulak tespiti veya kulak tanımanın bir ilk adımı olarak görülmekte, basit eşikleme ile bölütleme veya elle işaretlenen bölgeler üzerinde aktif çevritin işletilmesiyle kulak konumlandırma (localization) yapılmaktadır. Bu çalışmalarda yoğunlukla, tespit doğruluk oranı, yanlış tespit oranı veya benzeri nesnel ölçütler üzerinden bir başarımlı değerlendirme yapılmaya çalışılmaktadır. Doğruluk oranı (accuracy) ölçütü *Jaccard* endeksi yanı sıra çalışmalarda başarımlı değerlendirme için kullanılmaktadır. Aşağıdaki Eşitlik (3) ile doğruluk oranı ifade edilmektedir (Emeršič, Gabriel, Štruc ve Peer, 2015). Bu ölçüt ile [0,1] aralığında değerler elde edilmekle birlikte literatürdeki bazı çalışmalarda ise bu değerler yüzdelik orana dönüştürülerek de verilebilmektedir.

$$Doğruluk = \frac{|TP|+|TN|}{(|TP|+|TN|+|FP|+|FN|)} \quad (3)$$

Yakından bakıldığında Eşitlik (2) ile Eşitlik (3) arasındaki temel farkın doğru negatif (*TN*) terimi olduğu görülmektedir. *TN* değeri tespit edilmesi istenilen bölgenin dışında kalan piksellerin kaç tanesinin doğru olarak programatik biçimde bu bölgenin dışarısında tespit edilmiş olduğunu göstermektedir. Buna göre, doğruluk oranı ile *Jaccard* endeksi arasında temel farklılık ilgili bölgeler için bakılan örtüşüm oranlarını ne açıdan değerlendirdikleridir. Doğruluk oranı tüm görüntüde kaç tane pikselin doğru olarak sınıflandırıldığını göstermektedir. Çalışmamızda başarımlar değerlendirilirken, görüntüdeki piksellerin kulak bölgesine ait olarak ne derecede doğru sınıflandırılmış olduğunun oransal olarak ölçülmesi hedeflenmiştir. Buna göre çalışmamızdaki sinerjik yaklaşımla elde edilen kulak bölgesi ve kulak bölgesi dışında kalan bölgelerdeki (saç, cilt vb.) pikseller değerlendirilmede dikkate alınmıştır. Doğruluk oranı görüntü içerisinde dengesiz olarak dağılan piksellerden oluşan bölgeler için daha uygun bir ölçüttür. Kulak bölgesi dışında kalan bölgeler (çoğunluk bölgesi) daha fazla pikselin bulunduğu alanları, kulak bölgesi (azınlık bölgesi) ise daha az pikselin bulunduğu alanı ifade ettiği durumlarda doğruluk oranı daha kullanışlı olmaktadır. Literatürdeki çoğu veritabanının aksine AMIDB veritabanı koşullu, kısıtlar altında oluşturulmuş ve bahsi geçen bu iki bölgenin veritabanındaki görüntüden görüntüye farklılık gösterdiği bir yapıdadır. Bazı görüntülerde kulak bölgesi pikselleri çoğunluk sınıfındayken bazılarında ise azınlık sınıfında olabilmektedir. Bu açıdan *Jaccard* endeksi (*IoU*), çoğunluk ve azınlık sınıfları arasında piksellerin dağılımından daha az etkilenmesi nedeniyle daha doğru bir ölçüm yapmamızı sağlamaktadır (Emeršič, Gabriel, Štruc ve Peer, 2015). Çalışmamızda nihai çizge kesme adımıyla bölge tabanlı bölütleme (region-based segmentation) yapıldığı ve çoğunluk ile azınlık sınıflarına ait piksellerin AMIDB veritabanında çeşitli alt kümelerde (*subset-1*, *subset-2*, *subset-3* ve *subset-4*) farklı bölgesel oranlara sahip görüntülerde farklı sınıflara düşmesi nedeniyle özellikle doğruluk oranı yerine *Jaccard* endeksi (*IoU*) tercih edilmiştir. Bu açıdan çalışmamız AMIDB özelinde *Jaccard* endeksini kullanması nedeniyle literatürdeki diğer çalışmalardan farklı bir özgün bakış açısına sahiptir.

4. SONUÇ

Yapılan deneylerde elde edilen görsel nitel (qualitative) sonuçlarda da aynı zamanda nesnel nicel (quantitative) nesnel benzerlik oranı ölçütü (*Jaccard* endeksi) değerlerinde de oransal olarak kulak bölgesinin konumunun doğru olarak bulunması ve programatik yoldan oluşan nesne sınırları tarafından tam olarak kapsanır oluşu, kullanılan yöntemlerin parametrik değerlerine göre etkin sonuçlar üretmesi ve biyometrik alanında otomatik kulak bölütlemenin popüler bir araştırma konusu olması nedeniyle çalışmamız kendi kapsamında e-sağlık ve turizm alanlarında bireylere uygulanacak biyometrik yetkilendirme ve tanıma ile alakalı diğer çalışmalara temel oluşturabilecek bir yapıdadır.

Çalışmamızdaki deneylerden anlaşılacağı üzere uygun olmayan ön işleme algoritmasının oluşturulan sinerjiyi olumsuz etkileyebileceği, böyle bir sinerjinin en çok ilk adımlarda yapılan kümelemeye bağımlı olarak başarımının değişkenlik gösterebileceği anlaşılmaktadır. Önerilen sinerjik yaklaşımdaki aktif çevritin görüntüdeki gürültü olarak tarif edebileceğimiz çeşitli piksellerden ve azınlık sınıfına girmesi gerekirken çoğunluk sınıfına giren veya tam tersi durumdaki piksellerden etkilendiği görülmektedir. Bu açıdan oluşturulan süperpiksel kümelerinin olabildiğince homojen bölgeleri ifade etmesine dikkat edilmelidir. Buna göre SLIC ile elde edilebilecek kümelerde bölge içi homojenlik oluşması adına renk yaklaşıklığı ve uzamsal yaklaşıklık için yoğunluk parametresi etki ettirilerek aktif çevritin işletilmesi sağlanmıştır. Bu esnada oluşan maskenin gürültü oluşturacak piksellerden daha az etkilenmesi mümkün olmuştur. Böylece daha uygun bir çevritin elde edilmesiyle bölge tabanlı görüntü bölütlemeye daha uygun bir maske elde edilmesi için GrabCut kullanan çizge kesmeye verilecek kapalı dolgu alanının oldukça iyi biçimde belirlenmesi mümkün olmaktadır. Bu olgular çerçevesinde yapılan çalışmamızda deneysel sonuçlar başarımları açıkça ortaya koymaktadır.

Bu çalışma özelinde elde edilen sonuçlara dayanarak, ileriki çalışmalarda daha fazla çeşitte kulak görüntü veri kümelerinin üzerinde deneylerin yapılması, farklı kümeleme ve aktif çevrit yöntemlerinin denenerek ön plan ve arka plan ayırımında daha yüksek oranda kapsama ve kulak harici bölgelerin kulaktan daha yüksek doğrulukta ayrıştırılması sonuçlarının elde edilmesi planlanmaktadır.

Hakem Değerlendirmesi: Dış bağımsız.

Çıkar Çatışması: Yazar çıkar çatışması bildirmemiştir.

Finansal Destek: Yazar bu çalışma için finansal destek almadığını beyan etmiştir.

Peer-review: Externally peer-reviewed.

Conflict of Interest: The author has no conflict of interest to declare.

Grant Support: The author declared that this study has received no financial support.

Kaynaklar/References

- Achanta, R., Shaji, A., Smith, K., Lucchi, A., Fua, P. & Süsstrunk, S. (2012). SLIC Superpixels Compared to State-of-the-Art Superpixel methods. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 34(11), 2274–2282. <http://dx.doi.org/10.1109/TPAMI.2012.120>.
- Banerjee, S. & Mery, D. (2016). Iris Segmentation Using Geodesic Active Contours and GrabCut. *Revised Selected Papers of Image and Video Technology, PSIVT 2015 Workshops, LNCS 9555*, (pp. 48-60). Auckland, New Zealand, (Eds., Huang, F., & Sugimoto, A.), Springer International Publishing Switzerland. http://dx.doi.org/10.1007/978-3-319-30285-0_5.
- Chen, H., & Bhanu, B. (2005, January). Contour matching for 3D ear recognition. *Proceedings of the Seventh IEEE Workshop on Applications of Computer Vision Motion and Video Computing (WACV/MOTION2005)* (pp. 123-128). Breckenridge, CO, USA. <http://dx.doi.org/10.1109/ACVMOT.2005.38>.
- Cintas, C., Delriux, C., Navarro, P., Quinto-Sánchez, M., Pazos, B. & Gonzalez-José, R. (2019). Automatic Ear Detection and Segmentation over Partially Occluded Profile Face Images. *Journal of Computer Science Technology*, 19(1), 81-89. <http://dx.doi.org/10.24215/16666038.19.e08>.
- Cohen, L. D. (1991). On Active Contour Models and Balloons. *Computer Vision, Graphics, and Image Processing: Image Understanding*, 53(2), 211-218. [http://dx.doi.org/10.2016/1049-9660\(91\)90028-N](http://dx.doi.org/10.2016/1049-9660(91)90028-N).
- El-Naggar, S., Abaza, A., & Bourlai, T. (2018, August). Ear detection in the wild using Faster R-CNN Deep Learning. *Proceedings of the IEEE/ACM International Conference on Advances in Social Networks Analysis and Mining (ASONAM2018)* (pp. 1124-1130). Barcelona, Spain. <http://dx.doi.org/10.1109/ASONAM.2018.8508487>.
- Emeršič, Ž., Gabriel, L. L., Štruc, V., & Peer, P. (2015). Convolutional Encoder-Decoder Networks for Pixel-wise Ear Detection and Segmentation. *IET Biometrics, Special Issue: Unconstrained Ear Recognition*, 7(3), 175-184. <http://dx.doi.org/10.1049/iet-bmt.2017.0240>.
- Emeršič, Ž., Štruc, V., & Peer, P. (2017). Ear Recognition: More than a Survey. *Neurocomputing*, 255, 26-39. <http://dx.doi.org/10.1016/j.neucom.2016.08.139>.
- Gonzalez-Sánchez, E. (2008). Biometria de la Oreja. (Doctoral Dissertation, Universidad de Las Palmas de Gran Canaria, Spain). Mathematical Analysis of Images Database (AMIDB). http://ctim.ulpgc.es/research_works/ami_ear_database.
- Jacob, L., & Raju, G. (2011). Automatic Ear Localization Using An Effective Skin Segmentation Algorithm And Correlation Coefficient in 2D Images. *International Journal of Machine Intelligence (IJMI)*, 3(4), 327-332.
- Joshi, K. V., & Chauhan, N. C. (2011). Edge Detection and Template Matching Approaches for Human Ear Detection. *International Journal of Computer Applications, Special Issue for International Conference on Intelligent Systems and Data Processing (ICISD2011)*, 1(1), 50-55.
- Khattab, D., Ebied, H. M., Husein, A. S., & Tolba, M. F. (2014). Color Image Segmentation Based on Different Color Space Models Using Automatic GrabCut. *The Scientific World Journal*, 2014, 126025,1-10. <http://dx.doi.org/10.1155/2014/126025>.
- Márquez-Neila, P., Baumela, L., & Alvarez, L. (2014). A Morphological Approach to Curvature-based Evolution of Curves and Surfaces. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 36(1), 2-17. <http://dx.doi.org/TPAMI.2013.106>.
- Pflug, A., & Busch, C. (2012). Ear Biometrics: A Survey of detection, feature extraction and recognition methods. *IET Biometrics*, 1(2), 114-129. <http://dx.doi.org/10.1049/iet-bmt.2011.0003>.
- Schuckers, M. E., (2010). *Computational Methods in Biometric Authentication, Statistical Methods for Performance Evaluation*. Information Science and Statistics. Springer-Verlag London Limited.
- Tomczyk, A., & Szczepaniak, P. S. (2019). Ear Detection using Convolutional Neural Network on Graphs with Filter Rotation. *Sensors (MDPI)*, 19, 5510, 1-21. <http://dx.doi.org/10.3390/s19245510>.
- Tushar, F. I., (2018). Automatic Skin Lesion Segmentation Using GrabCut in HSV Colour Space. *Computer Vision and Pattern Recognition (cs.cv), ArXiv. 1-3*. Retrieved from: <https://arxiv.org/abs/1810.00871>.
- University of Notre Dame. UND Face Database. (2015). Erişim adresi: <http://www.nd.edu/cvrl/CVRL/DataSets.html>. (Erişim zamanı: 12.03.2021).
- University of Science and Technology Beijing. Ear Recognition Laboratory USTB Database. (2002). Erişim adresi: http://www1.ustb.edu.cn/resb/en/doc/Imagedb_123_intro_en.pdf. (Erişim zamanı: 12.03.2021).
- Walt, S. V. D., Schönberger, J. L., Nunez-Iglesias, J., Boulogne, F., Warner, J. D., Yager, N., Gouillart, E., & Yu, T. (2014). Sckit-Image: Image Processing in Python. *Peer J. Life & Environment*, 2(e453), 1-18. <http://dx.doi.org/10.7717/peerj.453>
- Wang, Z., Yang, J., & Zhu, Y. (2021). Review of Ear Biometrics. *Archives of Computational Methods in Engineering*, 28, 149-180. <http://dx.doi.org/10.1007/s11831-019-09376-2>.

Formel Yaklaşımlar ile Sürüm Değerlendirme

A Formal Methods Approach for Release Evaluation

Ebru Aydın Göl¹ 



ÖZ

Bu çalışmada sürüm değerlendirme sürecini otomatikleştirmek için formel metotlar kullanılarak bir yöntem geliştirilmiştir. Sürüm değerlendirme sürecinde, bir sunucunun yeni ve eski sürümleri benzer (veya aynı) konfigürasyonlarda çalıştırılır ve sistemlerin ürettikleri izler karşılaştırılır. Bu karşılaştırma grafik incelenmesi, izlerin sabit ölçütler ile karşılaştırılması veya regresyon analizi tabanlı olabilir. Bu makalede ise, yapılan önceki çalışmalardan farklı olarak formel metotlar tabanlı bir analiz yöntemi sunulmaktadır. Bu yöntem karşılaştırılacak sistemlerden verilerin toplanması, her iki veri kümesi için bu kümeleri tanımlayacak Sinyal Zamansal Mantık (STL) formüllerinin üretilmesi ve son olarak da yüksek başarımla kümeleri tanımlayabilen aynı yapıya sahip formüllerin karşılaştırılması ile sürüm değerlendirme sonucunun üretilmesi adımlarından oluşmaktadır. Bu yöntem ile sürüm değerlendirmede kullanılmak üzere, ölçütlerin STL formülü olarak ifade edilmesi ve bu ölçütlerin sistem izlerinden otomatik olarak üretilmesi sağlanmıştır. Zamansal mantıkların konuşma diline benzerlikleri sayesinde bu formüller açıklayıcıdır. Geliştirilen metot ile değerlendirme süreci otomatikleştirilmektedir. Elde edilen sonuçlar örnek veri kümeleri üzerinde incelenmiştir.

Anahtar Kelimeler: Formel Metotlar, Sürüm Değerlendirme, Formül Türetme, Zamansal Mantık

ABSTRACT

In this paper, a formal method-based release evaluation method was developed. During the release evaluation process, two versions of a server are run under similar (or the same) configurations and the system logs are compared. This comparison can be based on graphical analysis, applying fixed rules over logs or regression analysis. This paper presents a novel release evaluation approach based on formal methods. The proposed method consists of three main steps. The first step is the collection of data from both versions of the server. The second step is the synthesis of Signal Temporal Logic (STL) formulas for each dataset. The final step is the generation of the release evaluation result by comparing the formulas that have the same structure and that can represent both datasets with high accuracy. Thus, the proposed approach represents the release evaluation rules as STL formulas and generates such formulas from system logs in an automated way. Due to the resemblance of temporal logics to natural language, the resulting formulas explain the evaluation result. The proposed method automates the release evaluation process. The findings of the paper are shown over sample datasets.

Keywords: Formal Methods, Release Evaluation, Formula Synthesis, Temporal Logics

¹(Dr. Öğr. Üyesi), Orta Doğu Teknik Üniversitesi, Bilgisayar Mühendisliği Bölümü, Ankara, Türkiye

ORCID: E.A.G. 0000-0002-5813-9836

Corresponding author:

Ebru AYDIN GÖL
Orta Doğu Teknik Üniversitesi, Bilgisayar
Mühendisliği Bölümü, Ankara, Türkiye
E-mail address: ebrugol@metu.edu.tr

Submitted: 12.02.2021

Revision Requested: 02.04.2021

Last Revision Received: 15.04.2021

Accepted: 17.05.2021

Published Online: 29.06.2021

Citation: Aydın Gol, E. (2021). Formel yaklaşımlar ile sürüm değerlendirme. *Acta Infologica*, 5(1), 129-140. <https://doi.org/10.26650/acin.879052>

1. GİRİŞ

Günümüzde e-devlet, e-posta, e-alışveriş ve e-gazete gibi kullanıcıya doğrudan hizmet veren birçok servis elektronik ortamda internet tabanlı olarak sunulmaktadır. Bu tip servisler için güvenilirlik ve ulaşılabilirlik kritik öneme sahiptir. Diğer taraftan, sunucu tabanlı servis sağlayan bu sistemler düzenli olarak güncellenmekte, yani sunuculara yeni versiyonlar sürülmektedir. Bu işlem, sürüm yönetimi ile idame ettirilmektedir (Bays, 1999). Bir sürüm yönetimi döngüsünde takip edilen ana adımlar yeni versiyona eklenecek özelliklerin belirlenmesi, geliştirme planının yapılması, geliştirme ve test faaliyetlerinin yapılması, sürüm değerlendirme ve son olarak bütün sunucuların güncellenmesi olarak sıralanabilir. Bu adımlar içinde, test ve sürüm değerlendirme adımları sistemin güvenilirlik ve ulaşılabilirlik gibi özelliklerinin korunması için çok önemlidir (Sommerville, 2015).

Sürüm değerlendirme işleminde hedef, sistemin bütünü veya bir bileşeninin yeni bir versiyonu geliştirildiğinde, bu versiyonda yapılan değişikliklerin beklenmeyen etkilerinin olup olmadığını saptamaktır. Bu amaçla, yeni sürüm ve eski sürüm benzer (veya aynı) konfigürasyonlarda çalıştırılır ve sistemlerin ürettikleri metrikler karşılaştırılır (Howard, 2016). Bu değerlendirmede, A ve B versiyonları birbirleri ile karşılaştırıldığı için “A/B testi” adı da verilmektedir (Chatterjee & Simonoff, 2013). Bu değerlendirme etkilenen kullanıcı sayısına göre farklı seviyelerde yapılmaktadır. Bu sıralamada amaç, yeni versiyonun sürülmesini engelleyecek bir durumu olabildiğince az kullanıcıyı etkileyerek saptamaktır. Şekil 1’de örnek bir sürüm değerlendirme süreci verilmiştir. Burada kırmızı kutular ile gösterilen adımlarda, gerçek kullanıcıya hizmet eden sunucular üzerinde değerlendirme yapılmaktadır. Sürüm değerlendirme süreçlerinde, genel olarak en son değerlendirme, en son geçerli sürümün bulunduğu sunucuların konfigürasyonuna benzer bir konfigürasyona sahip sunucu üzerinde yapılır ve buna kanarya testi adı verilir. Bu aşamalardan herhangi birinde bir problem gözleendiğinde ileriki aşamalara geçilmez, sürüm geri çekilir ve sorun kaynağının araştırılması süreci başlatılır. Sistemin özelliklerine ve sürüm güncelleme sıklığına göre bu aşamalardan bir kısmı uygulanmayabilir veya farklı aşamalar eklenebilir.



Şekil 1. Örnek bir sürüm değerlendirme süreci.

Her bir sürüm değerlendirme adımında, değerlendirilen yeni versiyon ve bir önceki versiyonu çalıştıran iki yazılımın ürettikleri metrikler belirli ölçütlere göre değerlendirilerek beklenmeyen durumlar saptanmaya çalışılır. Yaygın olarak, yazılımlardan zaman serisi verisi toplanır. Bu veri üzerinde ölçütler sinyal filtreleri (en büyük değer/ en küçük değer gibi) sonuçlarının karşılaştırılması veya istatistiksel özelliklerin (ortalama, süregelen ortalama, standart sapma gibi) karşılaştırılması olarak tanımlanır. Örneğin, bazı değişiklikler sonucu ortalama CPU kullanımının artması (veya azalması) beklenmez. Bunun yanı sıra sunucuların CPU kullanım miktarları gün içinde kullanıcı aktivitelerine göre değişebilmektedir. Bu nedenle, CPU kullanımında beklenmeyen bir değişiklik olup olmadığını kontrol edebilmek için ‘iki sunucunun ortalama CPU kullanımı arasındaki fark 5 dakika boyunca %5’in üzerine çıkmaz’ gibi bir ölçüt kullanılabilir. Yapılan karşılaştırmalar sonucunda, ulaşılan fark önceden belirlenen bir değer üzerinde ise beklenmeyen bir durum olduğu sonucu çıkarılıp kök neden analizi adımına geçilir. Aksi durumda, sürüm değerlendirme sürecinin sonraki adımı ile devam edilir (bkz. Şekil 1).

Ölçüte dayalı değerlendirme çok hızlı yapılabilse de ölçütlerin tanımlanması uzmanlık gerektiren zor bir iştir ve olası hataların maliyeti yüksek olur. Örneğin sürüm değerlendirme adımında yakalanamayan bir durum son kullanıcıyı etkileyebilir. Diğer taraftan fazla hassas olarak belirlenen ölçütler aslında gereği olmayan analiz çalışmalarının yapılmasına yol açıp, yazılım mühendislerine zaman kaybettirebilir ve sürüm planının aksamasına yol açabilir. Birçok uygulamada sürüm değerlendirme adımı ilgili mühendisler tarafından manuel olarak yapılmaktadır. Bu yaklaşımda, zaman serisi verisi olarak üretilen metriklerden oluşturulan grafikler incelenmekte ve incelemeyi yapan kişi tarafından sonuç belirlenmektedir. Uzmanlar, sadece bir metriktaki değişikliğe göre değil, takip edilen bütün metrikleri birlikte değerlendirerek ek çıkarımlar yapabilmektedir. Örneğin bir uygulama sunucusu için, sunucunun yeniden başlatıldığı durumlarda, diğer metriklerin değerlerine bakılmayabilir veya kullanıcı sayısında anlık değişiklikler oldu ise diğer metriklerdeki bazı değişiklikler için hata değerlendirmesi yapılmayabilir. Uzmanların yaptığı bu tip değerlendirmeleri, otomatik olarak yapabilmek için birden fazla metriğe dayalı ölçütler tanımlamak gerekmektedir. Diğer taraftan uzman incelemesi hem kişi hatasına açık hem de zaman alan bir adımdır.

Basit ölçüt fonksiyonları yerine, sürüm analizinde kullanılan bir diğer yaklaşım ise regresyon analizidir (Chatterjee & Simonoff, 2013). Burada da uygun analiz yöntemini (doğrusal, çok terimli), uygulanacak ön işlemi (ortalama alma, sapma hesaplama gibi) ve sonuçta elde edilecek değerler için limitleri seçmek uzmanlık gerektirir ve bu adımlar her bir metrik için tekrarlanmalıdır. Yukarıda verilen örnekteki gibi birçok metriktaki değişikliklere göre çıkarım yapılması regresyon analizine doğrudan eklenemez ve ikincil değerlendirmelerin tanımlanmasını gerekebilir. Sürüm değerlendirme problemi, yeni sürümde öncekine göre herhangi bir anomali var mı şeklinde ele alınarak, makine öğrenmesi tabanlı metotlar da uygulanabilir. Fakat bu yaklaşım, regresyon analizinde olduğu gibi, metriklerin karakteristik özellikleri göz önünde bulundurularak, her bir metrik için farklı bir model geliştirilmesini gerektirebilir.

Bu çalışmada, sürüm değerlendirme ölçütlerinin formel dillerde ifade edilmesi ve bu ölçütlerin, değerlendirme başarısını en iyileştirecek şekilde otomatik olarak üretilmesi amaçlanmaktadır. Sistemlerin ürettikleri izlerin, gerçek zamanlı ve gerçek değerli sinyallerin özelliklerini tanımlamak için geliştirilmiş olan Sinyal Zamansal Mantık (STL) formülleri ile tanımlanması (Donzé, 2013), formüllerin arasındaki farklar üzerinden sürüm değerlendirme sonucu yapılması hedeflenmektedir. Önerilen yöntem şu temel adımlar ile özetlenebilir. İlk adım, A ve B versiyonlarından veri toplanmasıdır. İkinci adım her iki veri kümesi için de bu kümeleri en uygun şekilde tanımlayacak STL formüllerinin bulunmasıdır. Üçüncü adım ise, bir önceki adımda her iki veri kümesini de başarı ile tanımlayan formül yapılarının belirlenip, optimize edilmiş formüllerin parametrelerinin karşılaştırılması ile sürüm değerlendirme sonucunun üretilmesidir.

STL formülleri ile sinyal özellikleri konuşma diline yakın bir yazım ile ifade edilebilir. Örneğin $F_{[0,10]} x > 10$ formülü, takip eden on zaman biriminden en azından birinde x değerinin 10 üzerine çıkması gerektiğini, $x > 0 \ U_{[0,10]} y = 1$ formülü ise takip eden on zaman biriminden birinde y 'nin bir olması gerektiğini ve bu ana kadar x 'in 0'dan büyük olması gerektiğini ifade eder. Zamansal ve Boole mantık operatörlerinin iç içe kullanılması ile daha karmaşık formüller elde edilebilir. Fakat, formül yapısının belirlenmesi ($x > p_1 \ U_{[p_2,p_3]} y = p_4$) ve bu yapıdaki parametrelerin veriye uygun bir şekilde üretilmesi başlı başına zor bir problemdir.

Sistem izlerinden (zaman serisi verisi), bu izleri tanımlayacak STL formüllerinin üretilmesi üzerinde yakın zamanda birçok çalışma yapılmıştır (Aydın & Aydın Gol, 2020; Bartocci, Bortolussi & Sanguinetti, 2014; Bombara, Vasile, Penedo, Yasuoka & Belta, 2016; Ergurtuna & Aydın Gol, 2019; Hoxha, Dokhanchi & Fainekos, 2018; Jha, Tiwari, Seshia, Sahai & Shankar, 2019; Jin, Donzé, Deshmukh & Seshia, 2015; Ketenci & Aydın Gol, 2019; Kong, Jones, Ayala, Aydın Gol & Belta, 2014; Mohammadinejad, Deshmukh, Puranic, Vazquez-Chanlatte & Donzé, 2020; Yoo & Belta, 2017). Bu çalışmalar, veri kümesinin yapısı ve üretilen formül açısından farklılık göstermektedir. Veri kümesindeki bütün izleri tanımlayacak bir formül üretilmesi problemi gereksinim madenciliği olarak da tanımlanmaktadır: örneğin eski (legacy) bir sistemin sağladığı gereksinimin otomatik olarak bulunması (Hoxha et al., 2017; Jha et al., 2019; Jin et al., 2015). Bu konudaki bazı çalışmalarda, formül yapısı kullanıcı tarafından belirlenmiş ve belirlenen formül yapısı için parametre bulunmuştur (Hoxha et al., 2017; Jin et al., 2015). İyi ve kötü olarak etiketlenmiş sistem izlerinden, bu izleri doğru şekilde sınıflandırabilecek bir formül bulunması hem gerçek zamanlı sistem izleme hem de sınıflandırıcı üretmek uygulamaları için çalışılmıştır (Aydın & Aydın Gol, 2020; Bartocci et al., 2014; Bombara et al., 2016; Ergurtuna & Aydın Gol, 2019; Ketenci & Aydın Gol, 2019; Kong et al. 2014; Mohammadinejad et al., 2020; Yoo & Belta, 2017).

Bu çalışmanın literatüre ana katkısı formel metotlara dayalı bir sürüm değerlendirme sistemi tanımlamasıdır. Bu sistemin tamamen otomatik olması, elde edilen sonuçların kullanıcıları için açıklayıcı nitelikte olması ve ulaşılan sonucun geçerliliğine dair nicel değerler sağlanması, geliştirilen yöntemin en önemli özellikleri arasındadır. Temel olarak, daha önce bahsedilen ölçüt belirlenmesi, regresyon analizi veya makine öğrenmesi tabanlı yaklaşımlar, her bir metrik için bir ön çalışma ile sürüm değerlendirmede uygulanacak olan modelin (ölçüt sınırları, regresyon parametreleri, anomali tanıma modeli gibi) belirlenmesini gerektirir. Diğer taraftan, geliştirilen değerlendirme sistemi bu tip ön çalışmalar gerektirmez. Sonuçlar formül tabanlı oluşturulduğu için, açıklayıcı niteliktedir. Bu da makine öğrenmesi tabanlı metotlara kıyasla bir avantaj sunmaktadır. Ek olarak, geliştirilen yöntem, iki sistemin karşılaştırılması üzerine formel yaklaşımlar tabanlı bir yöntem olarak daha geniş bir pencereden değerlendirildiğinde, sürüm değerlendirme dışında farklı problemlere de uygulanabilir (örneğin enjeksiyon/kötü niyetli sistem tespiti).

Makalenin yapısı. 2. Bölüm'de STL formüllerinin söz-dizim kuralları ve semantiği verilmiş, bu bilgiler ışığında formül sentezleme problemi anlatılmış ve literatürde bulunan yaklaşımlar özetlenmiştir. Takip eden kısımda ise, önerilen formel yaklaşımlar ile sürüm değerlendirme yöntemi anlatılmış ve bu yöntemde izlenecek her bir adım detaylandırılmıştır. 4. Bölüm'de, örnek veri kümeleri tanımlanmış ve bu örnek durumlar üzerinde elde edilen sonuçlar anlatılmıştır. Son kısımda ise, yapılan çalışma özetlenmiş ve ileriki çalışma yönleri sunulmuştur.

2. SİNYAL ZAMANSAL MANTIK VE FORMÜL SENTEZİ

Bu kısımda geliştirilen sürüm değerlendirme sisteminde kullanılan araçlara yönelik ön bilgiler sunulmuştur.

2.1 Sinyal Zamansal Mantık

Sonlu kesikli ve n-boyutlu bir sinyal $x = x_0, x_1, \dots, x_N$ dizisi şeklinde gösterilir. Sinyalin i . değişkeni t -kesikli zamanında x_t^i notasyonu ile gösterilir.

Geçmiş zamanlı sinyal zamansal mantık (past time Signal Temporal Logic - ptSTL) için söz-dizim kuralları Denklem (1)'deki gibi tanımlanır (Donzé, 2013; Gabbay, 1989).

$$\phi = T[x^i \sim c | \neg\phi | \phi \wedge \phi | \phi \vee \phi | \phi S_{[a,b]} \phi | F_{[a,b]}^- \phi | G_{[a,b]}^- \phi] \quad (1)$$

Bu denklemde T mantık sabiti "doğru", x^i sinyal değişkeni, $\sim \in \{<, \leq, >, \geq\}$ bir eşitsizlik sembolü, c sabit bir değer, \neg -değil, \wedge -ve, \vee -veya ise standart mantık operatörleridir. $S_{[a,b]}$ (olduğundan beri), $F_{[a,b]}^-$ (geçmişte bir zaman) ve $G_{[a,b]}^-$ (geçmişte her zaman) zaman aralıkları $[a,b]$ olan geçmiş zaman operatörleridir.

Bir ptSTL formülünün semantiği bir sinyal ve belirli bir zaman üzerinde tanımlanmıştır. Özetle, T formülü her zaman sağlanır; $x_t^i \sim c$ eşitsizliği sağlanıyor ise sinyal t anında $x_t^i \sim c$ formülünü sağlar. Mantık operatörleri de standart semantiklerine göre değerlendirilir: sinyal t anında ϕ formülünü sağlamıyor ise, $\neg\phi$ formülünü sağlar; ϕ_1 ve ϕ_2 formüllerini sağlıyorsa, $\phi_1 \wedge \phi_2$ formülünü de sağlar; ϕ_1 ve ϕ_2 formüllerinden en az birini sağlıyorsa, $\phi_1 \vee \phi_2$ formülünü de sağlar. Görüldüğü gibi mantık operatörlerinin anlamlandırmasında içerdikleri formüllerin o andaki değerleri kontrol edilmektedir. Diğer taraftan, zamansal operatörler için içerdikleri formüllerin belirli bir aralıktaki değerlerinden anlamlandırma yapılmaktadır. ϕ formülü $[t - b, t - a]$ aralığında her an sağlanır ise, $G_{[a,b]}^- \phi$ formülü t anında sağlanır. ϕ formülü $[t - b, t - a]$ aralığında herhangi bir anda sağlanır ise, $F_{[a,b]}^- \phi$ formülü t anında sağlanır. ϕ_2 formülü $[t - b, t - a]$ aralığında bir t' anında sağlanır, ve ϕ' formülü o andan itibaren, yani $[t', t]$ aralığında her an sağlanır ise, $\phi_1 S_{[a,b]} \phi_2$ formülü t anında sağlanır. Bir x sinyalinin t anında ϕ formülünü sağlaması $x, t \models \phi$ şeklinde gösterilir. Diğer yandan, x sinyalinin t anında ϕ formülünü sağlamaması, diğer bir deyişle ihlal etmesi, $x, t \not\models \phi$ şeklinde gösterilir.

Bir x sinyalinin bir ptSTL formülünün ne ölçüde sağlayıp sağlamadığına ise nicel değerlendirme ile karar verilir. Nicel değerlendirme hesabı Denklem (2)'de verilmiştir. Bir x sinyali ve ϕ formülü için, t anında nicel değerlendirme sonucu pozitif ise, yani $\rho(\phi, x, t) > 0$ ise, o anda sinyal formülü sağlar. Diğer taraftan değerlendirme sonucu negatif ise, yani $\rho(\phi, x, t) < 0$ ise, o anda sinyal formülü ihlal eder. Nicel değerlendirme sonucunun 0 olduğu durumlarda ise yukarıda özetlenen nitel değerlendirmeye bakılır.

$$\begin{aligned}
\rho(\mathbf{T}, \mathbf{x}, t) &= T \\
\rho(x^i > c, \mathbf{x}, t) &= x_t^i - c \\
\rho(x^i < c, \mathbf{x}, t) &= c - x_t^i \\
\rho(\neg\phi, \mathbf{x}, t) &= -\rho(\phi, \mathbf{x}, t) \\
\rho(\phi_1 \wedge \phi_2, \mathbf{x}, t) &= \min(\rho(\phi_1, \mathbf{x}, t), \rho(\phi_2, \mathbf{x}, t)) \\
\rho(\phi_1 \vee \phi_2, \mathbf{x}, t) &= \max(\rho(\phi_1, \mathbf{x}, t), \rho(\phi_2, \mathbf{x}, t)) \\
\rho(\phi_1 \mathcal{S}_{[a,b]} \phi_2, \mathbf{x}, t) &= \max_{j \in [t-b, t-a]} \left(\rho(\phi_2, \mathbf{x}, j), \min_{k \in [j+1, t]} \rho(\phi_1, \mathbf{x}, k) \right) \\
\rho(\mathbf{F}_{[a,b]}^- \phi, \mathbf{x}, t) &= \max_{j \in [t-b, t-a]} \rho(\phi, \mathbf{x}, j) \\
\rho(\mathbf{G}_{[a,b]}^- \phi, \mathbf{x}, t) &= \min_{j \in [t-b, t-a]} \rho(\phi, \mathbf{x}, j)
\end{aligned} \tag{2}$$

Nicel değerlendirme sonucu gürbüzlük olarak da isimlendirilir (Jin et al, 2015; Asarin, Donzé, Maler & Nickovic, 2012). Nicel değerlendirme sonucunun mutlak değerinden küçük bir ϵ değeri için, (yani $\epsilon < |\rho(\phi, \mathbf{x}, t)|$), sinyal ϵ kadar sarsılırsa da nitel değerlendirme sonucu değişmez.

Eşitsizliklerde ($x^i \sim c$) veya zaman aralıklarında ($[a, b]$) sabit sayılar yerine parametreler kullanılması ile parametrik ptSTL formülleri tanımlanır (Asarin et al., 2012). Parametrik bir ptSTL formülünün parametrelerine sabit değerler (p) atanarak bir ptSTL formülü ($\phi(p)$) elde edilir. Örneğin $\phi = \mathbf{F}_{[0,p]}^- x^0 > p_2$ parametrik formülünden $p = (p_1 \leftarrow 5, p_2 \leftarrow 6)$ değerleri ile $\phi(p) = \mathbf{F}_{[0,5]}^- x^0 > 6$ formülü elde edilir.

2.2 Formül Sentezi

Formül sentezi problemi, verilen sinyal kümesini en iyi tanımlayacak zamansal mantık formülünün türetilmesi olarak tanımlanır. Bu problem, kullanılan zamansal mantık (STL veya ptSTL), veri kümesinin yapısı, parametrik formüllerin problem girdisi olup olmaması açısından farklı biçimlerde çalışılmıştır (Aydın & Aydın Gol, 2020; Bartocci et al., 2014; Bombara et al., 2016; Ergurtuna & Aydın Gol, 2019; Hoxha et al., 2018; Jha et al., 2019; Jin et al., 2015; Ketenci & Aydın Gol, 2019; Kong et al., 2014; Mohammadinejad et al., 2020; Yoo & Belta, 2017).

Verilen bir sinyal kümesini en iyi şekilde tanımlayacak olan STL formülünü bulmayı hedefleyen formül sentezi problemi gereksinim türetme olarak da adlandırılmıştır (Jin et al. 2015). Bu çalışmada formül sentezleme problemi, verilen parametrik bir STL formülü için, veri kümesini olabildiğince sıkı şekilde tanımlayacak parametre değerlerinin bulunması olarak ele alınmıştır. Her bir sinyal için değerlendirme sinyalin başında yapılmış ve sinyal kümesi üzerinde elde edilen en küçük sinyal değerlendirmesini (gürbüzlük değerini) sıfıra yakın pozitif bir değer yapan parametreler üretilmiştir. Bu sayede veri kümesindeki sinyallerin hepsini olabildiğince net bir şekilde tanımlayan formüle ulaşılmıştır. Bu çalışmada görülen en önemli eksiklik, parametrik formül yapısının tanımlanmasının gerekmesidir. Ek olarak, parametreler arasında bir sıralama yapılması gerekmekte ve parametre optimizasyonu bu sıralamaya göre yapılmaktadır. Sıralamada önde yer alan parametreler aldıkları değerler ile diğer parametreleri kısıtlamakta, standart altı (sub-optimal) sonuçlar elde edilmektedir. Sadece pozitif örneklerden formül türetme üzerine yapılan diğer bir çalışmada bu kısıtların üzerine gidilerek, nicel değerlendirme sonuçlarına dayalı bir sıklık ölçütü tanımlanmış ve bu ölçüt ile parametreler arası sıralama yapılmadan optimizasyon yapılmıştır (Jha et al., 2019). Ek olarak verilen bir parametrik formül kümesi için optimizasyon yapılmış ve bu formüllerin en uygun kombinasyonu üretilmiştir. Bu çalışmada (Jin et al. 2015)'de olduğu gibi tek bir parametrik formül ile kısıtlı kalınmasa da kullanıcının parametrik formül kümesinin tanımlanması gerekmektedir.

Formül sentezleme problemi, iyi ve kötü olarak etiketlenmiş olan sinyalleri ayırt edebilecek bir formül bulma olarak da ele alınmıştır. Bu alandaki çalışmalar da sinyal başına bir etiket olması (Bartocci et al., 2014; Bombara et al., 2016; Kong et al., 2014; Yoo & Belta, 2017) ve her sinyalin her noktası için bir etiket olması, yani etiketlerin zaman serisi verisi olması, olarak ayrılmaktadır (Aydın & Aydın Gol, 2020; Ergurtuna & Aydın Gol, 2019; Ketenci & Aydın Gol, 2019). Bu çalışmalar formül yapısının belirlenmesini, yani parametrik formül tanımlanmasını gerektirmemektedir. Formül yapısının üretilmesi için

karar ağacı tabanlı sentez metotları (Bombara et al., 2016; Ketenci & Aydın Gol, 2019; Yoo & Belta, 2017); yinelemeli formül birleştirme metotları (Kong et al., 2014; Ergurtuna & Aydın Gol, 2019); genetik algoritma tabanlı metotlar (Bartocci et al., 2014; Aydın & Aydın Gol, 2020); ve parametrik formül uzayının sıralanması ve her bir parametrik formül için parametre optimizasyonu yapılması yöntemleri (Aydın & Aydın Gol, 2020; Mohammadinejad et al., 2020) bulunmaktadır.

Bu çalışmada amaç, sürüm değerlendirme için formel metotlara dayalı bir yöntem geliştirmektir. Bu amaca ulaşmak için bir veri kümesini en iyi şekilde tanımlayacak formül sentezlenmiştir. Bu adımda, (Jha et al., 2019; Jin et al. 2015) çalışmalarında olduğu gibi sadece pozitif örnekler kullanılmıştır. Fakat bu çalışmalardan farklı olarak (Aydın & Aydın Gol, 2020; Mohammadinejad et al., 2020) çalışmalarında olduğu gibi parametrik formül uzayı sıralanmıştır ve her bir formül için parametre optimizasyonu yapılmıştır. Ek olarak, formülün veri kümesini tanımlamadaki başarısını ölçmek üzere nicel değerlendirmeye dayalı yeni bir ölçüt tanımlanmıştır. Bu bağlamda, pozitif örneklerden formül sentezi için parametrik formül tanımlanmasını gerektirmeyen özgün bir yöntem ortaya konmaktadır. Bu çalışmanın sonraki aşamalarında ise formüllerin karşılaştırılması için yeni bir formül farkı ölçüm metriği geliştirilmiştir. Bu metrik ile sürüm değerlendirme sonuçları üretilmiştir.

3. ptSTL İLE SÜRÜM DEĞERLENDİRME

Bu çalışmada sürüm değerlendirme sürecini otomatikleştirmek için formel metotlar tabanlı bir yöntem geliştirilmiştir. Bu yöntemin ilk adımı, karşılaştırılacak olan A ve B versiyonlarından veri kümelerinin oluşturulmasıdır (3.1. Bölüm). İkinci adımda ise bu kümeleri en uygun şekilde tanımlayan ptSTL formülleri üretilmiştir (3.2. Bölüm). Üçüncü adım ise, bir önceki adımda her iki veri kümesini de başarı ile tanımlayan formül yapılarının belirlenip, optimize edilmiş formüllerin parametrelerinin karşılaştırılması ile sürüm değerlendirme sonucunun üretilmesidir (3.3. Bölüm).

3.1 Veri Kümesi Oluşturma

Önerilen sürüm değerlendirme metodunun ilk adımı, her bir sürüm için bir sinyal kümesinin oluşturulmasıdır. Bu adım için, öncelikle takip edilecek olan metriklerin belirlenmesi gerekmektedir. Örneğin istek cevap süresi (response time) ile ilgili yapılacak bir değerlendirme için, aynı sunucuya bir dakika içinde gelen istek sayısının da (request rate) göz önünde bulundurulması gerekir.

Takip edilecek metrikler $M = \{1, \dots, n\}$ kümesi ile ifade edilir. Burada her bir indeks $i \in M$ bir metrik için kullanılmaktadır (CPU kullanım yüzdesi gibi). Metrik- i 'nin alabileceği değerler π^i ile gösterilir. Bu metriklerin belirlenmesinin ardından, değerlendirme yapılacak uygulamanın A ve B versiyonları, olabildiğince aynı konfigürasyonlarda çalıştırılır. Sistem günlüklerinden (loglar) belirlenen metriklerin anlık değerleri çıkarılarak Denklem (3)'deki gibi veri kümeleri elde edilir. Burada N bir sinyalin uzunluğunu ifade eder. Bu uzunluk, değerlendirme yapılan sistemin özelliklerine göre bir saat veya bir gün olarak seçilebilir. Farklı versiyonlardan oluşturulan veri kümeleri D_A ve D_B ile gösterilir. Sürüm versiyonları dışındaki faktörlerin etkisini istatistiksel olarak azaltmak için, bir veri kümesi oluşturulurken aynı konfigürasyona sahip birden fazla sunucudan izler kullanılabilir.

$$D = \{\mathbf{x} = x_0, x_1, \dots, x_N \mid \text{her bir } j \in M \text{ ve her bir } i \in \{0, \dots, N\} \text{ için } x_i^j \in \pi^j\} \quad (3)$$

3.2 Formül Sentezi

Önerilen metodun ikinci adımı, ilk adımda üretilen sinyal kümelerini (D_A ve D_B) en iyi şekilde tanımlayan ptSTL formüllerinin üretilmesidir. Bu amaçla, öncelikle bir formülün sinyal kümesine uyumunun tersini gösteren bir ölçüt tanımlanmıştır. Parametrik bir ptSTL formülü için, bu ölçütü en güçleştiren parametre değerleri ızgara araması (grid search) metodu ile hesaplanmıştır. Veri kümesini en iyi tanımlayan ptSTL formüllerinin bulunması için ise, belirli bir operatör sayısına sahip ve değişken olarak takip edilen metrikler (M kümesi) kullanılarak oluşturulabilen bütün parametrik formüller tanımlanmış ve her biri için parametre optimizasyonu yapılmıştır.

3.2.1 Optimizasyon Ölçütü

Bir formülün bir sinyale olan uyumsuzluğu, sinyalin belirli bir noktasından (k) başlanarak elde edilen nicel değerlendirme sonuçlarının mutlak değerlerinin ortalaması olarak tanımlanmıştır.

$$C(\phi, \mathbf{x}) = \frac{\sum_{i=k}^N |\rho(\phi, \mathbf{x}, i)|}{N-k} \quad (4)$$

Bir formülün, bir sinyal üzerindeki uyumsuzluk değerinin düşük olması, formülün sinyalin k - N aralığındaki kısmını her noktada sıkı bir şekilde tanımladığını gösterir. Burada k sayısı, 0 alınabileceği gibi formülde geçen zaman aralıklarına göre de belirlenebilir. Örneğin, $F_{[0,10]}^- G_{[0,3]}^- x^0 > 6$ formülü için $k=13$ değeri, değerlendirmenin ilk yapıldığı noktada da sinyalin uzunluğunun yeterli olmasını sağlar. Geliştirilen metotta elde edilen uyumsuzluk değeri farklı formüllerin karşılaştırılmasında kullanıldığı için, bütün formüllerde aynı k sabiti kullanılmıştır.

Bir formülün sinyal kümesi üzerindeki uyumsuzluk değeri bütün sinyaller üzerindeki uyumsuzluk değerlerinin toplamına eşittir.

$$C(\phi, D) = \frac{\sum_{\mathbf{x} \in D} C(\phi, \mathbf{x})}{\sum_{\mathbf{x} \in D} 1} \quad (5)$$

Denklem (5)'de tanımlanan başarımlı ölçütü, literatürde sadece pozitif örneklerden formül türetme çalışmalarında kullanılan ölçütlerden farklılık göstermektedir. Temel olarak, Jin et al. (2015) gelecek zamansal operatörlerini kullanmış ve sadece sinyalin başında değerlendirme yapmıştır. Ek olarak, bahsedilen çalışmada amaç bütün sinyallerin sıkı bir şekilde sağladığı bir formül türetmektir. Bu nedenle, pozitif bir c değeri için, $\bigwedge_{\mathbf{x} \in D} \rho(\phi, \mathbf{x}, 0) > c$ koşulunu sağlayan ve c değerini en küçükleştiren parametreler aranmıştır. Bu çalışmada ise, değerlendirme sinyalin her noktasında yapılarak sürüm değerlendirmede kullanılmak üzere sinyal karakterini tanımlayan bir formül aranmıştır. Ek olarak, nicel değerlendirme sonucunun pozitif olması koşulu aranmamaktadır. Sinyalin gürültülü olabileceği de göz önünde bulundurularak, Denklem (4)'de gösterildiği gibi, sinyal boyunca birçok noktada nicel değerlendirme yapılmış ve değerlendirme sonuçlarının mutlak değerlerinin ortalaması hesaplanmıştır. Bu şekilde tanımlanan ölçüt, sinyali her noktada olabildiğince sıkı şekilde tanımlayan formüller için küçük değerler alacaktır.

3.2.2 Parametre optimizasyonu

Bu işlem, bir parametrik ptSTL formülü ϕ için, tanımlanmış parametre uzayı P^ϕ içinde, verilen ölçütü en iyileştiren parametre değerlerinin bulunmasını amaçlar. Bu çalışmada, parametre optimizasyonu, parametre uzayının taranması, yani ızgara araması metodu ile yapılmıştır. Örneğin parametrik $\phi = F_{[0,p_1]}^- G_{[0,p_2]}^- x^0 > p_3$ formülü ve $p_1, p_2 \in \{1, 3, \dots, 2m+1\}, p_3 \in \{1, 2, \dots, n\}$ parametre uzayı için, $m^2 n$ adet parametre değeri bulunmaktadır. ızgara araması metodunda, her bir parametre değeri $p \in P^\phi$ için bir ptSTL formülü $\phi(p)$ tanımlanır ve bu formülün veri kümesi uyumsuzluk değeri $C(\phi(p), D)$ Denklem (5)'deki gibi hesaplanır. Bu hesap sonucunda, uyumsuzluk değerini en küçükleştiren parametre değeri p^* seçilir:

$$p^* = \underset{p \in P^\phi}{\operatorname{argmin}} C(\phi(p), D) \quad (6)$$

3.2.3 Formül Sentezi

Bir sinyal kümesini en iyi şekilde tanımlayan ptSTL formülünü üretmek için, öncelikle sinyal değişkenleri (M kümesi) kullanılarak oluşturulabilen parametrik formüller tanımlanmıştır. Ardından her bir formül için Denklem (6)'da belirtilen parametre optimizasyonu yapılmıştır. Temel olarak, operatör limiti c ve değişken kümesi M olarak verildiğinde, c tane operatör içeren bütün ptSTL formülleri F^c ile gösterilmiş ve özyinelemeli olarak Denklem (7)'deki gibi tanımlanmıştır. Buna dayalı olarak, en çok c operatör içeren bütün ptSTL formülleri ise $F^{\leq c}$ ile gösterilmiş ve Denklem (8)'deki gibi tanımlanmıştır (Aydın & Aydın Göl, 2020).

$$F^0 = \{x \sim p \mid x \in M, \sim \in \{<, >\}\} \quad (7)$$

$$F^c = \{\neg \phi \mid \phi \in F^{c-1}\} \cup \{X_{[a,b]} \phi \mid \phi \in F^{c-1}, X \in \{F^-, G^-\}\} \\ \bigcup_{i=1}^{c-1} \{\phi_1 \operatorname{op} \phi_2 \mid \phi_1 \in F^i, \phi_2 \in F^{c-i-1}, \operatorname{op} \in \{\wedge, \vee\}\} \\ \bigcup_{i=1}^{c-1} \{\phi_1 S_{[a,b]} \phi_2 \mid \phi_1 \in F^i, \phi_2 \in F^{c-i-1}\}$$

$$F^{\leq c} = \bigcup_{i=0}^c F^i \quad (8)$$

Formül sentezi aşamasında öncelikle bir operatör üst limiti (c) belirlenmiştir. Ardından, takip edilen metrikler üzerinde bu limitten fazla operatör içermeyen bütün parametrik ptSTL formülleri $F^{\leq c}$ tanımlanmıştır (Denklem (8)). Her bir parametrik formül $\phi \in F^{\leq c}$ için, farklı versiyonlardan üretilmiş olan D_A ve D_B sinyal kümeleri üzerinde parametre optimizasyonu yapılmıştır. Bunun sonucunda Denklem (9)'da verilen parametrik formül ϕ , D_A kümesine göre optimum parametre p_A^* , ve D_B kümesine göre optimum parametre p_B^* üçlülerinden oluşan bir küme tanımlanmıştır.

$$\mathcal{F}^P = \{(\phi, p_A^*, p_B^*) \mid p_A^* = \underset{p \in P^\phi}{\operatorname{argmin}} C(\phi(p), D_A), \quad p_B^* = \underset{p \in P^\phi}{\operatorname{argmin}} C(\phi(p), D_B), \quad \phi \in \mathcal{F}^{\leq c}\} \quad (9)$$

Denklem (9)'da verilen kümenin oluşturulması için, $F^{\leq c}$ kümesindeki her bir parametrik formül için Denklem (6)'da verilen ızgara araması işlemi iki kere gerçekleştirilmiştir. $F^{\leq c}$ kümesinin boyutu takip edilen metrik sayısı (M kümesinin boyutu) ve operatör üst limiti (c) ile üstel olarak artmaktadır. Izgara aramasında kullanılan parametre değeri sayısı (P^ϕ kümesinin boyutu), formüldeki parametre sayısı ile üstel olarak artmaktadır. Bir formülde bulunan parametre sayısı ise formülün operatör sayısı ile doğru orantılıdır. Bu değerlendirme ışığında, metrik sayısının ve operatör üst limitinin arttırılmasının hesap süresini ciddi ölçüde arttıracığı görülmektedir. Bu süre köşegen taraması (Ergurtuna & Aydın Gol, 2019) veya ızgara aramasında bilinen en iyi sonuca göre aramanın erken sonlandırılması gibi keşifsel metotlar ile kısaltılabilir.

3.2 Sürüm Değerlendirme Kararı

Sunulan formel metotlara dayalı sürüm değerlendirme sürecinin son aşaması, A ve B versiyonlarını tanımlayacak şekilde üretilen formüllerin ve parametrelerinin karşılaştırılarak sürüm değerlendirme sonucunun üretilmesidir.

Öncelikle, her iki versiyonu da en iyi şekilde tanımlayan bir parametrik formül ile bu formülün D_A ve D_B sinyal kümelerine göre optimum parametreleri Denklem (10)'da verilen kritere göre seçilir.

$$(\phi^*, p_A^*, p_B^*) = \underset{(\phi, p_A^*, p_B^*) \in \mathcal{F}^P}{\operatorname{argmin}} \max(C(\phi(p_A^*), D_A), C(\phi(p_B^*), D_B)) \quad (10)$$

Seçilen parametrik formül ve parametrelerden elde edilen formüller $\phi^*(p_A^*)$ ve $\phi^*(p_B^*)$, sırası ile D_A ve D_B sinyal kümelerini yüksek başarımla tanımlamaktadır. Bu formüllerin yapısı aynıdır fakat eşitsizlik sabitleri veya zamansal parametreler farklılık gösterebilmektedir. Burada ortaya çıkan farklılıklara göre ise sürüm değerlendirme sonucu üretilmektedir. Temel olarak ϕ^* formülünde geçen her bir parametre p_i için, bu parametrenin optimize edilmiş değerleri $p_A^*(p_i)$ ve $p_B^*(p_i)$ karşılaştırılır. Karşılaştırma sırasında zamansal ve metrik parametreler için değerlendirme yapılan sisteme göre önceden belirlenmiş olan sınırlar kullanılabilir. Bunun yanı sıra, sınırların otomatik olarak belirlenmesi gerektiği durumlarda ise, birbirlerine göre bağıl olarak kontrol edilebilirler. Örneğin, B versiyonu için bulunan parametrenin, A versiyonu için bulunan parametrenin yüzde 10 yakınında olması gibi, $p_B^*(p_i) \in [0.9p_A^*(p_i), 1.1p_A^*(p_i)]$. Parametrelerden herhangi biri belirlenen sınırların dışında bir değer almış ise sürüm değerlendirme sonucu olumsuz, bütün parametrelerin belirlenen limitler içinde değer alması durumunda ise sürüm değerlendirme sonucu olumlu olarak üretilir. Bu yaklaşım p_1, \dots, p_n parametrelerini içeren ϕ^* formülü için Denklem (11)'de özetlenmiştir.

$$\text{OLUMLU: Eğer her bir } i = 1, \dots, n \text{ için} \quad p_B^*(p_i) \in [(1-r)p_A^*(p_i), (1+r)p_A^*(p_i)] \quad (11)$$

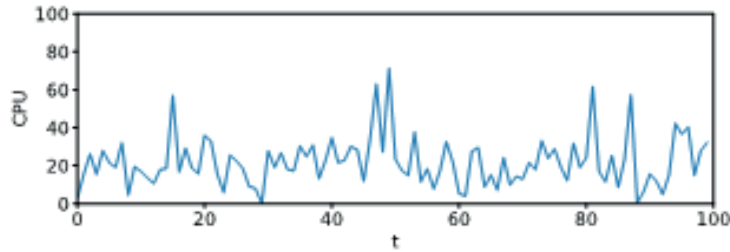
$$\text{OLUMSUZ: Eğer en azından bir } i = 1, \dots, n \text{ için} \quad p_B^*(p_i) \notin [(1-r)p_A^*(p_i), (1+r)p_A^*(p_i)]$$

Sürüm değerlendirme sonucunun olumsuz olması durumuna, $\phi^*(p_A^*)$ ve $\phi^*(p_B^*)$ formülleri ile Denklem (11)'de verilen koşulu sağlamayan parametreler kullanıcıya verilerek değerlendirmeye yönelik anlaşılır bir analiz sunulur.

4. ÖRNEK SİSTEM ÜZERİNDE SONUÇLAR

Bu bölümde geliştirilen sürüm değerlendirme sistemi merkezi işlemci birimi (central processing unit) kullanım oranına dayalı sürüm değerlendirme yapılması amacı ile kullanılmıştır. Çok çekirdekli CPU'larda, her çekirdekteki kullanım oranı ayrı bir metrik olarak tanımlanarak çok boyutlu bir sinyal üzerinde değerlendirme yapılabilir. Bu çalışmada ise sadece bir metrik, birleştirilmiş CPU kullanım oranı, kullanılmıştır. Belirlenen metrik üzerine iki deney yapılmıştır. Deneylerde

kullanılan veri kümelerindeki sinyaller yapay olarak Gauss dağılımlarından örneklemeler alınarak oluşturulmuştur. Tüm sinyaller 100 birim uzunluğundadır. Bir sinyal örneği Şekil 2’de görülebilir. Deneylede kullanılan sinyal kümelerinde 50 adet sinyal bulunmaktadır. Bütün uyumsuzluk hesaplarında, değerlendirmenin başladığı noktayı gösteren k sabiti (bkz. Denklem (4)) 20 olarak alınmıştır. Yani her bir ptSTL formülü için $(100-20) \times 50 = 4000$ adet nicel değerlendirme hesabı yapılmıştır.



Şekil 2. CPU Sinyal kümesinden örnek bir sinyal.

Ortalaması m sapması s olan bir Gauss dağılımı $G(m,s)$ ile gösterilmiştir. Bir veri kümesi üretilirken iki adet Gauss dağılımı kullanılmıştır: normal durum için $G(m^n, s^n)$ ve nadiren oluşan (gürültülü) durumu için $G(m^g, s^g)$. Buna ek olarak bir gürültü oranı parametresi $g \in [0,1)$ kullanılmıştır. Sinyal üretilirken, her bir zaman birimi için $1-g$ olasılık ile $G(m^n, s^n)$ dağılımından ve g olasılık ile $G(m^g, s^g)$ dağılımından örneklem alınmıştır. Ek olarak, CPU kullanım oranını modelleyebilmek için elde edilen veriler $[0,100]$ arasında tutulmuştur.

Yapılan deneylerde operatör üst limiti 1 olarak belirlenmiştir. Bir metrik için, toplam 8 parametrik ptSTL formülü tanımlanmıştır (Denklem (8)). Izgara araması için zamansal operatörlerin parametrelerin alt ve üst limitleri $\{0,5,10,15,20\}$ kümesinden, sinyal değişkeninin (CPU) karşılaştırma eşik değerleri ise $\{0,5,\dots,30\}$ kümesinden alınmıştır.

Yapılan ilk deneyde A ve B sürümleri için sinyal kümeleri D_A ve D_B aynı profil ile oluşturulmuştur. Her bir sinyal kümesi için $G(20,10)$ dağılımı normal durumlar için ve $G(60,10)$ dağılımı gürültülü durumlar için kullanılmıştır. Gürültü oranı parametresi ise $g=0.02$ olarak belirlenmiştir. Şekil 2’de D^4 kümesinden bir sinyal gösterilmiştir. Geliştirilen sürüm değerlendirme metodunun uygulanması sonucunda, her iki veri kümesini birden en iyi şekilde tanımlayabilen parametrik formül $\phi^* = \mathbf{G}^-_{[p_1, p_2]} x > p_3$ olmuştur. Buna ek olarak, iki veri kümesi içinde parametre optimizasyonu sonucunda aynı parametrelere $p_A^* = p_B^* = [0,15,5]$ ulaşılmıştır. Bu parametreler ile, veri kümeleri üzerinde elde edilen uyumsuzluk değerleri $C(\phi^*(p_A^*), D_A) = 3.6907$ ve $C(\phi^*(p_A^*), D_B) = 3.6917$ olmuştur. Buradaki farklılık, veri kümelerinin rastgele üretilmiş olmasından kaynaklanmaktadır. Bu değerler ile nicel değerlendirme değerinin ($|\rho(\phi, x, i)|$) her iki veri kümesi için de ortalama 3.7 olduğu görülmektedir. Şekil 2’de örneklendirilen $[0,100]$ aralığında değer alabilen bir sinyal kullanıldığı da hesaba katıldığında, elde edilen formüllerin veri kümelerini sıkı bir şekilde tanımladığı görülmektedir. Elde edilen formüller tamamen aynı olduğu için, Denklem (11)’de özetlenen sürüm değerlendirme sonucu olumludur.

İkinci deneyde ise A sürümü normal durumu $G(20,10)$ dağılımı ile B sürümü normal durumu ise $G(25,10)$ dağılımı ile üretilmiştir. Her iki sürüm için de $G(60,10)$ dağılımı ile gürültülü durumlar üretilmiş ve gürültü oranı parametresi $g = 0.02$ olmuştur. Yani, B sürümü için ortalama CPU kullanım oranı beş birim artarak 20’den 25’e çıkmıştır. Yeni sürümde bu artış tetikleyecek bir değişiklik yapılmamış ise bu beklenmeyen bir durum olarak görülür ve sürüm değerlendirme sisteminin bu konuda bir uyarı üretmesi beklenir.

Geliştirilen sürüm değerlendirme metodu uygulandığında, ilk deneyde olduğu gibi $\phi^* = \mathbf{G}^-_{[p_1, p_2]} x > p_3$ parametrik ptSTL formülü, her iki sinyal kümesini de en iyi şekilde tanımlayabilen formül olmuştur. Fakat bu deneyde, parametre optimizasyonu sonucunda $p_A^* = [0,15,5]$ ve $p_B^* = [0,15,10]$ parametrelerine ulaşılmıştır. Parametrelerden bir tanesi (p_3), B sürümünde A sürümüne oran ile iki katına çıkmıştır. Bu nedenle Denklem (11)’de özetlenen sürüm değerlendirme sonucu olumsuzdur. Olumsuz değerlendirme için kullanıcıya verilen $\phi^{A,*} = \mathbf{G}^-_{[0,15]} x > 5$ ve $\phi^{B,*} = \mathbf{G}^-_{[0,15]} x > 10$ formülleri de CPU kullanım oranında ciddi bir artış olduğunu net olarak göstermektedir.

Deneylerde kullanılan sinyal kümeleri üzerinde regresyon analizine dayalı bir sürüm değerlendirme metodu da uygulanmıştır. Bu metotta, öncelikle benzetimli tavlama algoritması kullanılarak D_A kümesindeki sinyalleri tanımlayabilecek 3. dereceden bir sinyal üretilmiştir. Üretilen tanımlayıcı sinyalin sinyal kümesine uyumluluğunu ölçmek için ortalama artık kareler toplamı kullanılmıştır (r_A). Ardından, üretilen bu sinyalin D_B kümesi üzerindeki ortalama artık kareler toplamı r_B hesaplanmıştır. Bu metotta sürüm değerlendirme sonucu r_A ve r_B arasındaki farka göre üretilir. İlk deneyde r_A ve r_B değerleri, sırası ile, 124 ve 125 olarak hesaplanmıştır. İkinci deneyde ise $r_A = 129$ ve $r_B = 176$ olarak hesaplanmıştır. İlk deneyde birbirine yakın değerler gözlenirken, ikinci deneyde r_B 'de görülen artış olumsuz bir sürüm değerlendirme sonucu üretilmesi gerektiğini göstermektedir. Geliştirilen metoda göre özetlenen regresyon analizi tabanlı sürüm değerlendirme yaklaşımı, kullanıcıya üretilen sonuç ile ilgili açıklayıcı bir bilgi sunamamaktadır. Bunun yanı sıra, farklı karakterlerdeki sinyaller için tanımlayıcı sinyalin derecesi (3.) ve artık kareler toplamları arasındaki farklar için sınır belirlenmelidir.

5. SONUÇ

Bu çalışmada formel metotlara dayalı bir sürüm değerlendirme metodu sunulmuştur. Önerilen sürüm değerlendirme metodu, takip edilen metrikler ile ilgili yazılabilecek bütün parametrik ptSTL formüllerinin sıralanması ve her bir parametrik formül için A ve B versiyonlarından üretilmiş olan sinyal kümeleri üzerinde parametre optimizasyonu yapılması adımlarını içermektedir. Son olarak da iki versiyonu da en iyi şekilde tanımlayabilen parametrik formül için üretilen parametre değerleri karşılaştırılarak sürüm değerlendirme sonucu üretilmiştir. Önerilen metot ile sürüm değerlendirmede kullanılmak üzere ptSTL formüllerinin otomatik olarak üretilmesi sağlanmıştır.

Hakem Değerlendirmesi: Dış bağımsız.

Çıkar Çatışması: Yazar çıkar çatışması bildirmemiştir.

Finansal Destek: Bu çalışma, TÜBİTAK tarafından desteklenmiştir. (Proje No: 117E242)

Peer-review: Externally peer-reviewed.

Conflict of Interest: The author has no conflict of interest to declare.

Grant Support: This study was supported by TÜBİTAK. (Project No: 117E242)

Kaynaklar/References

- Asarin E., Donzé A., Maler O. & Nickovic D. (2012). Parametric Identification of Temporal Properties. In: Khurshid S., Sen K. (eds). *Lecture Notes in Computer Science, vol 7186: Proceedings of Runtime Verification* (pp. 147-160). Springer, Berlin, Heidelberg. https://doi.org/10.1007/978-3-642-29860-8_12.
- Aydin, S. K. & Aydin Gol, E. (2020). Synthesis of monitoring rules with STL. *Journal of Circuits, Systems, and Computers*, 29(11), 1-26. <https://doi.org/10.1142/S0218126620501777>.
- Bartocci E., Bortolussi L. & Sanguinetti G. (2014) Data-Driven statistical learning of temporal logic properties. In: Legay A., Bozga M. (eds). *Lecture Notes in Computer Science, vol 8711: Proceedings of Formal Modeling and Analysis of Timed Systems* (pp 23-37). Springer, Cham. https://doi.org/10.1007/978-3-319-10512-3_3.
- Bays, M. E. (1999). *Software Release Methodology*. USA: Prentice-Hall.
- Bombara, G., Vasile, C.-I., Penedo, F., Yasuoka, H. & Belta, C. (2016). A decision tree approach to data classification using signal temporal logic. *Proceedings of the Hybrid Systems: Computation and Control*, 1–10. <https://doi.org/10.1145/2883817.2883843>.
- Chatterjee, S. & Simonoff, J. S. (2013). *Handbook of Regression Analysis*. Wiley.
- Donzé A. (2013). On signal temporal logic. In: Legay A., Bensalem S. (eds). *Lecture Notes in Computer Science, vol 8174: Proceedings of Runtime Verification* (pp 382-383). Berlin: Springer. https://doi.org/10.1007/978-3-642-40787-1_27.
- Ergurtuna, M. & Aydin Gol E. (2019). An efficient formula synthesis method with past signal temporal logic. *Proceedings of the IFAC Conference on Intelligent Control and Automation Sciences (ICONS)*, 43-48. <https://doi.org/10.1016/j.ifacol.2019.09.116>.
- Gabbay D. (1989) The declarative past and imperative future. In: Banieqbal B., Barringer H., Pnueli A. (eds) Temporal Logic in Specification. Lecture Notes in Computer Science, vol 398 (pp. 409-448). Springer, Berlin, Heidelberg. https://doi.org/10.1007/3-540-51803-7_36.
- Howard, D. (2016). *IT Release Management*. O'Reilly.
- Hoxha, B., Dokhanchi, A. & Fainekos, G. (2018). Mining parametric temporal logic properties in model-based design for cyber-physical systems. *Int J Softw Tools Technol Transfer* 20, 79–93. <https://doi.org/10.1007/s10009-017-0447-4>.
- Jha, S., Tiwari, A., Seshia, S.A., Sahai, T. & Shankar, N. (2019). TeLEx: learning signal temporal logic from positive examples using tightness metric. *Formal Methods in System Design* 54. 364–387. <https://doi.org/10.1007/s10703-019-00332-1>.

-
- Jin, X., Donzé, A., Deshmukh, J. V. & Seshia, S. A. (2015). Mining requirements from closed-loop control models. *IEEE Transactions on Computer-Aided Design of Integrated Circuits and Systems*, 34, 11. 1704-1717. <https://doi.org/10.1109/TCAD.2015.2421907>.
- Ketenci, A. & Aydin Gol, E. (2019). Synthesis of monitoring rules via data mining. *Proceedings of the IEEE American Control Conference (ACC)*, 1684-1689. <https://doi.org/10.23919/ACC.2019.8815002>.
- Kong, Z., Jones, A., Ayala, A. M., Aydin Gol, E. & Belta, C. (2014). Temporal logic inference for classification and prediction from data. *Proceedings of the Hybrid Systems: Computation and Control*, 273-282. <https://doi.org/10.1145/2562059.2562146>.
- Mohammadinejad, S., Deshmukh, J. V., Puranic, A. G., Vazquez-Chanlatte, M. & Donzé, A. (2020). Interpretable classification of time-series data using efficient enumerative techniques. *Proceedings of the Hybrid Systems: Computation and Control*, 1-10. <https://doi.org/10.1145/3365365.3382218>.
- Sommerville, I. (2015). *Software Engineering (10th ed.)*. Pearson.
- Yoo, C & Belta, C. (2017). Rich time series classification using temporal logic. *Proceedings of Robotic: Science and Systems*. 1-9.

Performance Comparison of Different Pre-Trained Deep Learning Models in Classifying Brain MRI Images

Beyin MR Görüntülerini Sınıflandırmada Farklı Önceden Eğitilmiş Derin Öğrenme Modellerinin Performans Karşılaştırması

Onur Sevli¹ 



ABSTRACT

A brain tumor is a collection of abnormal cells formed as a result of uncontrolled cell division. If tumors are not diagnosed in a timely and accurate manner, they can cause fatal consequences. One of the commonly used techniques to detect brain tumors is magnetic resonance imaging (MRI). MRI provides easy detection of abnormalities in the brain with its high resolution. MR images have traditionally been studied and interpreted by radiologists. However, with the development of technology, it becomes more difficult to interpret large amounts of data produced in reasonable periods. Therefore, the development of computerized semi-automatic or automatic methods has become an important research topic. Machine learning methods that can predict by learning from data are widely used in this field. However, the extraction of image features requires special engineering in the machine learning process. Deep learning, a sub-branch of machine learning, allows us to automatically discover the complex hierarchy in the data and eliminates the limitations of machine learning. Transfer learning is to transfer the knowledge of a pre-trained neural network to a similar model in case of limited training data or the goal of reducing the workload. In this study, the performance of the pre-trained Vgg-16, ResNet50, Inception v3 models in classifying 253 brain MR images were evaluated. The Vgg-16 model showed the highest success with 94.42% accuracy, 83.86% recall, 100% precision and 91.22% F1 score. This was followed by the ResNet50 model with an accuracy of 82.49%. The findings obtained in this study were compared with similar studies in the literature and it was found that it showed higher success than most studies.

Keywords: Brain MRI Classification, Transfer Learning, Convolutional Neural Networks

ÖZ

Beyin tümörleri, beyin hücrelerinin kontrolsüz bölünmeleri sonucu meydana gelen kitlelerdir. Tümörler zamanında ve doğru teşhis edilmezse ölümcül sonuçlara neden olabilir. Beyin tümörlerini tespit etmede yaygın olarak kullanılan tekniklerden biri olan MRI'dir. MRI, sağladığı yüksek çözünürlük ile beyindeki anormalliklerin kolay tespitine imkân verir. MR görüntüleri geleneksel olarak radyologlar tarafından incelenip yorumlanır. Ancak teknolojinin gelişmesi ile birlikte üretilen çok miktarda veriyi makul sürelerde yorumlamak daha zor hale gelmektedir. Bu nedenle bilgisayarlı yarı otomatik ya da otomatik yöntemlerin geliştirilmesi önemli bir araştırma konusu haline gelmiştir. Verilerden öğrenerek tahmin yapabilen makine öğrenmesi yöntemleri bu alanda yaygın olarak kullanılmaktadır. Ancak makine öğrenmesi için görüntü özelliklerinin çıkarımı özel bir mühendislik gerektirir. Makine öğrenmesinin bir alt dalı olan derin öğrenme, veri içerisindeki karmaşık hiyerarşiyi otomatik olarak keşfetmeye imkân sağlar ve makine öğrenmesinin sınırlılıklarını ortadan kaldırır. Transfer öğrenme ise eldeki eğitim verisinin az olması halinde ya da iş yükünü azaltmak için daha önceden eğitilmiş bir derin sinir ağının bilgilerinin benzer başka bir modele aktarılmasıdır. Bu çalışmada önceden eğitilmiş Vgg-16, ResNet50 ve Inception v3 modellerinin sınıflamadaki performansları değerlendirilmiştir. Vgg-16 modeli %94.42 doğruluk, %83.86 recall, %100 precision ve %91.22 F1 skoru ile en yüksek başarıyı göstermiştir. Bunu %82.49 doğrulukla ResNet50 modeli izlemektedir. Bu çalışmada elde edilen bulgular literatürdeki benzer çalışmalarla karşılaştırılmış ve çoğu çalışmadan daha yüksek başarı gösterdiği görülmüştür.

Anahtar kelimeler: Beyin MRI Sınıflama, Transfer Öğrenme, Evrimsel Sinir Ağları

¹(Asst. Prof. Dr.), Mehmet Akif Ersoy University, Engineering and Architecture Faculty, Department of Computer Engineering, Burdur, Turkey

ORCID: O.S. 0000-0002-8933-8395

Corresponding author:

Onur SEVLİ

Mehmet Akif Ersoy University, Engineering and Architecture Faculty, Department of Computer Engineering, Burdur, Turkey

E-mail address: onursevli@mehmetakif.edu.tr

Submitted: 15.02.2021

Revision Requested: 02.04.2021

Last Revision Received: 24.04.2021

Accepted: 05.05.2021

Published Online: 29.06.2021

Citation: Sevli, O. (2021). Performance comparison of different pre-trained deep learning models in classifying brain MRI images. *Acta Infologica*, 5(1), 141-154. <https://doi.org/10.26650/acin.880918>

1. INTRODUCTION

The brain is one of the most complex organs with billions of cells. A brain tumor is an abnormal group of cells that results from uncontrolled cell division in the brain. If precautions are not taken in time, this abnormal group of cells leads to impaired brain activity and damage to healthy cells. Brain tumors, like all other types of cancer, are a serious problem that threaten human life and can be fatal if not diagnosed and treated in time. 24,000 new brain cancer cases were diagnosed in the United States in 2019 (Cancer.Net, 2020). Accurate and rapid diagnosis of brain tumors is important for the effective treatment of the disease. Computer Aided Diagnosis (CAD) assists neuro-oncologists in a variety of ways, and therefore CAD-based brain tumor detection is a popular research topic (Kumar, Dabas, and Godara 2017). Diagnostic systems using machine learning and deep learning are examples of CAD. CAD systems are based on magnetic resonance (MR) images.

Magnetic resonance imaging (MRI) is one of several techniques used in medical imaging and is widely used in imaging abnormal brain tissues. Among other techniques, MRI is the most popular and risk-free one. MRI provides higher contrast for soft tissues of the brain than CT images. MRI provides valuable information about the brain with high resolution. With the help of MRI, abnormalities in the brain can be detected easily.

Advances in the field of medical imaging make it possible to extensively investigate human organs using a large number of images produced. Traditionally, MR images are examined by radiologists, and abnormalities in the brain are detected. However, with advances in medical imaging techniques and hardware acceleration, manual interpretation of large amounts of data produced by MRI becomes a laborious task. For this reason, semi-automatic or automatic analysis of images with computer-aided systems has become an important research topic (Bauer et al. 2013). As the issue is directly related to human health, high accuracy and clarity are needed. At this point, computer-aided diagnostics help to obtain accurate results faster.

Brain tumor segmentation aims to detect the location and spread of tumors, active tumor tissue, and edema (Rajinikanth et al. 2017). Segmentation and quantitative lesion assessment are vital for treatment planning methods, disease monitoring, and control of disease progression. This is done by comparing abnormal tissues with normal ones. Brain tumor segmentation from MR images is of great importance for advanced diagnosis and treatment planning. Current automated and semi-automated brain tumor segmentation methods are generally categorized as either generative or discriminative model-based methods (Menze et al. 2014). Generative models require knowledge of probabilistic image atlases of both healthy and tumor tissues. Brain tumor segmentation based on probabilistic image atlases is an outlier detection problem (Prastawa et al. 2004). Discriminative models require less information about the anatomy of the brain and instead use the basic features of the image. These basic features are pixels, textures, local histograms (Kleesiek et al. 2014), region shape difference and symmetry analysis (Tustison et al. 2013). The discriminative model methods classify MR images as either tumors or normal tissue according to their characteristics, and the performance of the method depends on the image properties and the classification algorithms used. In the discriminative model, machine learning algorithms are used, which handle handcrafted features. Machine learning is a set of algorithms that allow computers to make predictions over big data. However, the extraction of features that enable classification in these techniques requires special expertise (Akkus et al. 2017).

Different MR image classification techniques have been used in the literature to address brain abnormalities. In these studies, the techniques which are typically used in the identification of normal and abnormal images are preprocessing, feature extraction, and classification steps. These studies include various methods of supervised machine learning such as wavelet transform (El-Dahshan, Hosny, and Salem 2010), Independent Component Analysis (ICA) (Moritz et al. 2000), Support Vector Machines (SVM) (Chaplot, Patnaik, and Jagannathan 2006), Random Forest Classifier (RFC) (Reza and Iftekharuddin 2014) and productive Gaussian Mixture Model (GMM) (Domingues et al. 2018). Recent studies use advanced machine learning techniques. Nayak et al. used Discrete Wavelet Transform (DWT) for feature extraction as well as the PCA for feature reduction and the AdaBoost algorithm with the RFC (Nayak, Dash, and Majhi 2016). In another study, the same researchers increased success by developing a model that uses the SVM with the AdaBoost (Ranjan Nayak, Dash, and Majhi 2017). Mohsen et al. used the properties obtained by the DWT and the PCA technique on segmented brain MR images to train the neural network used for the classification of brain tumors (Mohsen et al. 2018). Zhang et al. developed a brain MRI

classifier by combining the SVM and Particle Swarm Optimization (PSO) (Zhang et al. 2013). The same researchers used the wavelet entropy method instead of SVM in another study (Zhang et al. 2015). Wang et al. used PSO, artificial bee colony and feed-forward neural networks in their studies (Wang et al. 2015). Saritha et al. used probabilistic neural networks (PNN) with wavelet entropy method (Saritha, Joseph, and Mathew 2013).

In most automated brain tumor segmentation studies, the detection process is carried out with the help of machine learning classifiers using handcrafted features. However, traditional machine learning algorithms cannot make very good generalizations. A machine learning technique called deep learning which has attracted great attention in recent years eliminates the limitations of classical machine learning algorithms and enables automatic discovery of the features of MR images with its self-learning ability. In deep learning, features are automatically extracted by various feature processing layers, and feature engineering is less needed (LeCun, Bengio, and Hinton 2015). Deep learning techniques provide solutions to a wide range of problems.

Recent studies on computer-aided medical diagnosis show improved performances with the emergence of the concept of deep learning, and the success of deep learning in feature extraction is expanding the research field (Cao et al. 2018). Deep learning models were used in many biomedical applications such as diagnosis of lung cancer (Gu et al. 2018), breast cancer (Yousefi, Krzyżak, and Suen 2018), skin cancer (Zuo et al. 2017), diagnosis of pulmonary nodules (Cheng et al. 2016) and histopathological diagnosis (Litjens et al. 2016).

Recently, deep learning techniques have been used in the analysis of brain MRI images. Despite effortful studies on medical image analysis, automated detection of brain abnormalities remains an unsolved problem due to variations in brain morphology, imaging techniques and devices (Akkus et al. 2017). Studies mention that convolutional neural networks (CNN) perform better in deep learning-based brain tumor segmentation methods (Kamnitsas et al. 2017). The advantage of CNN-based systems is that they do not require manual segmentation and can perform fully automatic classification. In their study, Pashaei et al. automatically extracted the features from brain MR images using a CNN architecture and the model achieved 81% classification accuracy (Pashaei, Sajedi, and Jazayeri 2018). Charron et al. used a CNN-based model to track metastases (Charron et al. 2018). Havaei et al. proposed a model that simultaneously detects both local and global contextual features of brain tumors by trying a series of CNN architectures (Havaei et al. 2017). Agn et al. proposed a deep learning-based model that adjusts the parameters in brain tumor radiotherapy planning to minimize the risk for healthy tissues (Agn et al. 2019). Afshar et al. used a capsule neural network, a modified CNN architecture, to classify brain tumors, but the improvement in performance was not sufficient (Afshar, Mohammadi, and Plataniotis 2018).

Examining MR images requires large amounts of data processing. For this data processing task, using the graphics processing units (GPU) of computers provides higher performance than using the central processing unit (CPU). Another solution for increasing performance in deep learning is using transfer learning method. Transfer learning is simply a matter of transferring the knowledge of another pre-trained neural network to a new and similar model. Transfer learning has great potential for computer-aided detection of medical problems. Zhou et al. used a pre-trained Inception v3 model to detect kidney tumors (Zhou et al. 2019). Deniz et al. used VGG-16 and AlexNet model for breast cancer diagnosis and then they classified tumors using SVM (Deniz et al. 2018). Hussein et al. proposed a transfer learning-based model for lung and pancreatic tumor characterization (Hussein et al. 2019). Transfer learning also attracts attention in neurooncological studies. Kaur and Gandhi conducted a study that showed that transfer learning is more effective in brain tumor classification than traditional machine learning (Kaur and Gandhi 2020). Yang et al. used AlexNet and GoogLeNet in their study to determine the tumor grade with brain MR images (Yang et al. 2018). Jain et al. used a pre-trained VGG-16 network to diagnose Alzheimer's disease via MR images (Jain et al. 2019). Deepak and Ameer used pre-trained GoogLeNet to distinguish three different types of tumors via MR images (Deepak and Ameer 2019). Rehman et al. performed a study using GoogleNet, AlexNet and Vgg models for a similar task (Rehman et al. 2020). Chelghoum et al. conducted a study demonstrating the effectiveness of transfer learning with limited time and small number of epochs in brain tumor classification (Chelghoum et al. 2020).

In this study, three different pre-trained models were used to classify 253 brain MR images. The performances of the Vgg-16, ResNet50 and Inception V3 models were evaluated in terms of accuracy, recall, sensitivity and F1-score measurements. With a relatively limited dataset, the aim was to achieve high accuracy values in low epoch numbers, and a success of over 94%

was achieved as a result of 30 epochs. The results obtained in this study were compared with similar recent studies in the literature. It was revealed that the success achieved with the Vgg-16 model was higher than similar studies in the literature. In this study, an alternative and successful solution for tumor detection on brain MRI images was presented.

In the following sections, transfer learning models used in the study, dataset, data processing method, deep learning techniques used in the study, experiments and findings are presented .

2. MATERIAL AND METHOD

Most previous studies are based on manual extraction of tumor features before classification. This prevents them from being fully automated. Only a small number of studies show solutions being produced if the available data are relatively small.

In this study, a fully automated classification approach for brain MRI images was presented, using the transfer learning method in deep convolutional neural networks. A dataset consisting of 253 brain MRI images was used. Before the learning process in deep neural networks, the images were preprocessed. The basic steps of the process are as follows:

1. Preprocessing: Raw MRI images were cropped and the borders of the brain tissue were determined with the help of the OpenCV library.
2. Deep learning model implementations: Three different pre-trained models (Vgg-16, ResNet50 and Inception V3) were used.
3. Data augmentation: Data augmentation technique was carried out during the training to increase learning success and prevent overfitting.
4. Learning-rate optimization: The learning rates of the models were optimized using the adaptive momentum method.
5. The performances of the models were evaluated with accuracy, precision, recall, and F1 score metrics.

The workflow of the method used is given in Fig. 1.

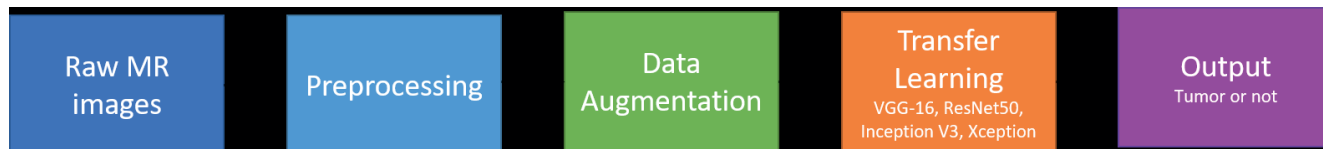


Figure 1. Flowchart of the study

2.1. Dataset

The dataset used in the study consists of 253 brain MRI images collected and shared publicly by Chakrabarty at the Jalpaiguri Government College of India (Chakrabarty, N. n.d.). 155 of these images were with tumor and 98 of them were without tumor. The graph showing the tumor and non-tumor image distributions within the dataset is given in Figure 2.

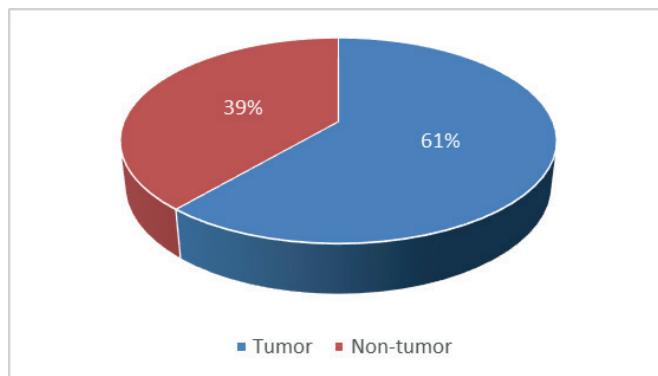


Figure 2. Image distributions within the dataset

The dataset is imbalanced in terms of the class distribution of the images. As a result of the dominance of a particular class in the dataset, the classifier becomes more prone to learning the dominant class. To eliminate this instability problem, the number of instances of the dominant class should be reduced or the instances of the less data class should be increased. Since the dataset used in this study was relatively limited, reducing the amount of data of the dominant class decreases the generalizability of the classification success. For this reason, data augmentation was applied to the class with a small number of samples.

Examples of MR images with tumors and without tumors in the dataset are given in Fig. 3. The dimensions of the images in the dataset and the black frames around them are different. For this reason, preprocesses were applied to standardize the images before the training.

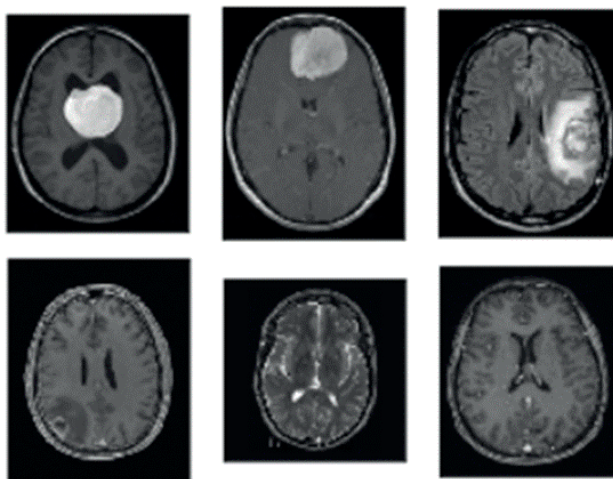


Figure 3. Sample images from the dataset

2.2. Image Preprocessing and Data Augmentation

During the preprocessing process, using the OpenCV library with Python, the boundaries of the brain tissue over the frame located around the image, namely, the endpoints in four separate poles, right, left, top, and bottom, were identified (Fig. 4).

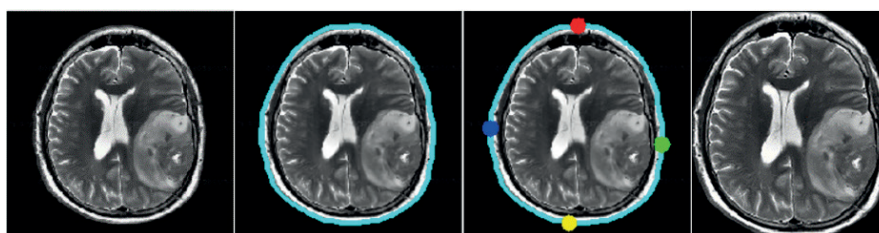


Figure 4. Preprocessing stages

The image was cropped according to the boundaries and polar points. In this way, the boundaries of the brain tissue were determined on the raw image and the part outside was discarded. This makes data processing easier. Images with different width and length values were resized to 224x224 pixel size after cropping.

Working with a relatively small amount of data and an imbalanced dataset leads to overfitting. The model tends to learn the dominant class and cannot generalize for other situations. Data augmentation technique is used to prevent this problem and to diversify the data. The data augmentation technique enables the generation of new data by applying processes such as zooming, reflecting, and rotating at varying rates on the existing dataset. The transformation methods and rates applied to the original data are obtained empirically, depending on the dataset and process. In this study, augmentation over the current dataset was done using 15% rotation, 5% shifting on the horizontal and vertical axis, 1/255 rescaling, and mirroring on the

horizontal and vertical axis. The rotation, shifting, and scaling ratios used for this study were also obtained experimentally in a way that maximizes the performance increase for the model.

Data augmentation was applied to reproduce tumor-free images with fewer samples to balance the dataset used. While the number of images with tumor remained the same, the number of images without tumor was increased by 50%. As a result of data augmentation, the number of images belonging to each class is given in Table 1.

Table 1

Number of images after data augmentation

Image class	Image count
Images with tumor	155
Images without tumor	147
Total	302

2.3. Deep Learning, CNN and Transfer Learning

Deep learning refers to the learning process performed with multiple layer neural networks to extract a feature hierarchy from raw input data. Contrary to manual feature extraction in classic machine learning, it is a popular technique that automatically extracts features from images. With the advances in hardware technology, large amounts of data can be processed quickly and high success and generalization can be achieved.

There are various deep learning techniques used for different purposes. CNN is a deep learning technique widely used in image classification and segmentation. CNN has also been widely used in medical image analysis to achieve better results (Belaïd and Loudini 2020). CNN enables automatic extraction and definition of features through images. A typical CNN includes convolution, pooling, activation and classification layers. Images that are classified with specific tags are transferred to the CNN in pixels so that the trainable parameters in the network are optimized to increase the classification accuracy. The convolution layer applies a kernel onto the input pixels, thereby revealing the features. The pooling layer reduces the data by taking the largest or average of the data from previous layers. The fully connected layer enables the transition to the prediction phase. The general structure of a CNN architecture is shown in Fig. 5.

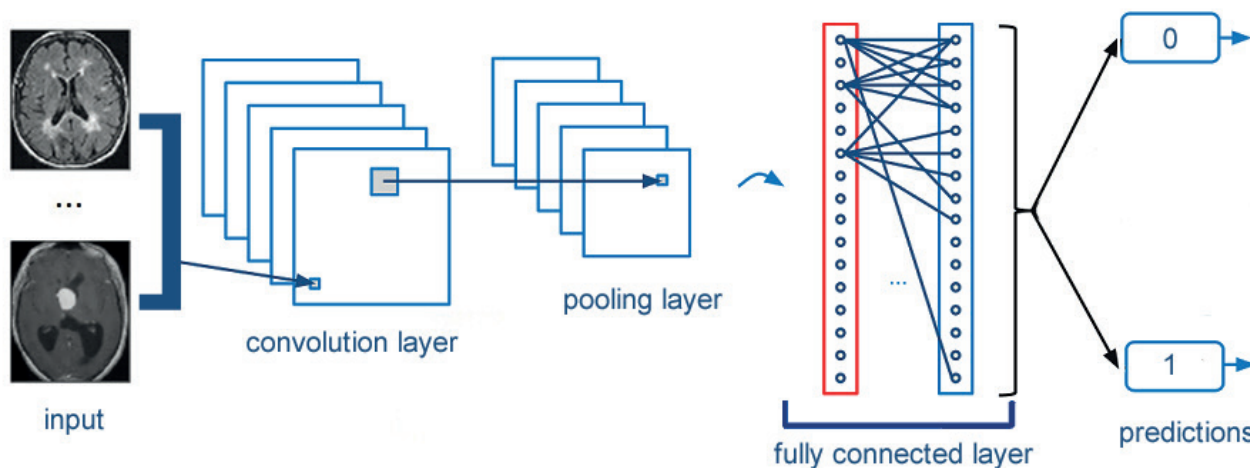


Figure 5. CNN architecture

Transfer learning is to transfer the weights of a network previously trained with large amounts of data to another model created to solve a similar problem. This method becomes important if there is not enough training data for the current problem. With a relatively small amount of data, an overfitting problem may occur in the network and successful generalizations cannot be made. If the data on which the pre-trained model was trained is comprehensive enough, the transferred network parameters ensure the correct classification of a small amount of data. The calculated weights of the pre-trained model are transferred to the new model and only the classifiers in the last part of the new model are trained.

Studies on medical image analysis also reveal that transfer learning is useful when a limited number of images are available (Deepak and Ameer 2020). In this study, a brain MRI classification task was performed using pre-trained VGG-16, ResNet50, Inception v3 transfer learning models, and then the success of the models was compared.

2.3.1. VGG-16 model

VGG network architecture is a CNN model introduced in 2014 by Simonyan and Zisserman (Simonyan and Zisserman 2020). VGG-16 is a special VGG type with 16 weighted layers. Its layers are convolution, maxpooling, activation, and fully connected layer. The general structure of the VGG-16 model is given in Fig. 6.



Figure 6. Structure of VGG-16 model

There are a total of 21 layers in the architecture: 13 convolution, 5 pooling, 3 dense layers. Only 16 of these are weighted layers. There are 64 filters in Convolution 1 layer, 128 filters in Convolution 2 layer, 256 filters in Convolution 3 layer, 512 filters in each of Convolution 4 and Convolution 5 layers.

The VGG-16 network was trained on the ImageNet dataset containing more than 14 million images and 1000 classes and achieved 92.7% accuracy.

2.3.2. ResNet-50 model

ResNet was introduced by He et al. in 2015 and the model became the winner of ImageNet competition with a 3.57% error rate (He et al. 2016). Unlike traditional sequential network architectures, ResNet has a structure based on microarchitecture modules. In theory, success is expected to increase as the number of layers in a model increases. However, increasing the number of parameters makes training and optimization difficult. During training, neurons that do not have high activations are ineffective in the deep neural networks, and in this case, residues appear in the network. ResNet is created by adding blocks that feed residual values to the next layers (Fig. 7).

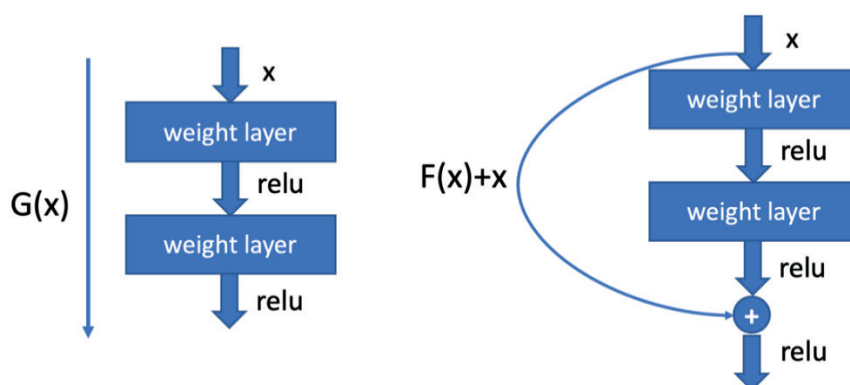


Figure 7. Typical CNN and ResNet

In a normal CNN, the model that runs sequentially from input to output is represented by a nonlinear $G(x)$ function. In ResNet, by creating a shortcut from the input to the output, the input value (x) is added to the $F(x)$ function arithmetically.

Then $F(x) + x$ is passed through ReLU. With the aim of conveying the values in the past layers to the next layers more strongly, the input at the end of the 2nd layer in ResNet is added.

ResNet50 is a special version of the ResNet model with 50 weighted layers.

2.3.3. Inception v3 model

Inception networks constitute one of the important steps in the development of CNN architectures. There are 4 versions with differences in performance and accuracy. Inception v3 was designed by Szegedy et al. and the model was successful in ImageNet competition with a low error rate (Szegedy et al. 2016). The auxiliary classifiers used in the Inception architecture did not contribute until the end of the training process in the second version. Inception v3 was designed with minor changes in this version. Inception v3 consists of 42 layers. In addition to the previous versions, there are label smoothing, normalization for auxiliary classifiers, 7x7 convolution, and RMSProp optimizer. The structure of the Inception v3 model is shown in Fig. 8.

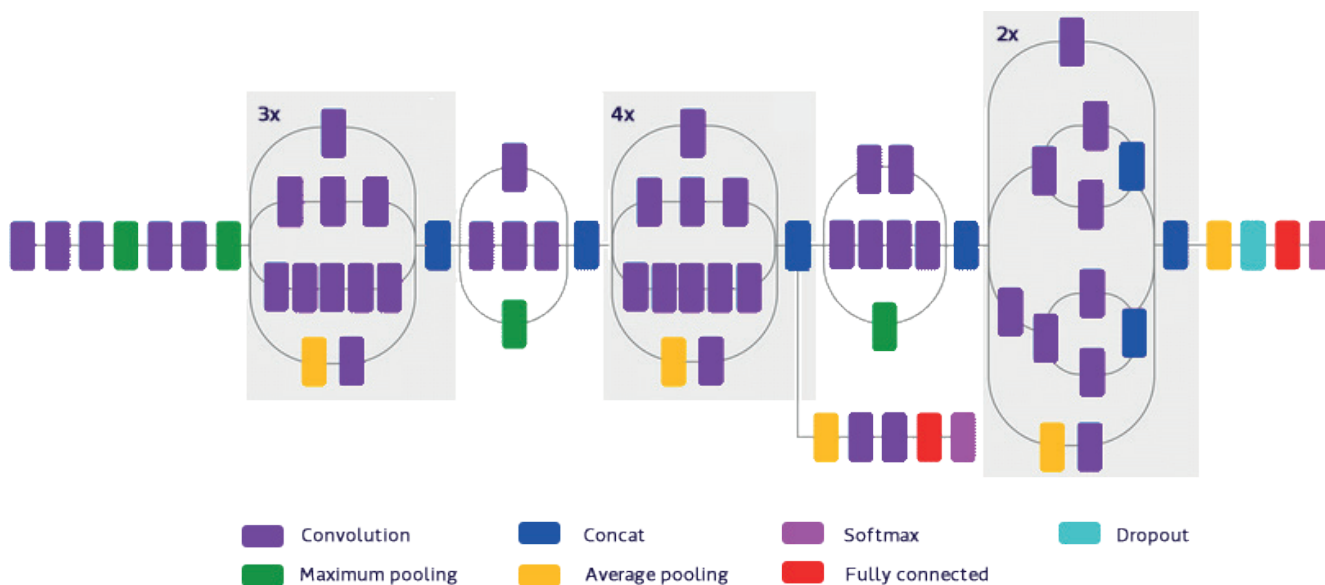


Figure 8. Structure of Inception v3 model (Agrawat and Raval 2018)

3. EXPERIMENTAL STUDY

In this study, a fully automated classification approach for brain tumors was presented. First, preprocessing was performed on the raw MRI images, the boundaries of the brain tissue were determined automatically, and the images were cropped. Then, the dataset was expanded by applying the data augmentation method. The transfer learning method was used to reduce the processing load and achieve successful results with a small amount of data.

3.1. Performance Metrics

There are several metrics used to evaluate a classifier. Accuracy is the most widely used metric. Accuracy in classification is defined as the ratio of the number of correctly classified samples to the total number of data (Eq. 1).

$$accuracy = \frac{\text{number of correctly classified samples}}{\text{total number of samples}} \quad (1)$$

Classification accuracy is an effective parameter to characterize performance if the test dataset contains an equal number of samples from each class, i.e., it is balanced. However, when working with an unbalanced dataset, the system should be evaluated with more performance metrics. For this purpose, confusion matrices are used. A confusion matrix gives true and false classifications as a table. A confusion matrix contains values for four possible outcomes: When the original positive data is correctly classified as positive it is called true positive (TP), when the original positive data is classified incorrectly

as negative it is called false negative (FN), when the original negative data is classified correctly as negative it is called true negative (TN), when the original negative data is classified incorrectly as positive it is called false positive (FP). With these values obtained from the confusion matrix, different metrics are calculated to express the performance of the classifier.

Recall (sensitivity) which is calculated from these results is expressed in Eq. 2. Recall indicates the ability of the classification to detect true positives.

$$recall = \frac{TP}{TP + FN} \quad (2)$$

Precision is calculated as shown in Eq. 3. Precision indicates the ability of the classification to eliminate false positives.

$$precision = \frac{TP}{TP + FP} \quad (3)$$

The F1 score used to express the balance between recall and precision is the harmonic mean of these two metrics. The F1 score is calculated as expressed in Eq. 4.

$$F1 \text{ score} = 2 * \frac{precision * recall}{precision + recall} \quad (4)$$

For each model used in this study, recall, precision, and F1 score values were evaluated along with the accuracy parameter.

3.2. Findings and Discussion

In this study, the dataset consisting of 253 brain MR images was classified using three different pre-trained models. All other external parameters such as learning rate optimization, batch size, and number of epochs were the same for each model applied. The dataset was divided into 80% training and 20% test sets. 20% of the test set was used as a validation set.

The accuracy and the loss graphs on epoch basis were given for each model. Besides, the performance of each model was compared in terms of the specified metrics.

VGG-16 model: The accuracy and loss graphs obtained as a result of the training process are given in Fig. 9.

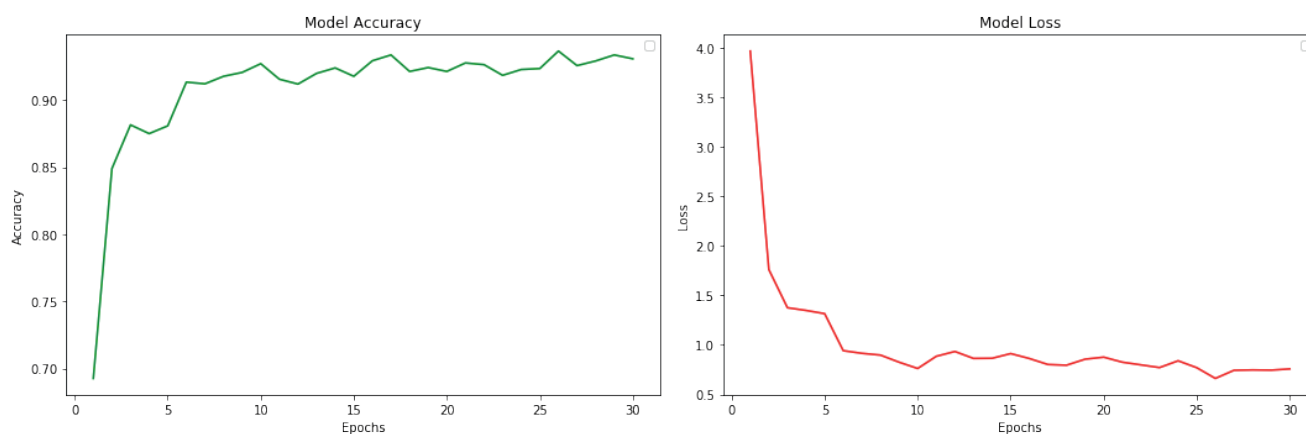


Figure 9. Accuracy and loss graphs of the VGG-16 model

The accuracy obtained at the end of the 30 epochs is 94.42%. Recall is 83.86%, precision is 100%, F1 score is 91.22%.

ResNet50 model: The accuracy and loss graphs obtained as a result of the training process are given in Fig. 10.

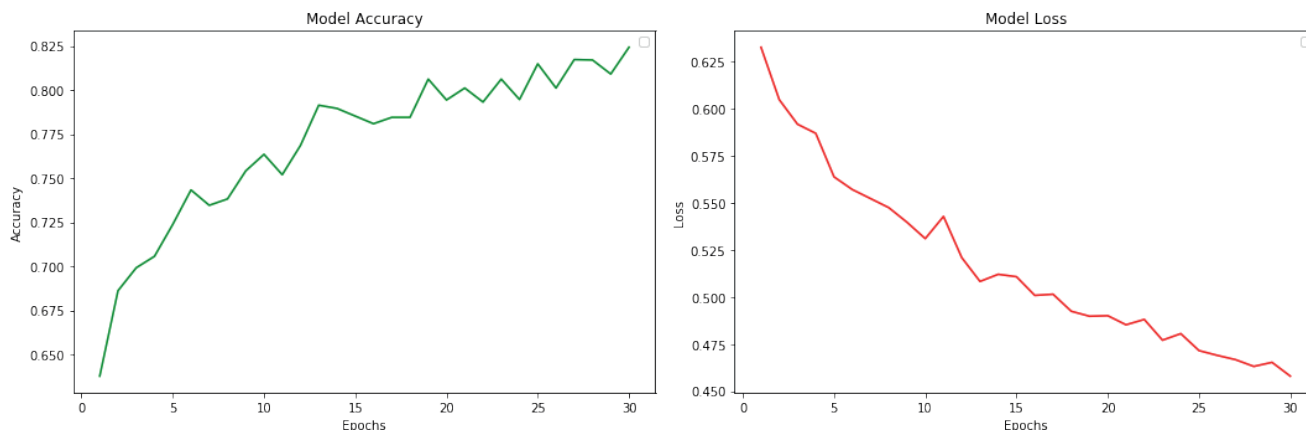


Figure 10. Accuracy and loss graphs of the ResNet50 model

The accuracy obtained at the end of the 30 epochs is 82.49%. Recall is 82.94%, precision is 100%, F1 score is 90.67%.

Inception v3 model: The accuracy and loss graphs obtained as a result of the training process are given in Fig. 11.

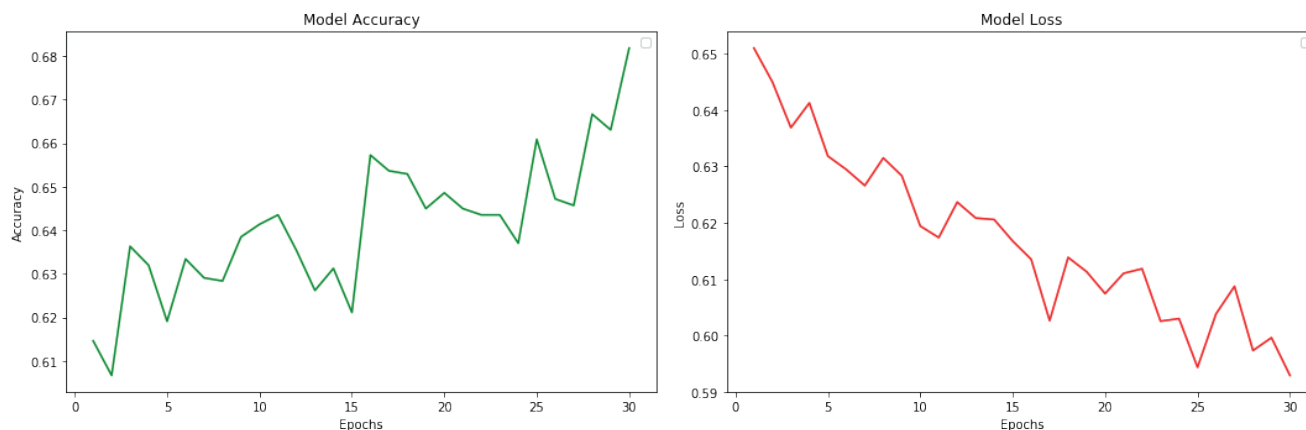


Figure 11. Accuracy and loss graphs of the Inception V3 model

The accuracy obtained at the end of the 30 epochs is 68.13%. Recall is 50.48%, precision is 40.12%, F1 score is 44.70%.

The performance results of the three pre-trained models in the same number of epochs are given in Table 2.

Table 2

Performance measurements

Model	Accuracy	Recall	Precision	F1 Score
Vgg-16	94.42%	83.86%	100%	91.22%
ResNet50	82.49%	82.94%	100%	90.67%
Inception V3	68.13%	50.48%	40.12%	44.70%

When the models are compared in terms of metric values, it is seen that the most successful model on the dataset used is Vgg-16 with an accuracy of 94.42%. The second successful model is ResNet50 with 82.49% accuracy. The Inception V3 model showed significantly lower success compared to these two models. The Vgg-16 and ResNet50 models are similar in terms of their ability to detect the right positives (recall) and to eliminate false positives (precision). However, Vgg-16 model was superior in terms of all metrics.

Since the dataset used in this study was imbalanced, we aimed to increase the success using the data augmentation technique. In order to reveal the effect of data increment on prediction success, each model was re-run without applying data augmentation,

keeping other parameters the same. The accuracy values obtained with and without data augmentation are comparatively given in Table 3.

Table 3

Model accuracies with and without data augmentation

Model	Without augmentation	With augmentation
Vgg-16	85.92%	94.42%
ResNet50	72.59%	82.49%
Inception V3	53.14%	68.13%

According to the measurements obtained, it was observed that the application of data augmentation brought an increase of up to 22% in the success of the model.

Most studies in the literature on the classification of brain MR images used machine learning techniques with handcrafted features. The manual extraction of features is a burden and increases the error rate. Deep learning enables automatic discovery of the features of MR images with its self-learning ability. A comparison of the findings obtained in this study with similar recent studies that classify brain MR images using deep learning methods in the literature is given in Table 4.

Table 4

Comparison of the results obtained in this study with the literature

Reference	Dataset	Method	The highest accuracy (%)
Shahamat and Abadeh 2020	140 MRI images	Custom CNN	70
Zhang et al. 2020	361 MRI images	R-CNN	79
Naser and Deen 2020	110 MRI images	CNN-UNet	89
Zhou et al. 2019	192 MRI images	InceptionV3	69
Saxena et al 2020	3064 MRI images	Vgg-16	90
Mlynarski et al. 2019	285 MRI images	CNN-UNet	80.92
· Afshar et al. 2019	3064 MRI images	CapsNet	90.89
· Abiwinanda et al. 2018	3064 MRI images	Custom CNN	84.19
Pashai et al. 2018	3064 MRI images	Custom CNN	81.09
Yang et al. 2018	113 MRI images	GoogLeNet	86.7
This study	253 MRI images	Vgg-16	94.42

Table 4 shows that the accuracy of the studies performed using various CNN models on different sizes of datasets in the literature varies between 70% and 90%. This study demonstrates that the best result of 94.42% obtained by using the Vgg-16 model was higher than similar studies in the literature and Saxena et al.'s study using the same model. Inception V3 was the least successful model of this study with 68.13% accuracy. In the literature, results obtained in other studies using the Inception V3 model in the classification of brain MR images were also between 55% and 69% (Saxena et al. 2020) (Zhou et al. 2019).

4. CONCLUSION

Due to the high image resolution it provides, MRI helps detect abnormalities in the brain with greater ease. Traditionally, MR images are interpreted by radiologists and tumors are diagnosed. However, with the advances in medical imaging techniques, it has become difficult to interpret manually large amounts of data produced in reasonable periods. As deep learning can effectively identify complex relationships, it has begun to attract great attention in the field of medical image analysis. Recently, deep learning has been used in the analysis of brain MRI images, but automatic detection of brain tumors is still a problem waiting to be solved due to variations in the morphological structure of the brain and differences in imaging technology. CNNs are commonly used in the deep learning-based brain tumor segmentation problems. One of the methods used to increase data processing performance is transfer learning. With the transfer of the learned parameters, high success is achieved while reducing the workload of the new model. Besides, transfer learning ensures successful results in case of low training data.

In this study, classifications were made using pre-trained Vgg-16, ResNet50, and Inception V3 models on a dataset consisting of 253 brain MRI images with and without tumors. Raw MR images were preprocessed to reduce the workload in the training

process. The boundaries of the brain tissue were determined over the raw image and the image was cropped and resized. Data augmentation was applied to balance the dataset class distributions. The classification process with three different models was evaluated with accuracy, recall, precision, and F1-score metrics. Vgg-16 model showed the highest success with 94.42% accuracy, 83.86% recall, 100% precision and 91.22% F1 score. This was followed by the ResNet50 model with an accuracy of 82.49%. Inception V3 model showed the lowest success in this study as in similar studies in the literature. Findings obtained in this study were compared with recent similar studies in the literature. The Vgg-16, which was the most successful model in this study, was shown to be more successful than the peer studies in the literature. In this study, it was shown that the transfer learning method provides successful results with a limited amount of data and fewer epochs. An alternative support system solution was presented for experts that will facilitate tumor detection from brain MR images.

Hakem Değerlendirmesi: Dış bağımsız.

Çıkar Çatışması: Yazar çıkar çatışması bildirmemiştir.

Finansal Destek: Yazar bu çalışma için finansal destek almadığını beyan etmiştir.

Peer-review: Externally peer-reviewed.

Conflict of Interest: The author has no conflict of interest to declare.

Grant Support: The author declared that this study has received no financial support.

References/Kaynaklar

- Abiwinanda, Nyoman, Muhammad Hanif, S. Tafwida Hesaputra, Astri Handayani, and Tati Rajab Mengko. 2019. "Brain Tumor Classification Using Convolutional Neural Network." Pp. 183–89 in World congress on medical physics and biomedical engineering 2018. Springer.
- Afshar, Parnian, Arash Mohammadi, and Konstantinos N. Plataniotis. 2018. "Brain Tumor Type Classification via Capsule Networks." Pp. 3129–3133 in 2018 25th IEEE International Conference on Image Processing (ICIP). IEEE.
- Agn, Mikael, Per Munck af Rosenschöld, Oula Puonti, Michael J. Lundemann, Laura Mancini, Anastasia Papadaki, Steffi Thust, John Ashburner, Ian Law, and Koen Van Leemput. 2019. "A Modality-Adaptive Method for Segmenting Brain Tumors and Organs-at-Risk in Radiation Therapy Planning." *Medical Image Analysis* 54:220–237.
- Agravat, Rupal R., and Mehul S. Raval. 2018. "Deep Learning for Automated Brain Tumor Segmentation in MRI Images." Pp. 183–201 in *Soft Computing Based Medical Image Analysis*, edited by N. Dey, A. S. Ashour, F. Shi, and V. E. Balas. Academic Press.
- Akkus, Zeynettin, Alfiia Galimzianova, Assaf Hoogi, Daniel L. Rubin, and Bradley J. Erickson. 2017. "Deep Learning for Brain MRI Segmentation: State of the Art and Future Directions." *Journal of Digital Imaging* 30(4):449–59.
- Bauer, Stefan, Roland Wiest, Lutz-P. Nolte, and Mauricio Reyes. 2013. "A Survey of MRI-Based Medical Image Analysis for Brain Tumor Studies." *Physics in Medicine & Biology* 58(13):R97.
- Belaid, Ouiza Nait, and Malik Loudini. 2020. "Classification of Brain Tumor by Combination of Pre-Trained VGG16 CNN." *Journal of Information Technology Management* 12(2):13–25.
- Cancer.Net, "Brain Tumor - Statistics." Retrieved October 28, 2020 (<https://www.cancer.net/cancer-types/brain-tumor/statistics>).
- Cao, Chensi, Feng Liu, Hai Tan, Deshou Song, Wenjie Shu, Weizhong Li, Yiming Zhou, Xiaochen Bo, and Zhi Xie. 2018. "Deep Learning and Its Applications in Biomedicine." *Genomics, Proteomics & Bioinformatics* 16(1):17–32.
- Chakrabarty, N. n.d. "Brain MRI Images for Brain Tumor Detection." Kaggle. Retrieved May 28, 2020 (<https://kaggle.com/navoneel/brain-mri-images-for-brain-tumor-detection>).
- Chaplot, Sandeep, Lalit M. Patnaik, and N. R. Jagannathan. 2006. "Classification of Magnetic Resonance Brain Images Using Wavelets as Input to Support Vector Machine and Neural Network." *Biomedical Signal Processing and Control* 1(1):86–92.
- Charron, Odeline, Alex Lallemand, Delphine Jarnet, Vincent Noblet, Jean-Baptiste Clavier, and Philippe Meyer. 2018. "Automatic Detection and Segmentation of Brain Metastases on Multimodal MR Images with a Deep Convolutional Neural Network." *Computers in Biology and Medicine* 95:43–54.
- Chelghoum, Rayene, Ameer Ikhlef, Amina Hameurlaine, and Sabir Jacquir. 2020. "Transfer Learning Using Convolutional Neural Network Architectures for Brain Tumor Classification from MRI Images." Pp. 189–200 in *Artificial Intelligence Applications and Innovations*, edited by I. Maglogiannis, L. Iliadis, and E. Pimenidis. Cham: Springer International Publishing.
- Cheng, Jie-Zhi, Dong Ni, Yi-Hong Chou, Jing Qin, Chui-Mei Tiu, Yeun-Chung Chang, Chiun-Sheng Huang, Dinggang Shen, and Chung-Ming Chen. 2016. "Computer-Aided Diagnosis with Deep Learning Architecture: Applications to Breast Lesions in US Images and Pulmonary Nodules in CT Scans." *Scientific Reports* 6(1):1–13.
- Deepak, S., and P. M. Ameer. 2019. "Brain Tumor Classification Using Deep CNN Features via Transfer Learning." *Computers in Biology and Medicine* 111:103345. doi: 10.1016/j.combiomed.2019.103345.
- Deepak, S., and P. M. Ameer. 2020. "Automated Categorization of Brain Tumor from MRI Using CNN Features and SVM." *Journal of Ambient Intelligence and Humanized Computing*. doi: 10.1007/s12652-020-02568-w.
- Deniz, Erkan, Abdulkadir Şengür, Zehra Kadiroğlu, Yanhui Guo, Varun Bajaj, and Ümit Budak. 2018. "Transfer Learning Based Histopathologic Image Classification for Breast Cancer Detection." *Health Information Science and Systems* 6(1):18.

- Domingues, Rémi, Maurizio Filippone, Pietro Michiardi, and Jihane Zouaoui. 2018. "A Comparative Evaluation of Outlier Detection Algorithms: Experiments and Analyses." *Pattern Recognition* 74:406–21.
- El-Dahshan, El-Sayed Ahmed, Tamer Hosny, and Abdel-Badeeh M. Salem. 2010. "Hybrid Intelligent Techniques for MRI Brain Images Classification." *Digital Signal Processing* 20(2):433–41.
- Gu, Yu, Xiaoqi Lu, Lidong Yang, Baohua Zhang, Dahua Yu, Ying Zhao, Lixin Gao, Liang Wu, and Tao Zhou. 2018. "Automatic Lung Nodule Detection Using a 3D Deep Convolutional Neural Network Combined with a Multi-Scale Prediction Strategy in Chest CTs." *Computers in Biology and Medicine* 103:220–231.
- Havaei, Mohammad, Axel Davy, David Warde-Farley, Antoine Biard, Aaron Courville, Yoshua Bengio, Chris Pal, Pierre-Marc Jodoin, and Hugo Larochelle. 2017. "Brain Tumor Segmentation with Deep Neural Networks." *Medical Image Analysis* 35:18–31.
- He, K., X. Zhang, S. Ren, and J. Sun. 2016. "Deep Residual Learning for Image Recognition." Pp. 770–78 in 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR).
- Hussein, Sarfaraz, Pujan Kandel, Candice W. Bolan, Michael B. Wallace, and Ulas Bagci. 2019. "Lung and Pancreatic Tumor Characterization in the Deep Learning Era: Novel Supervised and Unsupervised Learning Approaches." *IEEE Transactions on Medical Imaging* 38(8):1777–1787.
- Jain, Rachna, Nikita Jain, Akshay Aggarwal, and D. Jude Hemanth. 2019. "Convolutional Neural Network Based Alzheimer's Disease Classification from Magnetic Resonance Brain Images." *Cognitive Systems Research* 57:147–59. doi: 10.1016/j.cogsys.2018.12.015.
- Kamnitsas, Konstantinos, Christian Ledig, Virginia FJ Newcombe, Joanna P. Simpson, Andrew D. Kane, David K. Menon, Daniel Rueckert, and Ben Glocker. 2017. "Efficient Multi-Scale 3D CNN with Fully Connected CRF for Accurate Brain Lesion Segmentation." *Medical Image Analysis* 36:61–78.
- Kaur, Taranjit, and Tapan Kumar Gandhi. 2020. "Deep Convolutional Neural Networks with Transfer Learning for Automated Brain Image Classification." *Machine Vision and Applications* 31:1–16.
- Kleesiek, Jens, Armin Biller, Gregor Urban, U. Kothe, Martin Bendszus, and F. Hamprecht. 2014. "Ilastik for Multi-Modal Brain Tumor Segmentation." *Proceedings MICCAI BraTS (Brain Tumor Segmentation Challenge)* 12–17.
- Kumar, Sanjeev, Chetna Dabas, and Sunila Godara. 2017. "Classification of Brain MRI Tumor Images: A Hybrid Approach." *Procedia Computer Science* 122:510–17.
- LeCun, Yann, Yoshua Bengio, and Geoffrey Hinton. 2015. "Deep Learning." *Nature* 521(7553):436–444.
- Litjens, Geert, Clara I. Sánchez, Nadya Timofeeva, Meyke Hermesen, Iris Nagtegaal, Iringo Kovacs, Christina Hulsbergen-Van De Kaa, Peter Bult, Bram Van Ginneken, and Jeroen Van Der Laak. 2016. "Deep Learning as a Tool for Increased Accuracy and Efficiency of Histopathological Diagnosis." *Scientific Reports* 6:26286.
- Menze, Bjoern H., Andras Jakab, Stefan Bauer, Jayashree Kalpathy-Cramer, Keyvan Farahani, Justin Kirby, Yuliya Burren, Nicole Porz, Johannes Slotboom, and Roland Wiest. 2014. "The Multimodal Brain Tumor Image Segmentation Benchmark (BRATS)." *IEEE Transactions on Medical Imaging* 34(10):1993–2024.
- Mlynarski, Pawel, Hervé Delingette, Antonio Criminisi, and Nicholas Ayache. 2019. "Deep Learning with Mixed Supervision for Brain Tumor Segmentation." *Journal of Medical Imaging* 6(3):034002.
- Mohsen, Heba, El-Sayed A. El-Dahshan, El-Sayed M. El-Horbaty, and Abdel-Badeeh M. Salem. 2018. "Classification Using Deep Learning Neural Networks for Brain Tumors." *Future Computing and Informatics Journal* 3(1):68–71.
- Moritz, Chad H., Victor M. Haughton, Dietmar Cordes, Michelle Quigley, and M. Elizabeth Meyerand. 2000. "Whole-Brain Functional MR Imaging Activation from a Finger-Tapping Task Examined with Independent Component Analysis." *American Journal of Neuroradiology* 21(9):1629–35.
- Naser, Mohamed A., and M. Jamal Deen. 2020. "Brain Tumor Segmentation and Grading of Lower-Grade Glioma Using Deep Learning in MRI Images." *Computers in Biology and Medicine* 121:103758.
- Nayak, Deepak Ranjan, Ratnakar Dash, and Banshidhar Majhi. 2016. "Brain MR Image Classification Using Two-Dimensional Discrete Wavelet Transform and AdaBoost with Random Forests." *Neurocomputing* 177:188–97.
- Pashaie, Ali, Hedieh Sajedi, and Niloofar Jazayeri. 2018. "Brain Tumor Classification via Convolutional Neural Network and Extreme Learning Machines." Pp. 314–319 in 2018 8th International conference on computer and knowledge engineering (ICCKE). IEEE.
- Prastawa, Marcel, Elizabeth Bullitt, Sean Ho, and Guido Gerig. 2004. "A Brain Tumor Segmentation Framework Based on Outlier Detection." *Medical Image Analysis* 8(3):275–83.
- Rajinikanth, V., Suresh Chandra Satapathy, Steven Lawrence Fernandes, and S. Nachiappan. 2017. "Entropy Based Segmentation of Tumor from Brain MR Images—a Study with Teaching Learning Based Optimization." *Pattern Recognition Letters* 94:87–95.
- Ranjan Nayak, Deepak, Ratnakar Dash, and Banshidhar Majhi. 2017. "Stationary Wavelet Transform and Adaboost with SVM Based Pathological Brain Detection in MRI Scanning." *CNS & Neurological Disorders-Drug Targets (Formerly Current Drug Targets-CNS & Neurological Disorders)* 16(2):137–49.
- Rehman, Arshia, Saeeda Naz, Muhammad Imran Razzak, Faiza Akram, and Muhammad Imran. 2020. "A Deep Learning-Based Framework for Automatic Brain Tumors Classification Using Transfer Learning." *Circuits, Systems, and Signal Processing* 39(2):757–75. doi: 10.1007/s00034-019-01246-3.
- Reza, S., and K. M. Iftekharruddin. 2014. "Improved Brain Tumor Tissue Segmentation Using Texture Features." *Proceedings MICCAI BraTS (Brain Tumor Segmentation Challenge)* 27–30.
- Saritha, M., K. Paul Joseph, and Abraham T. Mathew. 2013. "Classification of MRI Brain Images Using Combined Wavelet Entropy Based Spider Web Plots and Probabilistic Neural Network." *Pattern Recognition Letters* 34(16):2151–2156.
- Saxena, Priyansh, Akshat Maheshwari, and Saumil Maheshwari. 2020. "Predictive Modeling of Brain Tumor: A Deep Learning Approach." Pp. 275–85 in *Innovations in Computational Intelligence and Computer Vision*. Springer.

- Shahamat, Hossein, and Mohammad Saniee Abadeh. 2020. "Brain MRI Analysis Using a Deep Learning Based Evolutionary Approach." *Neural Networks* 126:218–34.
- Simonyan, K., and A. Zisserman. 2020. "Very Deep Convolutional Networks for Large-Scale Image Recognition. ArXiv 1409.1556 (09 2014)." URL <https://arxiv.org/abs/1409.1556>. Accessed: February.
- Szegedy, Christian, Vincent Vanhoucke, Sergey Ioffe, Jon Shlens, and Zbigniew Wojna. 2016. "Rethinking the Inception Architecture for Computer Vision." Pp. 2818–2826 in *Proceedings of the IEEE conference on computer vision and pattern recognition*.
- Tustison, Nick, Max Wintermark, Chris Durst, and Brian Avants. 2013. "Ants and Arboles." *Multimodal Brain Tumor Segmentation* 47.
- Wang, Shuihua, Yudong Zhang, Zhengchao Dong, Sidan Du, Genlin Ji, Jie Yan, Jiquan Yang, Qiong Wang, Chunmei Feng, and Preetha Phillips. 2015. "Feed-Forward Neural Network Optimized by Hybridization of PSO and ABC for Abnormal Brain Detection." *International Journal of Imaging Systems and Technology* 25(2):153–164.
- Yang, Yang, Lin-Feng Yan, Xin Zhang, Yu Han, Hai-Yan Nan, Yu-Chuan Hu, Bo Hu, Song-Lin Yan, Jin Zhang, Dong-Liang Cheng, Xiang-Wei Ge, Guang-Bin Cui, Di Zhao, and Wen Wang. 2018. "Glioma Grading on Conventional MR Images: A Deep Learning Study With Transfer Learning." *Frontiers in Neuroscience* 12:804–804. doi: 10.3389/fnins.2018.00804.
- Yousefi, Mina, Adam Krzyżak, and Ching Y. Suen. 2018. "Mass Detection in Digital Breast Tomosynthesis Data Using Convolutional Neural Networks and Multiple Instance Learning." *Computers in Biology and Medicine* 96:283–293.
- Zhang, Min, Geoffrey S. Young, Huai Chen, Jing Li, Lei Qin, J. Ricardo McFaline-Figueroa, David A. Reardon, Xinhua Cao, Xian Wu, and Xiaoyin Xu. 2020. "Deep-Learning Detection of Cancer Metastases to the Brain on MRI." *Journal of Magnetic Resonance Imaging* 52(4):1227–36.
- Zhang, Yu-Dong, Shuihua Wang, Zhengchao Dong, Preetha Phillip, Genlin Ji, and Jiquan Yang. 2015. "Pathological Brain Detection in Magnetic Resonance Imaging Scanning by Wavelet Entropy and Hybridization of Biogeography-Based Optimization and Particle Swarm Optimization." *Progress In Electromagnetics Research* 152:41–58.
- Zhang, Yudong, Shuihua Wang, Genlin Ji, and Zhengchao Dong. 2013. "An MR Brain Images Classifier System via Particle Swarm Optimization and Kernel Support Vector Machine." *The Scientific World Journal* 2013.
- Zhou, Leilei, Zuoheng Zhang, Yu-Chen Chen, Zhen-Yu Zhao, Xin-Dao Yin, and Hong-Bing Jiang. 2019. "A Deep Learning-Based Radiomics Model for Differentiating Benign and Malignant Renal Tumors." *Translational Oncology* 12(2):292–300.
- Zuo, Haiqiang, Heng Fan, Erik Blasch, and Haibin Ling. 2017. "Combining Convolutional and Recurrent Neural Networks for Human Skin Detection." *IEEE Signal Processing Letters* 24(3):289–293.

Predicting Health Spending in Turkey Using the GPR, SVR, and DT Models

Türkiye Sağlık Harcamalarının GPR, SVR ve DT Modelleri ile Tahmini

Didem Gülerüz¹ 



ABSTRACT

Rising healthcare costs for countries and the long-term maintainability of this situation are at the center of the political agenda. The steady increase in health spending puts pressure on government budgets, healthcare, and personal patient financing. Policymakers would like to plan reforms to reduce these costs to adapt to problems that may arise. This has led planners to decision support systems and forecasting models. In this paper, three machine learning algorithms, namely Support Vector Regression (SVR), Decision Tree Regression (DT), and Gaussian Process Regression (GPR) are employed to design a forecasting model for Health Spendings (HS) of Turkey considering various determinants. Gross domestic product per capita, urban population rate, unemployment rate, population ages 65 and above, the life expectancy, the physicians' rate, and the total number of hospital beds are used as inputs. The data set consists of 30 years between 1990-2019, which splits as training and test sets. Developed models were compared considering performance metrics, and the most accurate model was identified. The coefficient of determinations (R^2) for SVR, GPR, and DT models are 0.9929, 0.9989, and 0.9611 in the training phase, 0.9536, 0.8944, and 0.1166 in the testing stage, respectively. Therefore, the SVR model has accurate prediction results with the highest R^2 and the least root mean square error values in the testing phase. The study showed that the proposed SVR model reduced RMSE value by 32.02% and 39.66% compared to the GPR and DT models, respectively. Consequently, the Health Spendings of Turkey can be predicted by employing SVR with high accuracy.

Keywords: Health Spending, Gaussian Process Regression, Decision Tree, Support Vector Regression

ÖZ

Ülkeler için artan sağlık maliyetleri ve bu durumun uzun vadeli sürdürülebilirliği siyasi gündemin merkezinde yer almaktadır. Sağlık harcamalarındaki sürekli artış, hükümet bütçeleri, sağlık hizmetleri ve kişisel hasta finansmanı üzerinde baskı oluşturmaktadır. Politika yapıcılar, ortaya çıkabilecek sorunlara uyum sağlamak ve bu maliyetleri düşürmek için reformlar planlamak isterler. Bu durum, planlayıcıları karar destek sistemlerine ve tahmin modellerine yönlendirmiştir. Bu çalışmada, Türkiye'nin Sağlık Harcaması (HS) için çeşitli belirleyicileri dikkate alan bir tahmin modeli tasarlamak amacıyla Destek Vektör Regresyonu (SVR), Regresyon Ağacı (DT) ve Gauss Süreç Regresyonu (GPR) olmak üzere üç makine öğrenme algoritması kullanılmıştır. Kişi başına gayri safi yurtiçi hasıla, kentsel nüfus oranı, işsizlik oranı, 65 yaş ve üstü nüfus, ortalama yaşam süresi, hekim oranı ve toplam hastane yatak sayısı girdi değişkenleri olarak belirlenmiştir. Veri seti eğitim ve test verisi olarak ayrılmış ve 1990-2019 yılları arası 30 yılı kapsamaktadır. Geliştirilen modeller performans ölçütleri dikkate alınarak karşılaştırılmış ve en iyi model belirlenmiştir. SVR, GPR ve DT modelleri için belirleme katsayısı (R^2) eğitim aşamasında sırasıyla 0.9929, 0.9989 ve 0.9611, test aşamasında sırasıyla 0.9536, 0.8944 ve 0.1166'dır. Ayrıca, SVR modeli, test aşamasında en yüksek R^2 ve en düşük kök ortalama kare hatası değerleri ile en iyi tahmin sonuçlarına sahiptir. Çalışma, önerilen SVR modelinin RMSE değerini diğer GPR ve DT modellerine kıyasla sırasıyla % 32.02 ve % 39.66 azalttığını göstermiştir. Sonuç olarak, Türkiye'nin sağlık harcamaları SVR modeli kullanılarak yüksek doğrulukta tahmin edilebilir.

Anahtar kelimeler: Sağlık Harcaması, Gauss Süreç Regresyonu, Regresyon Ağacı, Destek Vektör Regresyonu

¹(PhD.), Bayburt University, Faculty of Applied Sciences, Department of Management Information Systems Bayburt, Turkey

ORCID: D.G. 0000-0003-4198-9997

Corresponding author:

Didem GÜLERÜZ

Bayburt University, Faculty of Applied Sciences, Department of Management Information Systems Bayburt, Turkey

E-mail address: dguleryuz@bayburt.edu.tr

Submitted: 18.02.2021

Revision Requested: 21.03.2021

Last Revision Received: 15.04.2021

Accepted: 24.05.2021

Published Online: 29.06.2021

Citation: Gülerüz, D. (2021). Predicting health spending in Turkey using the GPR, SVR, and DT models. *Acta Infologica*, 5(1), 155-166. <https://doi.org/10.26650/acin.882660>

1. INTRODUCTION

People are living longer than ever before thanks to developing health technologies and improved living standards. According to Turkstat statistics, Turkey's life expectancy has increased in recent years (Turkstat, 2020). Therefore, healthcare expectations have also increased due to the developments of technological advances. As a corollary result of the increasing expectations for healthcare services, the expenditures of governments in this area will also increase. Determining the reasons for the increase in expenditures is very important for the planning of health systems.

The existing service system must be operated optimally for healthcare services to provide the proper services to the right person at the right time, the problems in the system must be identified, and solutions must be specific to these problems. With developing technology, many methods and technologies can be used in health systems. However, it is not possible to provide all of the technologies. For this reason, the concept of health economics is drawing attention today. Optimizing available resources under constraints on variables such as productivity, patient safety, and clinical efficacy will be an efficient tool for decision-makers.

Health and health services are in a reciprocal relationship with the economy. However, studies defining this relationship are divided into two classes. The first is to investigate the links between health status and income distribution, and the second is to identify the interrelation between health financing policies and economic outcomes (Martín et al., 2011). Health indicators play an important role in determining the development of countries. Some previous studies, such as Bauer et al. (2006) and Dalbokova and Krzyzanowski (2002), claim a positive correlation between countries' health levels and their development. Also, people's welfare, labor productivity, demographic and human capital factors directly impact health spending.

The number of studies that investigate the determinants of healthcare spending and estimate the expenditures according to related determinants is increasing day by day. In the light of previous studies, the determinants affecting health expenditures can be classified as environmental, social, and economic. Gross domestic product, unemployment rate, physicians, urbanization rate, and gas emissions were used frequently as determinants of health expenditure in previous studies. Also, most studies investigated the effects of environmental quality on health expenditures (Martín et al., 2011).

Health spending in Turkey, primarily financed by the government, should be optimized to avoid the extra cost of public spending and it is expected to increase in the coming years. This situation depends on ensuring financial efficiency in health services. Excessive healthcare financed by the Turkish government and the advancement of technology in health services negatively affect health expenditures, threatening the system's financial sustainability. The optimization of the system is possible with properly selected management strategies. For instance, the application of lean service tools in the health sector will increase the efficiency of the system and decrease expenditures (Efe and Efe, 2016). Therefore, decision support systems designed to manage health expenditures should predict health expenditure, determine the indicators of health expenditures and optimize the resources (Ministry of Development, 2014).

As healthcare needs increase, health expenditures are expected to increase to meet these demands. In most industrialized economies, healthcare spending is proliferating more than the growth of income. Figure 1 shows the average health spendings for OECD countries and Turkey's health spendings during the last fifteen years (OECD, 2021). According to this graph, health expenditures increased between 2003 and 2019. Unfortunately, Turkey's per capita health spending level remains below the OECD average. According to OECD data, there was a 15.9% increase in health spendings from 2012 to 2018. Although it is not valid for all OECD countries, some OECD countries with higher income and development levels spend more on health services. Identifying the factors that lead to increased health spending is essential for policymakers and scientists to control and plan the spendings.

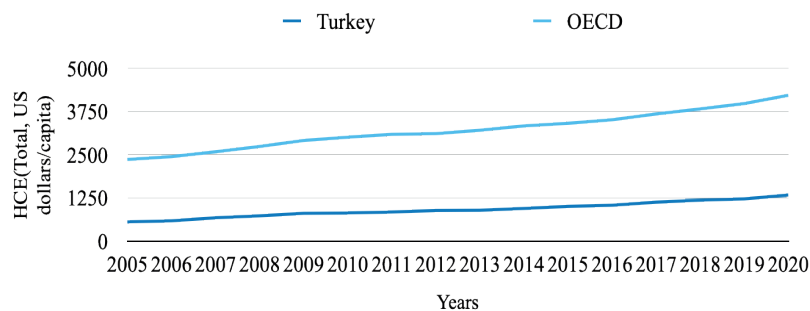


Figure 1. The annual health spendings of Turkey and OECD countries (OECD, 2021)

Improving the health status of citizens will have the potential to increase government spending on healthcare. As seen in Figure 2, the trend in rate of health spendings in GDP has risen and declined over twenty-nine years in Turkey.

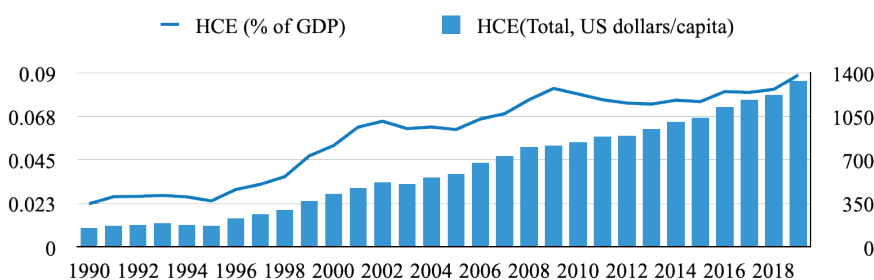


Figure 2. The change of HCE (% GDP) and HCE (Total, US dollars/capita) in Turkey (OECD, 2021)

Moreover, the government's ability to manage health spendings depends on determining the main factors influencing health spendings and selecting a suitable prediction model. With the development of machine learning methods, nonlinear relationships between inputs and outputs can be modeled quickly, and variables that do not explain the target can be included in the model. In previous studies, these relationships were mostly modeled using statistical models. Also, statistical methods widely used in the literature require large data sets. It can show low predictive performance in small-volume data sets. Machine learning models can overcome this problem and have better predictive results with smaller data sets. For this reason, Machine learning models have been widely used recently.

Various methods were employed to predict health spendings in previous studies. These methods can be classified as statistical methods and AI-based methods (Maksimović et al., 2017; Cinaroglu, 2017; Özcan and Tüysüz, 2018). It is impossible to examine all previous studies. For this reason, some selected studies to reflect the literature have been reviewed.

Maksimović et al. (2017) used soft computing methods to estimate gross domestic product (GDP) using health spending determinants. According to the obtained results, the SVR model has better estimation ability than the other employed methods (Maksimović et al., 2017). Mladenović et al. (2016) used a hybrid artificial neural network and fuzzy system (ANFIS) to examine health spending's influence on economic growth. It has been shown that the GDP growth rate is the most useful feature of health expenditures (Mladenović et al., 2016). Cinaroglu (2018) compared ML-based prediction models, namely Lasso Regression, Random Forest, and Support Vector Regression for Turkey. The total health expenditure per capita was selected as the dependent variable. The independent variables are: income, geographical region, total population over 65 years of age, life expectancy at birth, and total population. The different hyperparameters for each model were used in determined intervals, and the results were compared. As a result, the Random Forest model had better estimation results according to different hyperparameter values. Özcan and Tüysüz (2018) used a Grey estimation model to estimate Turkey's per capita healthcare spendings. Since the grey prediction model's performance depends on the model's parameters, these parameters were optimized using the genetic algorithm. Three different mechanisms have been developed to test the predictive ability of the models. As a result, the Grey model's prediction performance has increased with the genetic algorithm (Özcan

and Tüysüz, 2018). Dritsakı and Klazoglou (2019) modeled the total health expenditure of the USA using the Box-Jenkins method. The developed ARIMA (2,1,0) model is better than other employed models considering obtained performance criteria (Dritsakı and Klazoglou, 2019). Ram et al. (2019) employed ARIMA to forecast health expenditure in Iran. The ARIMA (1,1,1) model had the most accurate prediction results (Ram et al., 2019). Samah et al. (2020) aimed to observe health spendings and CO₂ emissions during the COVID-19 outbreak in Malaysia. The dynamic panel data system GMM prediction model was used. It was shown that health spendings and the growth in health services, inflation rate, and unemployment rate significantly impact CO₂ emissions (Samah et al., 2020). Shi et al. (2020) investigated the effects of urbanization on healthcare spendings and behavioral changes. As a result of the study, it was found that increasing urbanization affects the total treatment costs. Moreover, it has been observed that health spendings are affected by the urbanization rate and health insurance level (Shi et al., 2020). Ceylan (2020) estimated Turkey's health expenditures using greenhouse gas emission levels as input variables via optimized support vector regression, feedforward backpropagation neural network, and multivariate linear regression. The SVR model with Bayesian optimization has more effective health expenditure estimation results (Ceylan, 2020). Ceylan and Atalan (2020) used artificial intelligence techniques to estimate health expenditures per capita in Turkey. In the study, optimum determinants from the variables were selected using the genetic algorithm method. Random Forest model with genetic algorithm feature selection had the superior performance among all used models (Ceylan and Atalan, 2020). Cinaroglu (2020) presented a study comparing the predictive ability of logistic regression and random forest models for households that will face catastrophic out-of-pocket (OOP) health expenditure. The dataset was obtained from the household budget survey compiled by TURKSTAT for the year 2012. As a result, the developed random forest model had superior performance in showing households facing catastrophic OOP health expenditure. Durmaz and Pabuçcu (2020) used the ANFIS method to investigate the relevance between public health spendings and labor productivity. G20 countries data were used to analyze relations. The health spendings rate, trade gap, and GDP growth rate were used as input variables, and labor productivity as an output variable. It was determined that the relationship between health spendings and labor force productivity has different effects periodically, and there are linear and opposite relationships (Durmaz and Pabuçcu, 2020). Jia et al. (2021) developed a forecasting model using the gray system theory called the New Structure of the Multivariate Grey Prediction Model. The main driving factors affecting China's health sector were used as input variables, and health expenditures were estimated. The input variables in the model were determined considering the economy, population, health service utilization, and policy. The suggested NSGM (1, N) model provided good predictive results.

In the literature review, few studies used the SVR and DT methods for health spendings prediction. Ceylan (2020) and Maksimović et al. (2017) employed an SVR model to estimate health spendings. Akca et al. (2017) developed a decision tree model to identify features that have measurable effects on HS for OECD countries. In this study, GDP per capita (GDP), population ages 65 and above (PO65), urban population (UP), unemployment rate (UR), life expectancy (LE), the physicians' rate (PY), the total number of hospital beds (HB) were used as input variables to predict health spendings per capita (HS) in Turkey. Some or all of these variables have been used as inputs in the HS estimation literature (Häkkinen et al., 2008; Jakovljevic et al., 2020; Mladenović et al., 2016).

This study aims to contribute to the literature by developing forecasting models using machine learning methods. In this study, SVR, GPR, and DT models were used to estimate health spendings. According to the authors' knowledge, GPR has not been used previously to model health spendings. The SVR and GPR models' results were very close, so that GPR might also be used to model health spendings. Furthermore, the SVR and GPR models provided a stable estimation in a short time. The paper includes four sections. The materials and methods, evaluation of the model accuracy, and developed models' structure are placed in Section 2. Section 3 and Section 4 contain the results of the developed models and the conclusion, respectively.

2. MATERIAL AND METHOD

2.1. Data Collection Process

Turkey's health spending per capita was estimated using GDP per capita, population ages 65 and above, life expectancy, unemployment rate, the number of physicians, the total number of hospital beds, and urban population data in this case. The

data set was gained from several data sources covering the period 1990-2019. Table 1 explains the definitions, abbreviations (Abbr.), and units of the variables and shows the sources.

Table 1

Definition of the Variables

Variable	Abbr.	Unit	Source
Health spendings per capita	HS	Total, US dollars/capita	(OECD, 2021)
GDP per capita	GDP	Constant 2010 billion US\$	(Worldbank, 2021)
Population ages 65 +	PO65	% of the total population	(Worldbank, 2021)
Urban population	UP	% of the total population	(Worldbank, 2021)
Unemployment rate	UR	% of total labor force	(Worldbank, 2021)
Life expectancy	LE	years	(Worldbank, 2021)
Number of Physicians	PY	Per 1,000 people	(OECD, 2021)
Number of Hospital beds	HB	Unit	(Turkstat, 2020)

GDP, PO65, UP, UR, LE, PY, HB are inputs that affect the HS. Table 2 shows the descriptive statistics of inputs and output.

Table 2

The statistics of the data set

	GDP	PO65	UP	UR	LE	PY	HB	HS
Mean	9991.31	6.55	67.54	9.26	71.71	1.44	171828	627.85
Std	2710.69	1.16	4.96	1.70	4.11	0.31	34337	360.03
Minimum	6709.09	4.61	59.20	6.49	64.26	0.9	120738	151.80
Maximum	15068.98	8.73	75.63	13.67	77.54	1.88	237504	1337.17
Kurtosis	-0.97610	-0.96	-1.27	0.12	-1.17	-1.35	-1.12	-1.15
Skewness	0.60	-0.02	0.001	0.48	-0.28	-0.18	0.26	0.27

The input indicators were determined by examining the studies on health expenditures prediction in the literature. The change in the HS is also related to GDP. Many studies investigate the relationship between HS and GDP. Jakovljevic et al. (2020) and Rana et al. (2020) examined the relevance between HS and GDP. Since the elderly population is increasing, scientific studies on the elderly population have increased. It is inferred that the elderly population will rise in the upcoming period. There are many studies investigating the relationship between HS and PO65. Di Matteo (2005), Häkkinen et al. (2008) investigated the effects of the elderly population on health spendings.

One of the crucial effects of urbanization is that healthcare services' structure has changed from rural to urban. This situation necessitated the transition from local institutions, where preventive health services are provided in rural areas, to hybrid institutions, where more complex health services are provided in cities. Therefore, health expenditures are affected by the rate of urbanization. Aliyu and Amadu (2017), Biadgilign et al. (2019) have investigated HS and UR relations. Improvements in economic factors generally have positive effects on health services. Increasing household incomes facilitates the financing of healthcare needs. Also, the decrease in family unemployment prevents the loss of employment-based health insurance and provides an opportunity to meet healthcare needs. Several studies were conducted to analyze this relationship, particular examples being Jakovljevic et al. (2020) and Papanicolas et al. (2019). Turkey's life expectancy has increased from year to year, as seen in Figure 3. Therefore, in parallel with this life expectation, the quality expectation in health services will affect the expenditures. Some studies have investigated this relationship, for example, Anderson and Frogner (2008), Crémieux et al. (1999). The number of hospital beds and physicians are the direct determinants of health spendings. Therefore, these determinants are used to predict health spendings. The annual change of the indicators related to the health spendings of Turkey is seen in Figure 3.

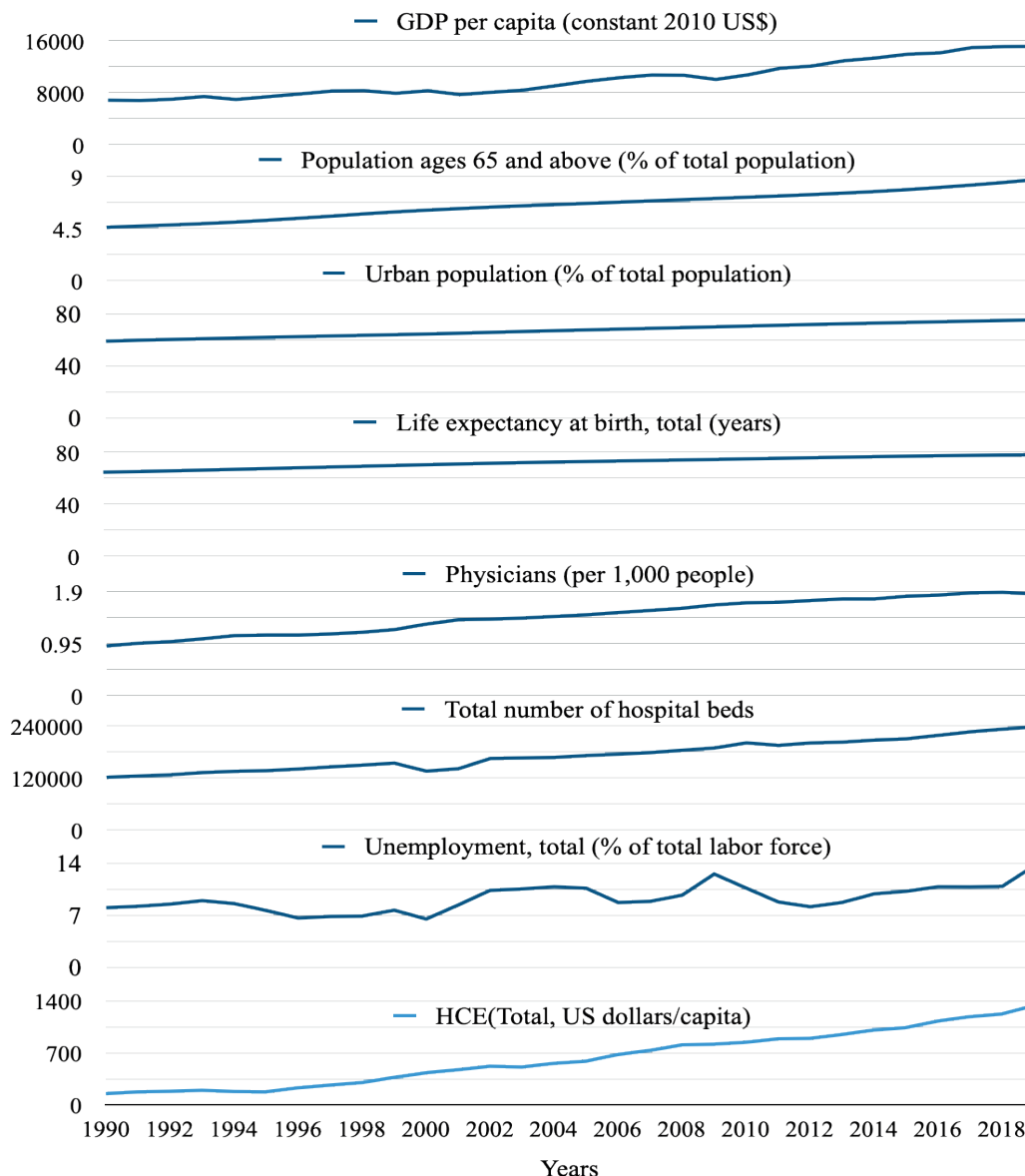


Figure 3. Change of inputs and output between 1990-2019

Preprocessing data for the machine learning algorithms is necessary and time-consuming. Normalizing data aims to transform variable values into a standard scale without disturbing the differences in value ranges. Normalization is only required when features have different ranges. In this case, since the variables' units are different, bringing the data into a standard scale will return more accurate results. Normalizing variables in the range of 0.05- 0.95 will give better results to make more meaningful comparisons in big data applications (Wang et al., 2015). The data set is divided into two, 70% used in the training stage and 30% in the test stage.

2.2. Support Vector Regression (SVR)

Support Vector Machine (SVM) is a kernel-based machine learning tool to address classification and regression problems. Although the support vector algorithm was initially employed for classification, it can be used in regression problems. The ability of this method to be applied for two purposes provides the solution to data science. Support vector regression (SVR) differs from SVM by having a single reference point. SVR does not scan to optimize the distance between two or more reference points to find the optimal hyperplane. The model can be used for linear and nonlinear problems. A linear or nonlinear range can be obtained when applying the SVR model according to the selected kernel function. It tries to find a regression function for the used data set. The SVR architecture is formulated as follows, as seen in Eq. (1) (Quan et al., 2020).

$$\begin{aligned} & \text{maximize} \begin{cases} \frac{1}{2} \sum_{i=1}^l (\omega_i - \omega_i^*) (\omega_i - \omega_i^*) K(x_i, x_j) \\ -\epsilon \sum_{i=1}^l (\omega_i + \omega_i^*) + \sum_{i=1}^l y_i (\omega_i - \omega_i^*) \end{cases} \\ & \text{s. t} \begin{cases} \sum_{i=1}^l (\omega_i - \omega_i^*) = 0 \text{ and } \omega_i, \omega_i^* \in [0, C] \\ 0 \leq \omega_i, \omega_i^* \leq \frac{C}{l} \\ i = 1, 2, \dots, l \end{cases} \end{aligned} \quad (1)$$

In Eq. (1), ω_i and ω_i^* are nonnegative multipliers for each observation. x_i is current data, l is data size, the penalty coefficient is depicted C , ϵ is the penalty dimension, and the kernel function is presented $K(x_i, x_j)$. Adjust the $\omega = [\omega_1, \omega_1^*, \dots, \omega_l, \omega_l^*]^T$ to get the optimum solution. The regression equation is formulated in Eq. (2) (Quan et al., 2020).

$$f(x) = \sum_{i=1}^l (\omega_i - \omega_i^*) K(x_i - x_j) + b^* \quad (2)$$

2.3. Gaussian Process Regression (GPR)

Gaussian process regression (GPR) is a probabilistic model based on kernels. The inputs (x_i) and the outputs (y_i) are taken from an unknown distribution. A GPR model searches for a solution for predicting the output variable via the input vector. As seen in Eq. (x), the form is based on linear regression (Zhang et al., 2016).

$$y = x^T \beta + \varepsilon$$

Where $\varepsilon \sim N(0, \sigma^2)$ and the error variance and β are predicted using the training data, considering the Gaussian process, $p(\mathbf{f})$ is zero; a matrix K depicts a kernel function.

$$p(\mathbf{f}) = N(\mathbf{f} | \mathbf{0}, K) \quad (3)$$

Let $K_{ij} = K(x_i, x_j)$. Eq. (4) shows the function of y .

$$p(\mathbf{y}) = \int p(\mathbf{y} | \mathbf{f}) p(\mathbf{f}) d\mathbf{f} = N(\mathbf{f} | \mathbf{0}, K_y) \quad (4)$$

There is a hidden variable $f(x_i)$ calculated for each x_i in the GPR model. Let $\mathbf{o}_* = [\mathbf{o}(\mathbf{x}_*, \mathbf{x}_1), \dots, \mathbf{o}(\mathbf{x}_*, \mathbf{x}_M)]^T$ and $\mathbf{o}_{**} = k(\mathbf{x}_*, \mathbf{x}_*) \mathbf{o}_{**} = k(\mathbf{x}_*, \mathbf{x}_*)$. Eq. (5) and Eq. (6) show the mean and variance of $P(y_i | y)$.

$$\mu(\mathbf{x}_*) = \mathbf{o}_*^T K_y^{-1} \mathbf{y} \quad (5)$$

$$\sigma^2(\mathbf{x}_*) = \mathbf{o}_{**} - \mathbf{o}_*^T K_y^{-1} \mathbf{o}_* + \sigma_n^2 \quad (6)$$

2.4. Decision Tree (DT)

The Decision Tree algorithm is frequently used in machine learning literature and the business environment to provide solutions to both classification and regression problems. The decision tree obtains the predictive value by asking a series of questions to the data until each model reaches a single predictive value. The model determines the order and content of the problem. During the training phase, the model is equipped with a historical data set of the relevant field. Thus, the decision tree model learns the relationship between input data and target value. After the training phase, the decision tree builds a tree and calculates the best questions, and nodes are sequenced to make the most accurate predictions. The developed model can predict by providing the same data format as in the training phase. In the regression model, Let $X = X_1, X_2, \dots, X_n$ be independent variables. Y_1, Y_2, \dots, Y_n is defined as dependent variables, and n is the number of the observed values. The candidate split can represent $\varkappa = (f, th)$, and node depicts t , feature, and threshold value present f and th , respectively. Eq. (7) presents the dividing data into \varkappa candidate, and the DT left side node is gained (Pekel et al., 2020).

$$Q_l(\mathfrak{x}) = (x, y) | x_f \leq th_t \tag{7}$$

Eq. (7) depicts the dividing data into \mathfrak{x} candidate, and the right side of the DT is found.

$$Q_r(\mathfrak{x}) = (x, y) | x_f > th_t \tag{8}$$

Also, Eq. (8) can be written as $Q_r(\mathfrak{x}) = Q/Q_l(\mathfrak{x})$. The predictive value is estimated based on the historical data.

2.5. Performance evaluation criteria

The model’s accuracy was measured using certain metrics - root mean square error (RMSE), the coefficient of determination (R^2), and mean absolute error (MAE). The comparison of these metrics is necessary to make relevant interpretations. The performance criteria equations are given in Eq. (9) to Eq. (11), respectively (Guleryuz and Ozden, 2020).

$$MAE = \frac{1}{n} \sum_{t=1}^n |HS_i^{observed} - HS_i^{predicted}| \tag{9}$$

$$RMSE = \sqrt{\frac{1}{n} \sum_{t=1}^n (HS_i^{observed} - HS_i^{predicted})^2} \tag{10}$$

$$R^2 = \left(\frac{\sum_{i=1}^n (HS_i^{observed} - \overline{HS_i^{observed}}) (HS_i^{predicted} - \overline{HS_i^{predicted}})}{\sqrt{\sum_{i=1}^n (HS_i^{observed} - \overline{HS_i^{observed}})^2 (HS_i^{predicted} - \overline{HS_i^{predicted}})^2}} \right)^2 \tag{11}$$

n represents the number of observations, $HS_i^{observed}$ is the observed value of HS and $HS_i^{predicted}$ is the estimated value of HS at the time i .

3. RESULTS

3.1. Results of SVR, GPR, and DT

The SVR model was established to define the relationship between seven inputs and the health spendings variable. The SVR model includes hyperparameters, parameters which affect the predictive performance of the model. The developer usually determines these values. There are studies on hyperparameter optimization in the literature. However, since three different models were compared in this study, hyperparameter optimization of each model was not targetted. The developed model was tested with different hyperparameters, and the parameters that led to the best prediction performance were selected. Kernel function, box constraint, kernel scale, and epsilon are hyperparameters of the SVR model. Table 3 includes the parameters of the SVR model.

The Gaussian process is a nonparametric Bayesian method that can model complex relationships by handling uncertainty. Like SVR, a dataset that includes seven inputs and one output is used to develop a regression model. In this case, Matern 3/2 was employed as a kernel function among many kernels since its predictive ability is better than others. Other parameters of the GPR model can be seen in Table 3.

The decision tree is an iterative algorithm that uses a tree structure that starts using a single node and splits branches to find the current value. The input values move in a specific way considering the nodes. In this case, the decision tree starts with a single node, then the fitness function is found, and the split is chosen with the best fitness function value. This step is repeated until it reaches the stopping criteria. Table 3 shows the parameters of the developed DT model. Also, 3-fold cross-validation was employed in the training phase to impede overfitting.

Linear, Polynomial, Gaussian, and Sigmoid kernel functions were used in the developed SVR models. SVR with Gaussian kernel function had the best accurate prediction result. Also, in the GPR model, various kernel functions such as Linear, Polynomial, Squared Exponential, Rational Quadratic, Matern 3/2, and Matern 5/2 were used. The model with the lowest prediction error among the developed GPR models was GPR with Matern 3/2 kernel function. The parameters of these models are given in Table 3.

Table 3

The parameters of the developed models

SVR		GPR		DT	
Kernel function	: Gaussian	Kernel function	: Matern3/2	Minimum leaf size	: 1
Box constraints (C)	: 750.8097	Sigma	: 0.010018	Prune criterion	: MSE
Epsilon, ϵ	: 0.00025521	Kernel scale	: 0.15866	Minimum parent	: 10
Kernel scale	: 29.917	Prediction speed (obs/sec)	: 2700	Number of nodes	: 7
Prediction speed (obs/sec)	: 2900	Training time (sec)	: 67.691	Prediction speed (obs/sec)	: 2900
Training time (sec)	: 85.414			Training time (sec)	: 26.236

3.2. Accuracy evaluations of the developed models

In this paper, the HS was a target value, and seven indicators related to HS were inputs for all developed models. The regression functions that establish the relationship between inputs and output were obtained using the training data set. The developed SVR, GPR, and DT models were tested using the test dataset to measure the predictability. The estimation results' performances were compared according to three metrics: MAE, RMSE, and R^2 .

The performance metrics values of all models are given in Table 4. When the dataset is small, there is much debate over whether the coefficient of determination is a valid performance measure for tree-based models. In this way, developed models are ranked according to their RMSE values. However, even if the models are ranked according to R^2 values, the developed SVR model has the best predictive performance during the test phase.

Table 4

Performance metrics of the developed models

	Models	MAE	R2	RMSE	Rank
Training Phase	SVR	0.0097	0.9929	0.0151	2
	GPR	0.0049	0.9989	0.0060	1
	DT	0.0292	0.9611	0.0350	3
Testing Phase	SVR	0.1452	0.9536	0.1520	1
	GPR	0.1802	0.8944	0.2236	2
	DT	0.2256	0.1166	0.2519	3

All models successfully estimated health spendings using determined indicators. The R^2 values of these three models changed between 0.9611 and 0.9929 during the training phase. This situation confirms that the developed models had good predictive performance. The predictive performance of the models depends on the relationships between inputs and outputs during the training phase. If this relationship is modeled correctly, the machine learning algorithm will give satisfactory results during the test phase.

For this reason, 30% of the data was reserved as the test data. Table 4 includes the prediction ability of the developed models in the testing phase. The R^2 value of SVR, GPR, and DT were 0.9929, 0.9989, and 0.9611 in the training phase, with 0.9536, 0.8944, and 0.1166 in the testing phase. Machine learning models establish the relationship between inputs and outputs in the training phase. Therefore, since this established relationship was evaluated during the test phase, the testing phase's performance metrics should be compared. Table 4 shows that the developed SVR model has stronger prediction ability than the DT and GPR models. The developed SVR model found the R^2 and RMSE values to be 0.9536 and 0.1520, respectively, in the testing phase. The computational results indicate that the SVR can use to predict HS, while meeting the least estimation error. Figure 4 shows the comparison of developed models based on MAE, RMSE, and R^2 .

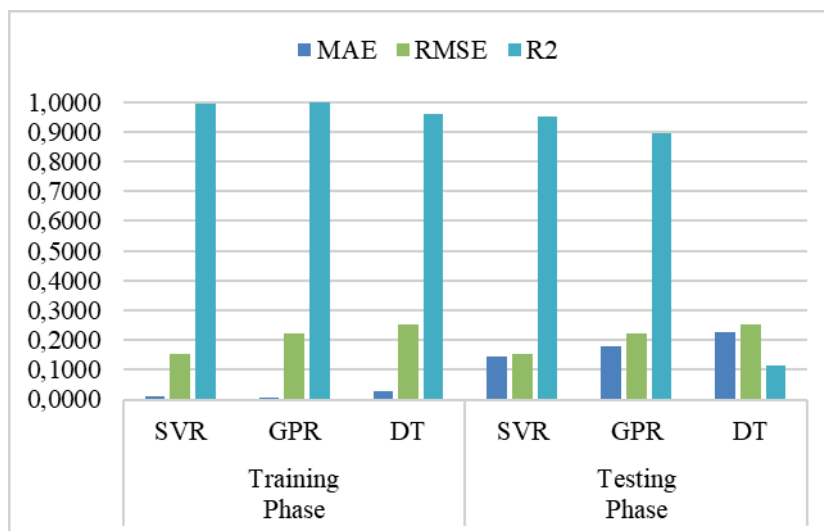


Figure 4. The comparison of developed models

4. DISCUSSION

Prosperous countries have a healthier population, mainly because there is a positive relationship between health services and economic power (Boyce and Brown, 2019). Since the problems related to countries’ health systems differ according to the country, the main reasons that affect the health systems of many countries, such as the aging population and chronic diseases, should be monitored carefully. Thus, problems that may arise in the health system can be easily identified. Data science and machine learning techniques play an active role in solving these problems and can support decision science.

For countries that want to provide sustainable health services, it is essential to measure and compare the performance of health processes reliably. Thus, it is crucial to obtain data on time to carry out the measurement and comparison process effectively. This can help to provide high-quality services at optimum costs continuously and responsively (WHO, 2018). This whole process requires more comprehensive analysis with new data. This situation opens the door to the use of machine learning and deep learning models. In order to manage expenditures effectively in health systems, the indicators affecting the system should be determined. Especially if a prediction model is developed to support the decision-maker, the variables and the relationships between these variables affect the model considerably. Linear and nonlinear relationships between data can be modeled with machine learning methods. In addition, higher accuracy prediction results can be obtained by using less data than statistical methods with machine learning methods.

This study was conducted to model the effects of selected indicators, such as population ages 65 and above, life expectancy, unemployment rate, the number of physicians, the total number of hospital beds, and urban population on health spending via ML-based models. Three different ML methods were developed, and the developed SVR model outperformed GPR and DT. Various models in the literature estimate health expenditures using various combinations of selected indicators with different models. High accuracy was obtained via developed models in this study, and this result is in accordance with Jakovljevic et al. (2020), Cinaroglu (2018), Aliyu and Amadu (2017), Häkkinen et al. (2008), and Anderson and Frogner (2008).

5. CONCLUSIONS

The Health economy has become an essential field in recent years. In modern economies, health spendings occupy a central place due to social policy. Health problems are considerably different for different countries. Thus, would be complicated to tell countries how much they will spend on health services without knowing the problems faced by each of them. Also, the indicators affecting the health expenditures of the countries should be appropriately determined. However, many different administrative and logistic activities and sustainability indicators can change health spending efficiency. The availability,

advantages, and disadvantages of each continuous improvement method differ by country. Health spendings are related to certain social, economic, and environmental indicators. In this study, prediction models have been developed with machine learning methods by selecting some of the indicators which are related to health expenditures.

Since GPR has not been used previously to model health spendings, this study contributes to the literature by developing forecasting models. Machine learning-based prediction models were designed with determined indicators via Support Vector Regression, Gaussian Process Regression, and Decision Tree architectures. The developed models were tested with test data, and their prediction performances were compared. The developed SVR model's prediction ability was better than GPR and DT. The MAE, RMSE, and R^2 values of the SVR model were 0.1452, 0.1520, and 0.9536, respectively, in the testing phase. Policymakers can benefit from these results to make healthcare spending plans. The study showed that the proposed SVR model reduced the RMSE value by 32.02% and 39.66% compared to the GPR and DT models, respectively.

The choice of determinants as input is important for accurate estimation results. Although this is the main limitation of the employed methods, good predictive performance has been achieved with this study's selected inputs. This limitation could be eliminated in future studies by combining the employed models with feature selection algorithms such as the genetic algorithm, grey wolf optimization, and particle swarm optimization. Also, the hyperparameters of the models could be optimized by heuristic or mathematical optimization methods.

Peer-review: Externally peer-reviewed.

Conflict of Interest: The author has no conflict of interest to declare.

Grant Support: The author declared that this study has received no financial support.

Hakem Değerlendirmesi: Dış bağımsız.

Çıkar Çatışması: Yazar çıkar çatışması bildirmemiştir.

Finansal Destek: Yazar bu çalışma için finansal destek almadığını beyan etmiştir.

References/Kaynaklar

- Akca, N., Sonmez, S., Yılmaz, A., 2017. Determinants of health expenditure in OECD countries: A decision tree model. *Pakistan J. Med. Sci.* 33, 1490–1494. <https://doi.org/10.12669/pjms.336.13300>
- Aliyu, A.A., Amadu, L., 2017. Urbanization, cities, and health: The challenges to Nigeria - A review. *Ann. Afr. Med.* 16, 149–158. https://doi.org/10.4103/aam.aam_1_17
- Anderson, G.F., Frogner, B.K., 2008. Health spending in OECD countries: Obtaining value per dollar. *Health Aff.* 27, 1718–1727. <https://doi.org/10.1377/hlthaff.27.6.1718>
- Bauer, G., Davies, J.K., Pelikan, J., Consortium, on behalf of the E.T.W.G. and T.E., 2006. The EUHPID Health Development Model for the classification of public health indicators. *Health Promot. Int.* 21, 153–159. <https://doi.org/10.1093/heapro/dak002>
- Biadgilign, S., Ayenew, H.Y., Shumetie, A., Chitekwe, S., Tolla, A., Haile, D., Gebreyesus, S.H., Deribew, A., Gebre, B., 2019. Good governance, public health expenditures, urbanization and child undernutrition Nexus in Ethiopia: An ecological analysis. *BMC Health Serv. Res.* 19, 1–10. <https://doi.org/10.1186/s12913-018-3822-2>
- Boyce, T., Brown, C., 2019. Economic and social impacts and benefits of health systems. *World Heal. Organ.* 56.
- Ceylan, Z., 2020. Investigation the insights between health expenditures and air quality. *Int. J. Glob. Warm.* 20, 203–215. <https://doi.org/10.1504/IJGW.2020.106594>
- Ceylan, Z., Atalan, A., 2020. Estimation of healthcare expenditure per capita of Turkey using artificial intelligence techniques with genetic algorithm-based feature selection. *J. Forecast.* 279–290. <https://doi.org/10.1002/for.2747>
- Cinaroglu, S., 2020. Modelling unbalanced catastrophic health expenditure data by using machine-learning methods. *Intell. Syst. Accounting, Financ. Manag.* 27, 168–181. <https://doi.org/10.1002/isaf.1483>
- Cinaroglu, Songül, 2017. Sağlık Harcamasının Tahmininde Makine Öğrenmesi Regresyon Yöntemlerinin Karşılaştırılması. *Uludağ Univ. J. Fac. Eng.* 22, 179–200. <https://doi.org/10.17482/uumfd.338805>
- Cinaroglu, Songul, 2017. International Journal of Economics and Financial Issues A Fresh Look at Out-of-Pocket Health Expenditures after More than a Decade Health Reform Experience in Turkey: A Data Mining Application. *Int. J. Econ. Financ. Issues* 7, 33–40.
- Crémieux, P.Y., Ouellette, P., Pilon, C., 1999. Health care spending as determinants of health outcomes. *Health Econ.* 8, 627–639. [https://doi.org/10.1002/\(SICI\)1099-1050\(199911\)8:7<627::AID-HEC474>3.0.CO;2-8](https://doi.org/10.1002/(SICI)1099-1050(199911)8:7<627::AID-HEC474>3.0.CO;2-8)
- Dalbokova, D., Krzyzanowski, M., 2002. Environmental health indicators: Development of a methodology for the WHO European Region. *Stat. J. UN. Econ. Comm. Eur.* 19, 93–103. <https://doi.org/10.3233/SJU-2002-191-209>
- Di Matteo, L., 2005. The macro determinants of health expenditure in the United States and Canada: Assessing the impact of income, age distribution and time. *Health Policy (New York)*. 71, 23–42. <https://doi.org/10.1016/j.healthpol.2004.05.007>

- Dritsakis, N., Klazoglou, P., 2019. Time Series Analysis using ARIMA Models: An Approach to Forecasting Health Expenditure in USA. *Econ. Internazionale / Int. Econ.* 72, 77–106.
- Durmaz, A., Pabuçcu, H., 2020. Kamu Sağlık Harcamaları - İşgücü Verimliliği İlişkisi: Sinirsel Bulanık Yaklaşım 8, 1375–1382.
- Efe, B., Efe, O.F., 2016. An Application of Value Analysis for Lean Healthcare Management in an Emergency Department. *Int. J. Comput. Intell. Syst.* 9, 689–697. <https://doi.org/https://doi.org/10.1080/18756891.2016.1204117>
- Guleryuz, D., Ozden, E., 2020. The Prediction of Brent Crude Oil Trend Using LSTM and Facebook Prophet. *Eur. J. Sci. Technol.* 1–9. <https://doi.org/10.31590/ejosat.759302>
- Häkkinen, U., Martikainen, P., Noro, A., Nihtilä, E., Peltola, M., 2008. Aging, health expenditure, proximity to death, and income in Finland. *Heal. Econ. Policy Law* 3, 165–195. <https://doi.org/10.1017/S174413310800443X>
- Jakovljevic, M., Timofeyev, Y., Ranabhat, C.L., Fernandes, P.O., Teixeira, J.P., Rancic, N., Reshetnikov, V., 2020b. Real GDP growth rates and healthcare spending - Comparison between the G7 and the EM7 countries. *Global. Health* 16, 1–13. <https://doi.org/10.1186/s12992-020-00590-3>
- Jia, H., Jiang, H., Yu, J., Zhang, J., Cao, P., Yu, X., 2021. Total Health Expenditure and Its Driving Factors in China: A Gray Theory Analysis. *Healthcare* 9. <https://doi.org/10.3390/healthcare9020207>
- Maksimović, G., Jović, S., Jovanović, R., Aničić, O., 2017. Management of health care expenditure by soft computing methodology. *Phys. A Stat. Mech. its Appl.* 465, 370–373. <https://doi.org/10.1016/j.physa.2016.08.035>
- Martín, J.J.M., del Amo González, M.P.L., García, M.D.C., 2011. Review of the literature on the determinants of healthcare expenditure. *Appl. Econ.* 43, 19–46. <https://doi.org/10.1080/00036841003689754>
- Ministry of Development, 2014. Onuncu Kalkınma Planı 2014-2018 Sağlık Hizmetlerinin Etkinliğinin Artırılması ve Mali Sürdürülebilirlik Özel İhtisas Komisyonu Raporu 2023 (In Turkish)
- Mladenović, I., Milovančević, M., Sokolov Mladenović, S., Marjanović, V., Petković, B., 2016. Analyzing and management of health care expenditure and gross domestic product (GDP) growth rate by adaptive neuro-fuzzy technique. *Comput. Human Behav.* 64, 524–530. <https://doi.org/10.1016/j.chb.2016.07.052>
- OECD, 2021. Health Resources (indicators) [WWW Document]. URL <https://data.oecd.org/health.htm> (accessed 2.1.21).
- Özcan, T., Tüysüz, F., 2018. Healthcare expenditure prediction in Turkey by using genetic algorithm based grey forecasting models, *International Series in Operations Research and Management Science*. https://doi.org/10.1007/978-3-319-65455-3_7
- Papancolas, I., Woskie, L.R., Orlander, D., John Orav, E., Jha, A.K., 2019. The relationship between health spending and social spending in high-income countries: How does the US compare? *Health Aff.* 38, 1–9. <https://doi.org/10.1377/hlthaff.2018.05187>
- Pekel, E., Akkoyunlu, M.C., Akkoyunlu, M.T., Pusat, S., 2020. Decision tree regression model to predict low-rank coal moisture content during convective drying process. *Int. J. Coal Prep. Util.* 40, 505–512. <https://doi.org/10.1080/19392699.2020.1737527>
- Quan, Q., Hao, Z., Xifeng, H., Jingchun, L., 2020. Research on water temperature prediction based on improved support vector regression. *Neural Comput. Appl.* 4. <https://doi.org/10.1007/s00521-020-04836-4>
- Ram, M., Akbar Haghdoost, A., Hossein Mehroliassani, M., Abolhallaje, M., Dehnavieh, R., Najafi, B., Akbar Fazaeli, A., 2019. Forecasting health expenditures in Iran using the ARIMA.
- Rana, R.H., Alam, K., Gow, J., 2020. Health expenditure and gross domestic product: causality analysis by income level. *Int. J. Heal. Econ. Manag.* 20, 55–77. <https://doi.org/10.1007/s10754-019-09270-1>
- Samah, I.H.A., Rashid, I.M.A., Husain, WAFW, Ibrahim, S., Hamzah, H., Amlus, M.H., 2020. The impact of healthcare expenditure and healthcare sector growth on CO2 emission using dynamic panel data system GMM estimation model during COVID 19 crisis. *Int. J. Energy Econ. Policy* 10, 235–241. <https://doi.org/10.32479/ijeep.9769>
- Shi, L., Smit, E., Luck, J., 2020. Panel Survey Estimation of the Impact of Urbanization in China: Does Level of Urbanization Affect Healthcare Expenditure, Utilization or Healthcare Seeking Behavior? *Chinese Econ.* 0, 1–12. <https://doi.org/10.1080/10971475.2020.1848472>
- TURKSTAT, 2020. Turkish Statistical Institute [WWW Document]. URL <http://www.tuik.gov.tr/Start.do> (accessed 6.13.20).
- Wang, Y., Li, J., Gu, J., Zhou, Z., Wang, Z., 2015. Artificial neural networks for infectious diarrhea prediction using meteorological factors in Shanghai (China). *Appl. Soft Comput. J.* 35, 280–290. <https://doi.org/10.1016/j.asoc.2015.05.047>
- Worldbank, 2021. World Bank Open Data [WWW Document]. URL <https://data.worldbank.org/>
- WHO, 2018. Delivering quality health services, World Health Organization, World Bank Group, OECD.
- Zhang, C., Wei, H., Zhao, X., Liu, T., Zhang, K., 2016. A Gaussian process regression based hybrid approach for short-term wind speed prediction. *Energy Convers. Manag.* 126, 1084–1092. <https://doi.org/10.1016/j.enconman.2016.08.086>

Çocuklar İçin Otizm Spektrum Bozukluğunun Veri Madenciliği Yöntemleri ile Analizi*

Analysis of Autism Spectrum Disorder in Children with Data Mining Methods

Sümeyye Çelik¹ , Melike Şişeci Çeşmeli² 



*Bu çalışma "Çocuklar İçin Otizm Spektrum Bozukluğunun Veri Madenciliği Yöntemleri ile Analizi" başlıklı yüksek lisans tezinden üretilmiştir.

¹Adana Alparslan Türkeş Bilim ve Teknoloji Üniversitesi, İşletme Fakültesi, Yönetim Bilişim Sistemleri Bölümü, Adana, Türkiye

²Burdur Mehmet Akif Ersoy Üniversitesi, Bucak Teknoloji Fakültesi, Yazılım Mühendisliği Bölümü, Burdur, Türkiye

ORCID: S.Ç.0000-0003-3621-8653;
M.Ş.Ç. 0000-0001-9541-2590

Corresponding author:

Sümeyye ÇELİK
Adana Alparslan Türkeş Bilim ve Teknoloji Üniversitesi, İşletme Fakültesi, Yönetim Bilişim Sistemleri Bölümü, Adana, Türkiye
E-mail address: sumeyyecelik@atu.edu.tr

Submitted: 17.02.2021

Revision Requested: 06.04.2021

Last Revision Received: 08.04.2021

Accepted: 08.04.2021

Published Online: 29.06.2021

Citation: Çelik, S. ve Şişeci Çeşmeli, M. (2021). Çocuklar için otizm spektrum bozukluğunun veri madenciliği yöntemleri ile analizi. *Acta Infologica*, 5(1), 167-186. <https://doi.org/10.26650/acin.882187>

ÖZ

Veri madenciliği teknikleri, veriler arasında gizli kalmış olan örüntüleri ortaya çıkarmayı amaçlamaktadır. Bu kapsamda, tıp gibi birçok alanda yaygın bir biçimde kullanılmaktadır. Teşhis ve tedavisi oldukça zor ve uzun bir süreçten oluşan otizm spektrum bozukluğu doğuştan gelen ya da yaşamın ilk yıllarında ortaya çıkan karmaşık bir nöro-gelişimsel bozukluktur. Bu çalışmada 292 çocuktan toplanan gerçek ve güncel otizm spektrum bozukluğu verileri kullanılmıştır. Veri seti 20 girdi özniteliği ve 1 çıktı öznitelğine sahiptir. Çıktı özniteliği otizmin bulunup bulunmadığını ifade etmektedir. Çalışma da öncelikle veri seti üzerinde eksik verilerin tamamlanması, kategorik verilerin sayısallaştırılması, normalizasyon gibi veri ön işleme aşamaları gerçekleştirilmiştir. Devamında ise öznitelikler yapay sinir ağları ve dilsel kuvvetli sinir-bulanık sınıflayıcı ile sınıflandırılmış, k-means ve x-means ile kümelennmiştir. Her bir yöntemin sonuçları değerlendirilmiş ve performanslar karşılaştırılmıştır.

Anahtar kelimeler: Tıp Alanında Veri Madenciliği, Hastalık Teşhisi, Otizm Spektrum Bozukluğu, Sınıflandırma, Kümeleme

ABSTRACT

Data mining techniques aim to reveal hidden patterns in data. They are widely used in many fields, such as medicine. Autism spectrum disorder, whose diagnosis and treatment are difficult and lengthy, is a complex neurodevelopmental disorder that is congenital or occurs in the first years of life. Actual and current autism spectrum disorder data collected from 292 children were used in this study. The data set has 20 input attributes and 1 output attribute. The output attribute expresses whether autism is present or not. In the study, data pre-processing stages, such as completing missing data on the data set, digitizing categorical data, and normalization, were first carried out. Subsequently, the features were classified by artificial neural networks and linguistic strength neuro-fuzzy classifier and clustered with k-means and x-means. The results of each method were evaluated and the performances were compared.

Keywords: Data Mining in Medicine, Diagnosis of Disease, Autism Spectrum Disorder, Classification, Clustering

1. GİRİŞ

Veri madenciliği, bilgiyi keşfetme yolculuğudur (Hong ve Wu, 2011). Veri madenciliği ile var olan ama daha keşfedilmemiş örüntüler aranmaktadır. Yani veri madenciliği; ham halde iken sadece yığın şeklinde duran ve bir anlam ifade etmeyen veriler arasından, amaca uygun olarak anlamlı ifadeler oluşturulmasına olanak tanıyan yöntemler bütünüdür. Veri madenciliğinin en temel amacına bakıldığında zaman cevher madenciliği ile aynı olduğu görülmektedir (Diler, 2016).

Teknolojik gelişmeler her alanda olduğu gibi tıp alanında da oldukça kullanılmaktadır (Moskov, 2016). Günümüzde gelişen bu teknoloji ile birlikte hastanelerde kullanılan cihazlar ve uygulanan tedavi yöntemleri de geliştirilmektedir. Donanımsal ve yazılımsal olarak sürekli gelişen bu yöntemler sayesinde teşhis ve tedavilerin daha profesyonel şekilde uygulanmasına olanak tanınmaktadır (Tsumoto vd., 2017).

Tıp alanında söz konusu insan sağlığı olduğu için Ar-Ge çalışmalarının olabildiğince fazla olması ve desteklenmesi önemlidir (Reddy vd., 2012). Birçok tedavide erken teşhis büyük bir öneme sahiptir. Tedaviye geç kalındığında zaman hastalık daha fazla ilerleyebilmekte ve tedavisi daha zor olabilmektedir. Hatta bazı durumlarda geri dönüşümü mümkün olmayacak kayıplara sebep olabilmektedir. Bu gibi sebeplerden dolayı teşhis ve tedavi noktasında doktora yardımcı olması için kullanılan veri madenciliği yöntemlerine literatürde sıkça rastlanmaktadır (Akdemir, 2016).

Tıp alanında yapılacak olan teknolojik çalışmalar birçok farklı disiplinin bir araya gelmesi ile oluşturulmaktadır. Çünkü hastalık ve teşhis noktasında uzman olan bir ya da birkaç hekim olmadan hastadan gerekli bilgiler alınamayacak ya da muayene sonucunda elde edilen bulgular değerlendirilemeyecektir. Hastalık ile ilgili olarak elde edilen bulguların değerlendirmesini alanında uzman olmayan kişilerin yapması mümkün değildir. Her bir hastalık ayrı uzmanlık gerektirir ve hastalığın kişide olup olmadığının tespit edilmesi için spesifik detayların uzman tarafından biliniyor olması gereklidir. Aksi takdirde teşhis yanlış konulabilmekte ve tedavi aksayabilmektedir (Budak ve Bozkurt, 2013). Bu durum hastanın daha ciddi sorunlar yaşamasına sebep olabilmektedir. Sonuç olarak teşhis konusunda bilişim teknolojileri kullanılarak yapılacak bir çalışmada, bilgilerin alanında uzman kişilerden alınması gereklidir. Kullanılan parametreler doğruluğu uzmanlar tarafından kabul görmüş hali ile toplanmış olmalıdır ve analizler daha sonra gerçekleştirilmelidir. Bu analizleri gerçekleştirmek için de bilişim sistemlerini kullanabilen, uzmanından aldığı verileri işleyerek anlamlı bilgilere dönüştürebilecek kişilere ihtiyaç duyulmaktadır. Sonuç olarak başarılı bir çalışma yürütmek için farklı disiplinlerin bir arada çalışması oldukça önemlidir (Nilashi, 2017).

Otizm, Türkiye de ve dünyada oldukça fazla görülen bir rahatsızlıktır (Kılınç ve Söğüt, 2018). Otizm spektrum bozukluğu, doğuştan olabilen ya da doğumdan sonra bir kaç yıl içerisinde ortaya çıkabilen karmaşık bir nöro gelişimsel bozukluktur (Torun ve İşeri, 2018). Özellikle çocukluk döneminde görülmektedir (Uğur vd., 2018). Beynin yapısını ya da işleyişini etkileyen bazı sinir sistemi sorunlarından kaynaklanmaktadır (Torun ve İşeri, 2018). Otizm spektrum bozukluğu gözlemlenen hastalarda sosyal ve bilişsel olarak gelişmelerin normalden daha yavaş olduğu görülmektedir (Uğur ve Göker, 2018). Örneğin çevrelerinden gelen uyarılara karşı tepkileri yeterli düzeyde değildir (Gülsöz ve Çıkkılı, 2018). Birçok hastalıkta olduğu gibi otizmde de erken tedavi ve eğitim çok önemlidir. Bunun içinde erken teşhis önem kazanmaktadır. Çünkü çocukların hayattan uzak kalmadan, topluma kazandırılmaları gerekmektedir (Akman, 2018). Literatürde otizm spektrum bozukluğu konusunda birçok çalışma yer almaktadır. Bu çalışmaya da Thabtah (2017) tarafından gerçekleştirilen bilimsel çalışma yön vermiştir. Yazar çalışmasında otizm bozukluğu ile ilgili veri madenciliği çalışmalarını değerlendirmiştir. Yapılan araştırma ile son dönemde otizm verileri ile gerçekleştirilen veri madenciliği çalışmalarında verilerin güncelliğini yitirdiği ve yeni ölçeklendirmelerle toplanan verilere ihtiyaç duyuldu vurgusu yapılmıştır (Thabtah, 2017). Devamında ise Thabtah ve arkadaşları (2018) tarafından gerçekleştirilen bilimsel çalışma ile güncel olarak oluşturulan ölçekten faydalanılarak, ihtiyaç duyulan yeni veriler toplanmış ve sonuçlar yayınlanmıştır. Tarama testi ASDTests isimli bir mobil uygulama ortamında gerçekleştirilmiştir. Uygulamanın içinde yaş gruplarına göre ayrı kategoriler mevcuttur. Her kategoride on soru bulunmaktadır ve her biri kullanıcıların uygun cevabı dikkatlice seçmelerini sağlamak için bir resimle ilişkilendirilmiştir. Katılımcılara, bilgilerinin isimsiz tutulacağı ve sadece araştırma amacıyla paylaşılacağı bildirilmiştir. Değerlendirme tamamlanmadan önce, katılımcılardan araştırmanın amacını, gizlilik politikasını ve verilerin kullanımını açıklayan bir bilgilendirme

yapılmıştır. Kullanılan veri setinde girdi olarak kullanılan öznelikler arasında otizm spektrum bozukluğu'nu ölçümleyen on sorunun yanında demografik özellikleri içeren bilgilerde mevcuttur. Çıktı parametresi ise teste giren bireyin otizm spektrum bozukluğu olup olmadığını belirtmek için kullanılan evet ya da hayır anlamı taşıyan sınıf değişkenidir. (Thabtah vd., 2018).

Otizm spektrum bozukluğu için Thabtah ve arkadaşların oluşturduğu güncel verilerin çocuklar için olan alt kümesi kullanılarak yapılan çalışmalar incelendiği zaman logistic regression, naive bayes gibi yöntemlerin kullanılması ile gerçekleştirilen sınıflandırma işlemleri olmakla birlikte yapay sinir ağları ve dilsel kuvvetli sinir-bulanık sınıflayıcı ile yapılan bir sınıflandırmaya ve herhangi bir yöntem ile kümelemeye rastlanmamıştır. Bu sebepten dolayı bu çalışmada öncelikle eksik verilerin tamamlanması, normalizasyon işleminin gerçekleştirilmesi gibi veri ön işleme yöntemleri uygulanmış devamında ise bu konuda literatürde eksikliği görülen kümeleme işleminin gerçekleştirilmesi ve sınıflandırma işlemi için denenmediği fark edilen iki yöntemin veri seti üzerinde uygulanması sağlanmıştır. Literatüre katkısı olması açısından çalışmada denenilen yöntemlerin tahminliliğinin doğruluğu Doğruluk, Duyarlılık, Belirleyicilik, F-Ölçütü gibi birçok parametre ile verilmiş ve yöntemlerin başarı oranları karşılaştırılmıştır. Kümeleme işlemi için k-means ve x-means yöntemleri, sınıflandırma işlemi için ise yapay sinir ağları ve dilsel kuvvetli sinir-bulanık sınıflayıcı yöntemleri kullanılmıştır. Sonuçlara bakıldığı zaman tahminleme doğruluğu bakımından otizm spektrum bozukluğu'nun çocuklar için olan verilerinde, sınıflandırma yöntemlerinin kümeleme yöntemlerinden daha başarılı olduğu görülmektedir. Özellikle dilsel kuvvetli sinir-bulanık sınıflayıcı yönteminin tüm verileri doğru sınıflandırarak literatürdeki birçok yöntemden daha yüksek başarı oranına sahip olduğu için tercih edilebilecek yöntemler arasında en iyi yöntemlerden bir tanesi olabileceği sonucuna varılmaktadır.

2. KAVRAMSAL ÇERÇEVE

2.1. Veri Madenciliği

Veri madenciliğinin literatürde birçok farklı tanımının olmasının yanında, en temel anlamına bakıldığında elde edilen verilerin veri madenciliği yöntemleri ile analiz edilerek anlamlandırılmaya çalışıldığı süreçtir. Veri madenciliği adından da anlaşıldığı üzere bir madenciliktir. Madencilik işinin çeşitli yöntemler yardımı ile verilere uygulanmasını sağlamaktadır (Akdemir, 2016). Böylece veriler üzerine uygulanan madencilik ile bilgi keşifleri gerçekleştirilmektedir. Yani var olan bilgiler, toplu halde duran ve ayrıştırılmamış olan veri yığınları içerisinde bulunarak keşfedilmektedir (Hong ve Wu, 2011).

2.2. Veri Madenciliğinin Tıp Alanında Kullanımı

Tıp alanında hastalık teşhisi yapılırken dikkate alınması gereken birçok kıstas vardır. Hastalığın ciddi olabileceği ve alınacak yanlış bir karar ile çok daha büyük problemlere yol açılabileceği göz önünde bulundurulduğunda teşhis konulurken hızlı ve doğru karar verilmesi oldukça önemlidir (Çelik, 2020). Bu noktada veri madenciliği yöntemlerinden faydalanmak gereklidir (Reddy vd., 2012). Tıp alanında, veri madenciliği hastalıklı dokuların tespit edilmesi ve tanı konulmasında ya da tedavi edilmesinde destek olacak sistemlerin tasarlanmasında kullanılabilir (Brossette vd., 1998).

2.3. Otizm Spektrum Bozukluğu

Otizm spektrum bozukluğu (OSB) dünya çapında oldukça hızlı büyüyen gelişimsel sakatlık tanısından biridir (Thabtah, 2017). OSB, doğuştan gelen ya da yaşamın ilk yıllarında ortaya çıkan karmaşık bir nöro-gelişimsel bozukluktur (Thabtah, 2018). Sosyal ilişkilerde, iletişim kurmada, öğrenmede ve konuşma yeteneklerinde sıkıntılar yaşanmasına sebep olmaktadır (De Campos Souza ve Guimaraes, 2018). OSB'de ilgi duyulan şeyler sınırlıdır ve bunun yanında kendini yineleyen davranışlarda ve hareketlerde bulunmaktadır (Akyol ve Karaci, 2018).

OSB'nin teşhis ve tedavisi oldukça uzundur. Aynı zamanda teşhisin erken konulması ve tedaviye başlanması çok önemlidir (Sabuncuoğlu vd., 2015). Bu hastalığın erken fark edilmesi ileride yaşanacak sağlık sorunlarını önemli ölçüde azaltabilmekte ve çocuğun genel zihinsel sağlığını iyileştirici yöntemlerin uygulanmasına imkân tanımaktadır (van den Bekerom, 2017). Teşhis için en güncel olarak Amerikan Psikiyatri Birliği'nin yayımladığı kılavuz (DSM-V) kullanılır (Olçay Gül ve Tekin İftar, 2012).

2.4. Otizm Spektrum Bozukluğu Konusunda Çocuklardan Toplanan Veri Seti ile Yapılan Çalışmalar

Literatür incelendiği zaman OSB konusunda veri madenciliği veya yapay zekâ tekniklerini içeren, OSB veri setinin farklı altkümelerini kullanan çok fazla çalışma olduğu görülmektedir. Bu sebepten dolayı literatürü sınırlandırabilmek için OSB veri setinin çocuklar için olan alt kümesini kullanan çalışmalara yer verilmiştir. Bu çalışmaların temelini topladığı verileri paylaşarak Thabtah (2018) oluşturmuştur.

Thabtah (2018) tarafından, detaylı olarak literatür değerlendirilmiş, daha önceki çalışmalara yer verilmiştir. Verilerin nasıl toplandığı anlatılmış ve uygulama aracı olarak kullanılan ASDTests tanıtılmıştır. Veriler arasındaki ilişki incelenerek grafikler oluşturulmuştur. ASD için oluşturulan veri seti için üç alt küme olarak belirlenen çocuk, ergen, yetişkin grupları için Logistic Regression ve Naïve Bayes algoritmaları ile analizler gerçekleştirilmiştir. Elde sonuçlara göre çocuklar için elde edilen doğruluk Logistic Regression için % 97.94, Naïve Bayes için ise % 92.80 olmuştur.

Akyol ve Karaci (2018), Thabtah'ın çalışması ile oluşturulan güncel veri setinin çocuklar için olan alt kümesini kullanarak öznitelik seçimi gerçekleştirilmiş ve ardından sınıflandırılma yapılmıştır. Çalışmada öncelikle ön işleme aşamaları gerçekleştirilmiş ve boş değerler içeren örnekler veri setinden atılmıştır. Böylece toplamda 292 örnekten oluşan veri seti 247 örneğe indirilmiştir. Veri setindeki kategorik veri türüne sahip öznitelikler sayısal değerlere çevrilmiştir. Bu ön işleme aşamasından sonra sınıflandırma gerçekleştirilmiştir. Sınıflandırma için Lojistik Regresyon ve Bulanık Kurallı Lojistik Regresyon kombinasyonu kullanılmıştır. Kullanılan yöntemlerin başarı oranına bakıldığı zaman Lojistik Regresyon %92.00, Bulanık Kurallı Lojistik Regresyon kombinasyonu %97.33 oranında doğruluğa sahiptir. Sonuç olarak bu çalışma için lojistik regresyonun bulanık kurallar ile birlikte çalıştırıldığında daha doğru bir tahminleme yaptığı gösterilmektedir (Akyol ve Karaci, 2018).

De Campos Souza ve Guimaraes (2018), Thabtah'ın çalışması ile oluşturulan güncel veri setinin çocuklar için olan alt kümesini kullanarak bir sınıflandırma gerçekleştirmiştir. Sınıflandırma işlemi gerçekleştirilmeden önce veriler ön işleme aşamasından geçirilmiş ve kategorik olan veri türleri sayısal ifadelerle dönüştürülmüştür. Yöntem olarak Bulanık Sinir Ağı Mimarisi kullanılmıştır ve bulanık kümelerin sayısının performansa etkisi değerlendirilmiştir. Modelin başarı oranını test etmek için veriler %70 eğitim %30 test verisi olarak ayrılmıştır. Sonuçlara bakıldığı zaman en düşük performansın bulanık küme sayısının 3 olarak kullanıldığı test verisinde %59 olarak, en yüksek performansın da bulanık küme sayısının 9 olarak kullanıldığı test ve eğitim setinde %100 olarak görülmüştür (De Campos Souza ve Guimaraes, 2018).

Hutchinson ve arkadaşları (2019), Thabtah'ın çalışması ile oluşturulan güncel veri setinin çocuklar için olan alt kümesini kullanarak sınıflandırma yapmışlardır. Sınıflandırmayı gerçekleştirmek için yöntem olarak Weka hazır paket programı kullanarak J48 Decision Tree ve Naive Bayes algoritmasını kullanmışlardır. İki yöntem ile elde edilen sonuçlar accuracy, sensitivity ve specificity değerleri olarak verilmiş ayrıca iki yöntemin başarıları karşılaştırılmıştır. Çalışmanın sonuçlarına bakıldığı zaman; J48 Decision Tree için accuracy değerinin %92.8, sensitivity değerinin 0.901 ve specificity değerinin 0.934 olduğu Naive Bayes için ise accuracy değerinin %96.2, sensitivity değerinin 0.972 ve specificity değerinin 0.954 olduğu görülmektedir. Bu sonuçlar kapsamında kullanılan veri setini Naive Bayes algoritmasının J48 Decision Tree algoritmasından daha doğru sınıflandırdığı görülmektedir (Hutchinson vd, 2019).

3. YÖNTEM

3.1. Sınıflandırma Yöntemi

3.1.1. Yapay sinir ağları

Yapay sinir ağları (YSA), insan beyninin çalışma şekli ilham alınarak oluşturulmuş bir algoritmadır (Deb vd., 2018). Temelinde nörofizyologların ve psikologların sinir ağlarının nasıl çalıştığı konusunda ortaya koyduğu biyolojik bulgular kullanılmaktadır (Küçükkoçoğlu vd., 2005). Bu biyolojik bulgular yapısal ve işlevsel olarak sistemleştirilerek matematiksel bir model oluşturulmaya çalışılmıştır. Oluşturulan bu modele sinir ağları modeli denilmiştir (Akkaya vd., 2009). Diğer bir

ifade ile YSA insan beyninin çalışma sistemini taklit ederek (Mehmet ve Akpınar, 2018), bilgisayarlı sistemlere karar verdirme becerisini kazandırmayı amaçlamaktadır (Kalantar vd., 2018). YSA, İnsan beyninin ezberleme, öğrenmeye, anlam çıkarma, bilgiler arasındaki ilişkileri ortaya koyma gibi birçok işlemini yapabilme amacıyla tasarlanmıştır. (Özsoy ve Fırat, 2004) (Akyılmaz ve Ayan, 2006). YSA ortadaki durumdan bir öğrenme gerçekleştirebildiği için yeni durumlara adapte olabileme yeteneğine sahiptir.

3.1.2. Dilsel Kuvvetli Sinir-Bulanık Sınıflayıcı

Dilsel kuvvetli sinir-bulanık sınıflayıcı (DKSBS), verileri sınıflandırmak için kullanılan yöntemlerden bir tanesidir. En temel özelliği sınıflandırma yapmadan önce özniteliklerin önem derecelerinin belirlenmesidir (Pençe vd, 2013). Özniteliklerin önem dereceleri belirlenirken bulanık çıkarsama yapılmaktadır. Böylece sınıflandırma sırasında, önem derecesi yüksek olan öznitelikler seçilirken, önem derecesi düşük olan öznitelikler devre dışı bırakılmış olmaktadır (Çetişli, 2006).

Klasik mantıkta ile bulanık mantığın en temel farkı; klasik mantıkta sınıflandırmaların kesin olması, bulanık mantıkta ise bir elemanın birden fazla kümeye ait olabilmesi durumudur. Bulanık mantıktaki elemanlar kümelere belirli oranlarda üye olabilmektedir. Yani elemanı olma durumunu 1, olmama durumunu 0 ile temsil edersek 0.3, 0.5 derecelerde elemanı olma durumu ortaya çıkabilmektedir. Bunun yanı sıra elemanlar tek bir kümede olmak zorunda da olmadığı için bir eleman 0.3 oranında bir kümeye, 0.5 oranında başka kümeye dahil olabilmektedir. Bu durum klasik kümelerdeki siyah beyaz şeklindeki keskin ayrımın ortadan kaldırılarak gri alanların meydana gelmesine olanak tanımaktadır. Böylece bulanık yaklaşım ile daha esnek davranılarak kısmi üyelikler oluşturulabilmektedir (Tiryaki ve Kazan, 2007).

Bulanık çıkarsama ile özniteliklerin önem dereceleri belirlendikten sonra sınıflandırma için YSA kullanılmaktadır. YSA eğitilirken, bulanık çıkarsama ile seçilen öznitelikler kullanıldığı için başarı oranı oldukça yüksektir (Çelik vd., 2018). Aynı zamanda kümeleri birbirinden ayırtırmayan öznitelikler kullanılmadığı için bellekten de tasarruf sağlanmaktadır. Bu durum hızın artmasının yanında, maliyetlerin de azalmasına yardımcı olmaktadır (Çetişli, 2006).

Sınıflandırma yapılırken bazı öznitelikler çıkarıldığında sınıflandırma başarısı düşmüyor ise o özniteliklerin sınıflandırmada kullanılmasına gerek yoktur (Huang ve Chow, 2005). Çünkü gerçek hayat problemlerinde de çözüme gidilirken oluşturulan alternatiflerin etki düzeyi farklı olabilmektedir. Önemli olan problemi çözüme götüreceği olan daha hızlı ve daha doğru alternatifleri bulmaktır (Liu vd., 2005). Bu açıdan bakıldığında zaman dilsel kuvvetli sinir-bulanık sınıflayıcıda da gerçek hayat problemlerinin çözümüne yakın bir çözüm yolu izlenmektedir. Veri madenciliği açısından izlenen yol ise iki aşamada gerçekleştirilmektedir. Öncelikle ön işleme aşaması olarak sayılan özniteliklerin önem derecesinin belirlenmesi işlemi gerçekleştirilmektedir. Daha sonra belirlenen bu önem düzeylerine göre yapay sinir ağı eğitilerek sınıflandırma işlemi gerçekleştirilmektedir (Çelik vd., 2018). Böylece dilsel kuvvetli sinir-bulanık sınıflayıcı'yı diğer sınıflandırma yöntemlerinden ayıran en önemli özelliği bulanık çıkarsama ile öznitelik tespiti gerçekleştirmesidir (Fırat, 2008).

Dilsel kuvvetli sinir-bulanık sınıflayıcı, özniteliklerin önem düzeyini hesaplarken bulanık kuralları kullanmaktadır. Öznitelikler, bulanık kurallar ile aşamalı olarak belirlenmektedir (Çağlar vd., 2010). Bulanıklık, sistemi evet hayır gibi kesin cevapların kısıtlarından kurtarmaktadır (Koçyiğit ve Korürek, 2010). Yani bir özneliğin sonuç üzerinde etkisinin olup olmadığı var ya da yok şeklinde kesin olarak nitelendirilmek yerine, “az var”, “var” yada “çok var” gibi aşamalı olarak nitelendirilmektedir. Böylece bulanık çıkarsamaların insan çıkarsamalarına benzerliği arttırılmış olmaktadır (Cetişli ve Kalkan, 2011).

3.2. Kümeleme Yöntemi

3.2.1. K-Means

Kümeleme yöntemi, bir dizi elemanı benzer gruplarda toplama işlemidir. Kümeleme yöntemi veri setlerini gruplara ayırmak ve verilerdeki önemli ve gizli kalmış bilgileri açığa çıkarmak için en çok kullanılan yöntemlerden bir tanesidir (Çınaroğlu ve Bulut 2018). Kümeleme yönteminin, diğer bir veri madenciliği yöntemi olan sınıflandırmadan farkı şudur; sınıflandırmada sınıflar önceden bellidir ama kümelemede sınıflar önceden belli değildir (Doğan vd., 2018). K-means en eski kümeleme algoritmalarından bir tanesidir (MacQueen, 1967) ve hiyerarşik olmayan kümeleme yöntemleri arasında yer almaktadır

(Yavuz vd., 2011). 1967 yılında J.B. MacQueen tarafından geliştirilen algoritma (Işık ve Çamurcu, 2007), yıllardır birçok alanda kullanılmaya devam etmektedir (Bradley vd., 1998). K-means'ın tercih edilmesinin en temel sebebi yapısının sadeliği yanında model başarı oranının oldukça yüksek olmasıdır. K-means, kümeleri eşit büyüklükte ve homojen olarak oluşturma eğilimindedir (Işık ve Çamurcu, 2011) ve her elemanı bir kümeye atamaktadır. Böylece her eleman bir kümeye ait olmakta ve kümeler birbirinden kesin olarak ayrılmaktadır. Kümelemenin kullandığı öğrenme tekniği gözetimsiz öğrenmedir (Sariman, 2011). Yani kümeleme yapılırken önceden belirlenmiş kıstaslar yoktur. Kümeler, kümeyi temsil eden merkez noktaların belirlenmesi fikrine dayalı olarak oluşturulmaktadır (Güner vd., 2018). Kaç kümenin olacağı k parametresi ile ifade edilmektedir (Khaled vd., 1998). K parametresi aynı zamanda kaç tane küme merkezinin olacağı bilgisini taşıdığı için, K-means algoritması kullanılırken k parametresinin önceden girilmesi gerekmektedir (Bilgin ve Çamurcu, 2005). Kümeler belirlenirken; grup elemanların grup içi benzerliklerin en fazla ve gruplar arası benzerliklerin en az olması hedeflenmektedir (Demiralay ve Çamurcu 2005). Benzer olan kümeler oluşturulurken veriler arasındaki uzaklıklara bakılmaktadır. Bunun için eldeki veri kaç kümeye ayrılmak isteniyor ise o kadar küme merkezi belirlenmektedir (k parametresi) ve her elemanın bu belirlenen küme merkezlerine olan uzaklıkları tek tek hesaplanmaktadır. Hesaplamama sonucu her bir veri, kendine en yakın olan küme merkezine dâhil edilmektedir. Oluşan yeni kümelerin, küme merkezleri yeniden hesaplanmaktadır ve her bir elemanın yeni küme merkezlerine olan uzaklıkları da yeniden hesaplanmaktadır. Bu işlem hiçbir elemanın dâhil olduğu küme değişmeyinceye kadar devam etmektedir. En sonunda hiçbir elemanın dahil olduğu küme merkezi değişmiyor ise o zaman hesaplama durdurulmaktadır ve her eleman en son bulunduğu kümeye atanarak algoritma bitirilmektedir (Tekin vd., 2011).

Algoritmada uzaklık hesaplamalarından sonra birbirine en yakın olan elemanlar en çok benzerlik gösteren elemanlardır ve aralarında benzer oldukları tespit edilen elemanlar bir kümede toplanmaktadır (Çalışkan ve Soğukpınar, 2008).

K-means'ın en büyük eksiği verilerin kaç kümeye ayrılacağını, yani k parametresinin kaç olacağını tespit edememesidir. Bunun için, eğer verilerin kaç kümeden oluştuğu biliniyorsa k parametresi olarak bilinen küme sayısı girilebilmekte ya da farklı k parametreleri girilerek en uygun olanı tespit edilebilmektedir.

3.2.2. X-Means

K-means, kümeleme algoritmaları arasında oldukça popülerdir. K-means'ın çok fazla tercih ediliyor olmasının sebebi kolay anlaşılabilir bir yapıya sahip olmasının yanında model başarı oranının oldukça yüksek olmasıdır. Diğer bir açıdan popüler olmasının yanında bazı eksikleri mevcuttur (Akçapınar vd., 2016). Bu eksiklerden en belirgin olanı; küme sayısını ifade eden k parametresinin kullanıcı tarafından sabit bir değer olarak sağlanması gerektiğidir. Küme sayısının sabit bir k değeri ile sınırlandırılması küme sayısı için daha iyi olabilecek diğer alternatiflerin göz ardı edilmesi anlamına gelmektedir (Pelleg ve Moore, 2000). X-means'te veriler bir kd-tree'ye yerleştirilerek her aşamadaki istatistiksel veriler depolanmaktadır. Depolanan istatistiksel veriler içinde, belirli bir bölge için göz önünde bulundurulması gereken merkezlerin listesi de tutulmaktadır (Pelleg ve Moore, 2000). Böylece tüm alternatifler değerlendirilerek en iyi olanın seçilmesi imkânı sağlanmış olmaktadır (Muhr ve Granitzer, 2009). X-means algoritması, K-means algoritmasının geliştirilmiş hali olarak (Tsai ve Chiu, 2008) 2000 yılında Pelleg ve Moore tarafından geliştirilmiş bir kümeleme (bölütleme) algoritmasıdır (Pelleg ve Moore, 2000). Temelinde K-means algoritmasının çalışma şeklini kullanmasının yanı sıra eksik görülen yönünün tamamlanması amacıyla oluşturulmuştur (Kalogeratos ve Likas, 2012). X-means ile K-means'in eksikliği olarak görülen küme sayısının algoritma tarafından belirlenememesi sorununa çözüm sunulmaktadır (Akçapınar vd., 2016). X-means'te küme sayısı için sabit bir değer belirtmek yerine makul olabilecek bir aralık belirtilmektedir. Böylece X-means belirtilen bu aralıktaki değerlerden (Steinley, 2006) en doğru oluşuna karar verdiği küme sayısını kendisi belirleyebilmektedir (Bholowalia ve Kumar, 2014). X-means yapısı içerisinde aşamalı olarak K-means algoritması çalıştırılmaktadır (Feng ve Hamerly, 2007). Alt kümelerin daha doğru oluşturulması için K-means her çalıştıktan sonra, hangi merkezlerin kendilerini bölmeleri gerektiği konusunda yerel kararlar oluşturularak ilerlenmektedir. Bu bölünme kararları bayesin bilgi kriteri (BIC) hesaplanarak verilmektedir (Jain, 2010; Hamerly ve Elkan, 2004). Her aşamada mevcut merkezlerin (ebeveyn) ve yeni oluşturulan yavru (çocuk) merkezlerin en iyi sonuçları blaclisting'e kaydedilmektedir. Kayıtlar tutulurken küme merkezlerinin yanında her k parametresi için model seçim kriterini puanlamak için kullanılan değerlerde tutulmaktadır (Pelleg ve Moore, 2000). Böylece

merkez konumlarının dikkatlice analiz edilmesine imkan sağlanmaktadır (Tucker vd., 2010). Temelde, algoritma verilen aralığın alt sınırına eşit k ile başlamakta ve üst sınıra ulaşılan kadar aşamalı olarak yeni merkezler eklemeye devam etmektedir. Bu işlemler sırasında, en iyi puanı elde eden centroid seti kaydedilmektedir. Bir sonraki işlemde elde olan puan, sistemde kayıtlı olan puandan daha yüksek ise yeni puan listeye eklenmektedir. Böylece liste sürekli olarak güncellenmektedir. Güncelleme prosedürü özyinelemelidir ve bir bölge içindeki noktalar için olası centroidlerin listesini tutmaktadır. Görevi, söz konusu bölgelerin merkez noktalarını uygun değerler ile güncellemektir. İlk olarak, girilen k parametre listesinin en küçük değerindeki sayı kadar rasgele oluşturulan merkez noktalarının yazılması ile başlamaktadır (Pelleg ve Moore, 2000). Yeni değerler hesaplandıktan sonra daha iyiler bulunduğça güncellenmektedir. En sonunda en yüksek değere sahip merkezler çıktı olarak kullanılmaktadır (Bholowalia ve Kumar, 2014).

3.3. Ön İşleme Teknikleri

Veri madenciliği yöntemleri ile analiz edilmek için toplanan veriler eksik olabilmektedir. Bu gibi durumlarda analizleri eksik veri ile gerçekleştirmek mümkündür. Aynı zaman da veri setinin kalitesini artırmak için eksik verilerin düzeltilmesi de tercih edilmektedir. Analizlerin eksiksiz veri üzerinde gerçekleştirilmesi kullanılan yöntemin başarı oranını arttırmada da oldukça önemli bir paya sahiptir. Eksik verinin düzeltilmesi veri ön işleme aşamasında yer almaktadır. Eksik veriyi düzeltmek için eksik verinin olduğu kayıtların çıkarılması (Sezgin ve Çelik, 2013), veya çeşitli yöntemlerle tamamlanması mümkündür (Aydın, 2007). Eksik olan verilerin tamamlanması için literatürde kabul görmüş birçok tamamlama yöntemi mevcuttur. Eksik değerler tamamlanırken dışardan genel bir sabit değer kullanılabilen, diğer verilerin özellikleri dikkate alınarak bir değer atanabilen (ortalama, mod ya da medyan gibi) (Aydın, 2007) ya da tahminlemeler sonucu bulunan değerler kullanılabilen (Regresyon analizi, Hot Deck, Naive Bayes İle Değer Atama, Son Gözlemi İleri Taşıma, Karar Ağaçlar, Beklenti Maksimizasyonu, Çoklu atama gibi) (Sezgin ve Çelik, 2013). Bu çalışmada eksik verilerin tamamlanması için diğer verilerin özellikleri dikkate alınarak bir değer ataması yapılmıştır. Veri seti içerisindeki her bir özneliğin frekansına bakılmış ve eksik olan veriler tamamlanırken frekansı en yüksek olan değer kullanılmıştır.

Günümüzde, veriler içinden anlam çıkarmak için veri madenciliği yöntemleri oldukça fazla kullanılmaktadır. Veri madenciliği ile bilginin keşfi süreci; veri seçimi ve toplanması, veri ön işleme, veri dönüştürme, veri madenciliği yöntemlerini uygulama ve en sonunda sonuçları değerlendirme şeklinde gerçekleşmektedir (Sebik ve Bülbül, 2018). Elde edilen verilerin kalitesi kullanılan yöntemin başarı oranını arttırmada çok büyük bir paya sahiptir (Oğuzlar, 2003). Verinin kalitesinin artırılması için kullanılan birçok veri ön işleme tekniği vardır (Mackinnon ve Glick, 1999). Veri setindeki eksik verilerin tamamlanması (Sezgin ve Çelik, 2013), gürültülü verilerin düzeltilmesi veya çıkarılması (Akgöbek ve Çakır, 2009), özneliklerin önem derecesinin belirlenmesi (Çetişli, 2006), bazı öznelikler üzerinde normalizasyon işleminin yapılması (Jayalakshmi ve Santhakumaran, 2011) kullanılan ön işleme tekniklerinden bazılarıdır. Bu çalışmada ön işleme tekniklerinden eksik verilerin tamamlanması ve normalizasyon kullanılmıştır. Literatürde kabul görmüş birçok normalizasyon çeşidi mevcuttur. Z-Score Normalizasyonu, Min-Max Normalizasyonu, Medyan Normalizasyonu, Sigmoid Normalizasyonu kabul gören normalizasyon yöntemlerinden bazılarıdır (Yavuz ve Deveci, 2012). Normalizasyon yöntemlerinin biri ya da birkaçı aynı anda kullanılabilir. Bu çalışmada Min-Max normalizasyon yöntemi kullanılmıştır.

4. UYGULAMA

4.1. Otizm Spektrum Bozukluğu Veri Seti

Bu çalışmada güncel ve gerçek OSB verilerinin çocuklar için olan alt kümesi kullanılmıştır. Verilerin ana kaynağı <http://fadifayez.com/autism-datasets/> internet sitesidir. Aynı zamanda Fadi Thabtah'ın 2017 yılında yaptığı bilimsel yayınlı birlikte UCI machine learning repository sitesinde paylaşılmış ve kullanmak isteyenler için erişime açılmıştır. Kullanılan veri setinin adı Autistic Spectrum Disorder Screening Data for Children Data Set'tir. Veri seti, OSB'nin teşhisi için literatürde kabul gören son parametrelere göre hazırlanan soruların cevapları ile oluşturulmuştur. Cevaplar ASDTests uygulaması ile toplanmıştır. Uygulamanın içinde yaş gruplarına göre ayrı kategoriler mevcuttur. Her kategoride on soru bulunmaktadır ve her biri kullanıcıların uygun cevabı dikkatlice seçmelerini sağlamak için bir resimle ilişkilendirilmiştir. Katılımcılara, bilgilerinin isimsiz tutulacağı ve sadece araştırma amacıyla paylaşılacağı bildirilmiştir. Değerlendirme tamamlamadan önce, katılımcılardan

araştırmanın amacını, gizlilik politikasını ve verilerin kullanımını açıklayan bir bilgilendirme yapılmıştır. Kullanılan veri setinde girdi olarak kullanılan öznitelikler arasında otizm spektrum bozukluğunu ölçümleyen on sorunun yanında demografik özellikleri içeren bilgilerde yer almaktadır. Toplamda 1100 kişiye uygulanan bu test, üç alt grup verilerini içermektedir. Gruplar çocuk, genç ve yetişkinler şeklindedir. Bu grupların oluşturulmasının sebebi ise OSB'nin teşhisi için sorulan bazı soruların gruplar arasında farklılık göstermesidir. Oluşturulan veri setinin çocuklar için olan alt kümesinde toplamda 292 örnek vardır ve bu çalışmada yalnızca çocuklar için olan küme kullanılmıştır. Veri seti sayısal ve kategorik veri türlerini içermektedir. Veri seti içerisindeki bazı örneklerde eksik veriler bulunmaktadır. Toplamda 21 öznitelikten oluşan veri setinde; 20 öznitelik bireylerin genel ve sağlık durumu ile ilgili bilgiler içerdiği için girdi olarak kullanılmış, 1 öznitelik ise bireyin otizmlili olma durumunu içerdiği için çıktı olarak kullanılmıştır. Girdi olarak kullanılan öznitelik parametrelerinden biri olan 'tarama yönteminin türü' bireyin yaş aralığının hangi grupta olduğunu ifade etmektedir. Yani otizm tarama testini gerçekleştiren bireyin çocuk, genç ya da yetişkin gruplarının hangisine dâhil olduğu bilgisini göstermektedir. Bu çalışmada sadece çocuklar için olan alt küme kullanılmıştır. Tüm veri seti çocuklara ait olduğu için grup bilgisine ihtiyaç duyulmamıştır. Bu sebepten dolayı veri seti içerisinde bu öznitelik çıkarılmıştır. Böylece girdi parametresinin sayısı 19' a düşürülmüştür. Özniteliklerin türü ve soruların detayları Tablo 1'de detaylı olarak verilmiştir.

Tablo 1

Veri setindeki özniteliklerin detayları

Öznitelik numarası	Öznitelik	Veri türü	Açıklama
1	Soru 1'in cevabı	Binary (0, 1)	Diğerleri duymadığında genellikle küçük sesler fark eder.
2	Soru 2'in cevabı	Binary (0, 1)	Genelde küçük detaylardan ziyade resmin tamamına odaklanır.
3	Soru 3'in cevabı	Binary (0, 1)	Bir sosyal grupta, birkaç farklı insanın konuşmasını kolayca takip edebilir.
4	Soru 4'in cevabı	Binary (0, 1)	Farklı etkinlikler arasında gidip gelmeyi kolay buluyor.
5	Soru 5'in cevabı	Binary (0, 1)	Adaylarla sohbete nasıl devam edeceğini bilmiyor.
6	Soru 6'in cevabı	Binary (0, 1)	Sosyal sohbetinde iyidir.
7	Soru 7'in cevabı	Binary (0, 1)	Bir hikaye okuduğunda, karakterin niyetlerini veya duygularını çözmekte zorlanır.
8	Soru 8'in cevabı	Binary (0, 1)	Okul öncesi eğitimindeyken, diğer çocuklarla oyun oynamayı sevmekten zevk alır.
9	Soru 9'in cevabı	Binary (0, 1)	Birinin ne düşündüğünü veya hissettiğini, sadece yüzlerine bakarak kolayca anlayabilir.
10	Soru 10'in cevabı	Binary (0, 1)	Yeni arkadaşlar edinmekte zorlanıyor.
11	Yaş	Number	Bireyin yıl olarak kaç yaşında olduğu bilgisi.
12	Cinsiyet	String	Bireyin erkek ya da kadın olduğu bilgisi.
13	Etnik köken	String	Bireyin hangi etnik kökene sahip olduğu bilgisi.
14	Sarılık ile doğmuş olmak	Boolean (evet-hayır)	Bireyin sarılık ile doğup doğmama bilgisi.
15	PDD'li aile üyesi	Boolean (evet-hayır)	Bireyin herhangi bir aile üyesinin PDD'sinin olup olmadığı bilgisi.
16	İkamet edilen ülke	String	Bireyin ikamet ettiği ülkenin bilgisi.
17	Daha önce tarama uygulamasının kullanılması	Boolean (evet-hayır)	Kullanıcının daha önce bir tarama uygulaması kullanıp kullanmadığı bilgisi.
18	Puanlama sonucu	Integer	Nihai puan, kullanılan tarama yönteminin puanlama algoritmasına dayanarak elde edilmiştir.
19	Teşti kim tamamlıyor	String	Bireyin testini kimin gerçekleştirdiği (Veli, kendisi, bakıcı, sağlık personeli, klinisyen, vb.) bilgisi.
20	Otizmlili olma durumu (Çıktı)	Boolean (evet-hayır)	Bireyin otizmlili olma durumunun bilgisi

(Thabtah vd., 2018)

Veri setinde bulunduğu halde veri setinin hangi yaş kategorisine uygulandığı bilgisini taşıyan ve tüm bireylerin çocuk olduğunu ifade etmek için 1 ile doldurulmuş olan "tarama yöntemi türü" özniteliği veri setinden çıkarılmıştır. Boolean türünde olan bireyin otizmlili olup olmama durumu ise sınıf etiketi olarak kullanılmıştır.

Çalışmada öncelikle öncelikle tüm öznitelikler sayısal değerlere çevrilmiştir. Bir sonraki aşamada veri seti üzerinde daha etkin analiz yapabilmek için eksik verilerin tamamlanması, verilerin 0 ile 1 arasında normalize edilmesi gibi veri ön işleme aşamaları gerçekleştirilmiştir. Daha sonra veriler yapay sinir ağları ve DKSBS ile sınıflandırılmış, k-means ve x-means ile kümelendirilmiştir. Tasarlanan ileri beslemeli YSA için kullanılan nöron sayılarının başarı oranına etkisi incelenmiş ve farklı yapıda düğümleri olan modeller test edilmiştir. Sınıflandırma yöntemlerinin model başarı oranını test etmek için veriler

%70 eğitim, %30 test olacak şekilde ayrılmıştır. Sonuçlara bakıldığı zaman veriler hem yapay sinir ağları hem de DKSBS ile sınıflandırıldığında başarı oranını hem test hem de eğitim setinde %100 olduğu görülmüştür. Kümeleme yöntemleri ön öğreticisiz yöntemlerdir. Dolayısı ile kümelenen veriler eğitim ve test olarak ikiye ayrılmamıştır ve verilerin tamamı eğitim için kullanılmıştır. Kümelenen verilerin başarı oranı k-means de % 89,73, x-means de %88,0 şeklindedir

4.2. Bulgular ve Değerlendirme

Literatür incelendiğinde otizm konusunda veri madenciliği yöntemleri kullanılarak gerçekleştirilen çok fazla çalışma olduğu görülmektedir. Thabtah (2017), çalışmasında bu konuya dikkat çekmiş, gerçekleştirilen çalışmalarda oldukça güzel sonuçların elde edilmiş olmasıyla birlikte kullanılan verilerin güncelliğini yitirdiğinden bahsetmiştir. Bu sebepten dolayı da güncel verilerle çalışılması gerektiğini vurgulamıştır. Devamında ise Thabtah (2018), yeni bir çalışma ile güncel ölçeğe uygun şekilde toplanan verileri paylaşmıştır. Bu çalışmada da Thabtah (2018), tarafından güncel ölçeğe uygun şekilde toplanan ve yapılan bilimsel yayın ile paylaşılan verilerin çocuklar için olan alt kümesi kullanılmıştır. Kullanılan veri seti 292 örnekten oluşmaktadır ve veri setinde toplamda 21 öznitelik mevcuttur. Veriler çok eski olmadığından bu veriler ile yapılan çalışmalarda sınırlıdır ve bu çalışma ile elde edilen bulgular diğer çalışmalarla karşılaştırılmıştır.

Veri setinin orijinal hali sayısal ve string ifade türünde değişkenler içermektedir. Bu çalışma kapsamında kullanılan normalizasyon işlemi ve sınıflandırma işlemi için kullanılan yöntemler ise string ifadelerle çalıştırılmamaktadır. Bu sebepten dolayı öncelikle tüm veri seti sayısal değerlere dönüştürülmüştür. Bu dönüşüm yapılırken string ifadelerle etiketlenmiş olan kategorik veriler, sayısal ifadeler ile etiketlenecek şekilde değiştirilmiştir. Örneğin bayan ya da erkek olarak etiketlenen kategorik veriler 0 ve 1 ile değiştirilmiştir. Etnik köken, ikamet edilen ülke gibi değerler de aynı şekilde 1'den başlanarak sıralı olarak etiketlenmiştir.

Tüm veriler sayısal hale getirildikten sonra her bir özneliğin kendi arasında frekansı çıkarılmıştır. Veri seti içerisindeki boş değerler frekansı en yüksek olan değer ile doldurulmuştur. Devamında ise birçok normalizasyon yöntemi denemiş en yüksek başarı oranına sahip literatürde de oldukça fazla kullanılan Min-Max normalizasyon yöntemi ile elde edilen değerler oluşturulmuştur. Yöntemlerin başarılarının karşılaştırılması için veri setinin bir kısmı eğitim bir kısmı test verileri olarak ayrılmıştır. Bu oranın yüzde kaç olarak ayrılması gerektiği konusunda literatür incelendiğinde çok farklı oranların kullanılmış olduğu görülmekle birlikte, en yaygın oran %70 eğitim verisi, %30 test verisi olarak ayrıldığı görülmektedir. Genel yaklaşıma uyması açısından bu çalışmada da %70'e karşılık gelen ve rastgele olarak belirlenen 205 örnek eğitim, %30'a karşılık gelen ve rastgele olarak belirlenen 87 örnek ise test için kullanılmıştır. Veriler YSA ve DKSBS ile sınıflandırılmış, K-Means ve X-Means ile kümelenebilir.

4.2.1. Yapay Sinir Ağı ile Elde Edilen Sonuçlar

Tasarlanan ileri beslemeli YSA ile kullanılan nöron sayılarının başarı oranına etkisi incelenmiş ve farklı yapıda düğümleri olan modeller test edilmiştir. YSA'nın performansı Tablo 2 de verilmiştir

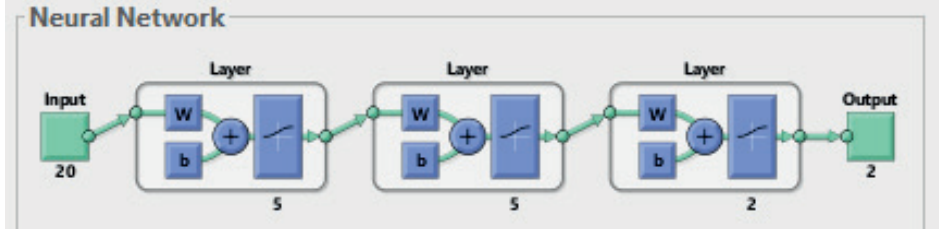
Tablo 2

Yapay Sinir Ağının performansı

Giriş katmanındaki nöron sayısı	1.Gizli katmandaki nöron sayısı	2.Gizli katmandaki nöron sayısı	Çıkış katmanındaki nöron sayısı	Eğitim verisindeki doğruluk oranı	Test verisindeki doğruluk oranı
20	1	1	2	%91,21	%90,80
20	10	10	2	%100	%97,70
20	5	5	2	%100	%98,85
20	30	30	2	%100	%96,56
20	15	15	2	%100	%96,56
20	50	50	2	%100	%97,70
20	20	20	2	%100	%96,56

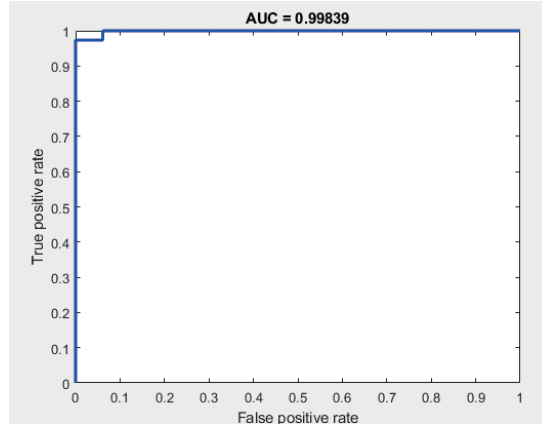
Tasarlanan modellerin başarı oranına bakıldığı zaman oldukça yüksek oranda başarıya ulaşıldığı görülmüştür. Farklı nöron sayıları ile tasarlanan modeller içerisinde en yüksek performans giriş katmanının da 20, iki gizli katmanda 5 ve çıkış katmanında 2 nöron kullanılarak tasarlanan modelde görülmüştür. Eğitim setinde %100 test setinde ise %98.85 oranında doğrulukla

başarı elde edilmiştir. Bu durum eğitim setindeki tüm örneklerin doğru sınıflandırıldığı, test veri setinde ise yalnızca bir örneğin yanlış sınıflandırıldığı anlamına gelmektedir. Tablo 4’de görüldüğü gibi gizli katmandaki nöron sayısını arttırmak başarı oranının artması ile doğru orantılı değildir. Nöron sayısını arttırmak bazı veri setleri için modelin başarı oranını arttırmakla beraber, bazı veri setleri için modelin başarı oranını düşürmektedir. Dolayısı ile her zaman nöron sayısının çok fazla ya da çok az kullanılması ile alakalı sabit bir yaklaşımı benimsemek yerine birkaç alternatif belirlenerek en başarılı olan modelin tercih edilmesi başarı oranını arttırabilmektedir. Başarı oranının en yüksek olarak görüldüğü YSA modeli Şekil 1de gösterilmiştir.



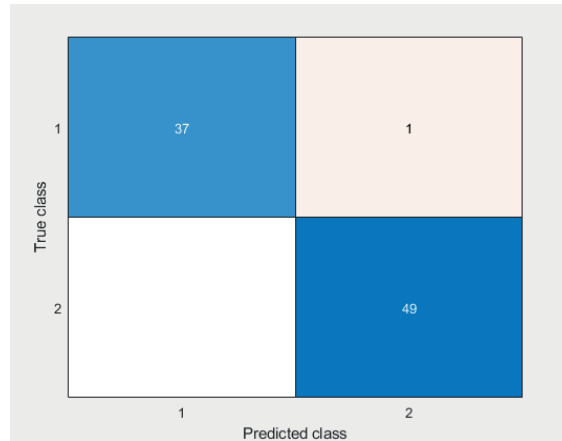
Şekil 1. Başarı oranının en yüksek olarak görüldüğü YSA modeli

Eğitilen ağın test verisi için çizdirilen ROC eğrisi (Receiver Operating Characteristic) Şekil 2 de verilmiştir.



Şekil 2. Eğitilen ağın test verisi için çizdirilen ROC eğrisi

Şekil 2 de görüldüğü üzere ROC eğrisi altında kalan alan (AUC) değeri 1’e çok yakın çıkmıştır. Bu durum, test verisinde yapılan sınıflandırmada oldukça yüksek bir başarı sağlandığını göstermektedir. Eğitilen ağın test verisi için çizdirilen hata matrisi Şekil 3 de verilmiştir.

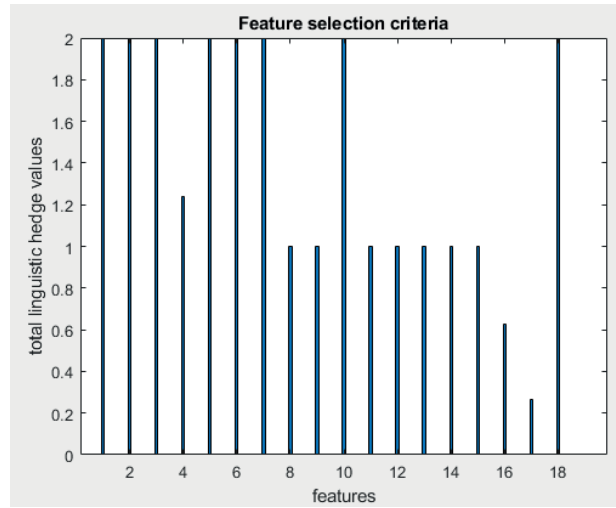


Şekil 3. Eğitilen ağın test verisi için çizdirilen hata matrisi

Şekil 3 de görüldüğü gibi eğitilen ağın test veri seti için gerçekleştirdiği sınıflandırmanın çok yüksek bir kısmı doğru tahminlenmiş, yalnızca bir örnek yanlış sınıflandırılmıştır. Bu durum test verisi için oldukça yüksek bir başarı anlamı taşımaktadır. Ağın test verisi için ürettiği hata değerlerine detaylı olarak bakıldığında MSE değerinin $4,822e-04$; RMSE değerinin ise $0,022$ olduğu görülmektedir. Eğitilen ağın eğitim verileri için MSE ve RMSE değerlerinden de anlaşıldığı gibi ağın hata değerleri oldukça düşüktür. Eğitilen ağın, test verileri için gerçekleştirdiği Doğruluk, Duyarlılık, Belirleyicilik ve F-Ölçütü değerleri sırası ile $0,989$ $0,974$ 1 $0,987$ şeklindedir. Bu durum eğitilen ağın test verileri içinde oldukça başarılı sonuçlar verdiğini göstermektedir.

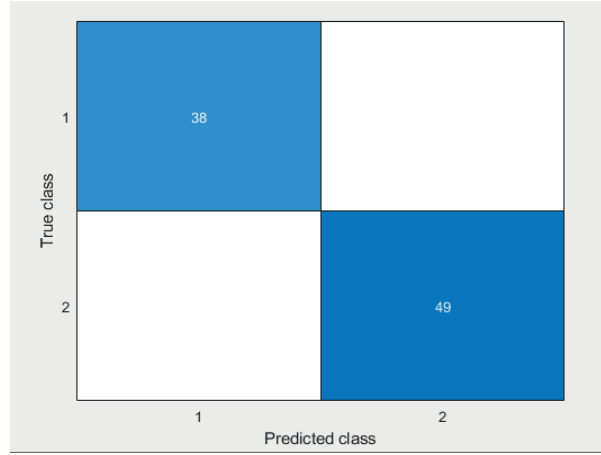
4.2.2. Dilsel Kuvvetli Sinir-Bulanık Sınıflayıcı ile Elde Edilen Sonuçlar

DKSBS ile öncelikle öznelik seçimi gerçekleştirilmiş devamında ise sınıflandırma işlemi yapılmıştır. Yapılan sınıflandırmada veriler %70 eğitim, %30 test verisi şeklinde ayrılarak gerçekleştirilmiştir. Sınıflandırma sonuçlarına bakıldığı zaman eğitim ve test verilerinde %100 oranda başarı elde edildiği görülmüştür. Yöntemin eğitim verileri için hesaplanan hata değerlerine bakıldığı zaman MSE değerinin $3,099e-32$, RMSE değerinin ise $1,760e-16$ olduğu görülmektedir. Gerçekleştirilen öznelik seçimi, özneliklerin önem düzeyine bakılarak belirlenmiştir. Özneliklerin DKSBS tarafından belirlenen önem düzeyleri Şekil 4 de verilmiştir.



Şekil 4. Özneliklerin DKSBS tarafından belirlenen önem düzeyleri

Şekil 4 de numaralandırılmış özneliklerin sırası Tablo 1 deki sıra ile aynı olacak şekilde 1'den 19'a kadar sıralı olarak ilerlemiştir. Dilsel kuvvetli sinir-bulanık sınıflayıcı'nın sınıflandırma için öznelikler üzerinde belirlemiş olduğu önem derecelerine bakıldığı zaman beş farklı önem düzeyine sahip grupların oluştuğu görülmektedir. Bu grupları en büyük öneme sahip onları 5 olacak şekilde, 5 den 1'e kadar derecelendirilecek olursa; 5 önem düzeyine sahip öznelikler 1,2,3,5,6,7,10,18 numaralı olanlar, 4 önem düzeyine sahip öznelik 4 numaralı olan, 3 önem düzeyine sahip öznelikler 8,9,11,12,13,14,15 numaralı olanlar, 2 önem düzeyine sahip öznelik 16 numaralı olan ve 1 önem düzeyine sahip olan 17 numaralı olan özneliktir. 19 numaralı özneliğin ise önem derecesi 0 olarak belirlenmiştir. Dilsel kuvvetli sinir-bulanık sınıflayıcı'nın ifade ettiği özneliklerin katkı dereceleri Şekil 5 de görülmektedir.



Şekil 7. Test verisi için çizdirilen hata matrisi

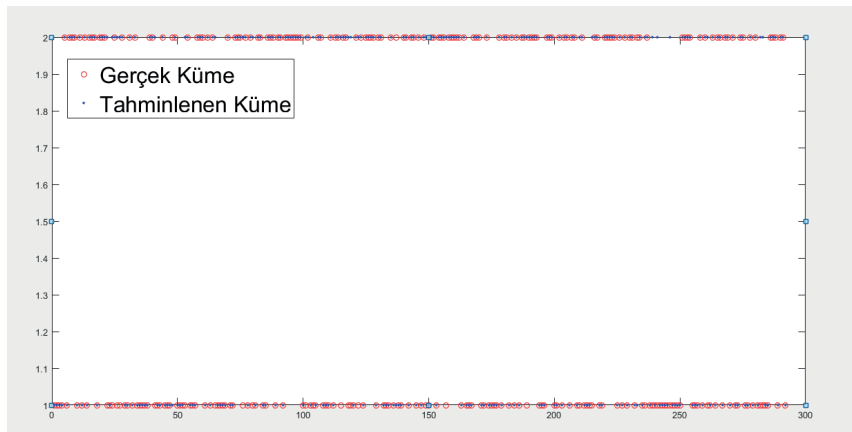
Şekil 7 de görüldüğü gibi eğitilen dilsel kuvvetli sinir-bulanık sınıflayıcı'nın test veri seti için gerçekleştirdiği sınıflandırmanın tamamı doğrudur.

Eğitilen dilsel kuvvetli sinir-bulanık sınıflayıcı'nın, test verileri için gerçekleştirdiği Doğruluk, Duyarlılık, Belirleyicilik ve F-Ölçütü değerleri ise sırası ile 1,1,1,1 şeklindedir. Bu durum eğitim verilerinin tamamının doğru sınıflandırıldığını göstermektedir.

4.2.3. K-Means ile Elde Edilen Sonuçlar

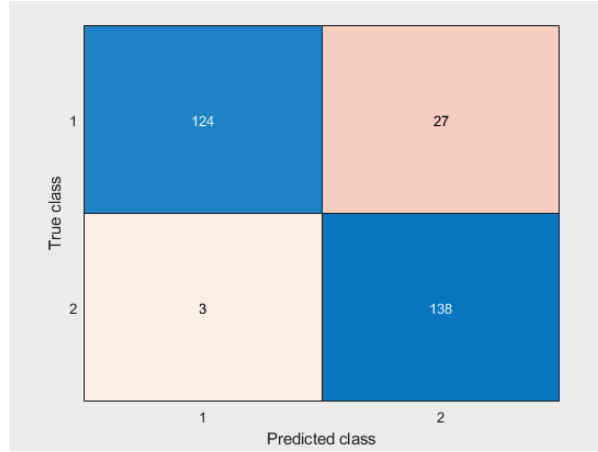
Kümeleme algoritmaları ön öğreticisiz yöntemlerdir ve verilerin sınıfları önceden belli değildir. Bu sebepten dolayı veliler eğitim ve test verisi olarak ayrılmamış, tamamı eğitim için kullanılmıştır.

K-means algoritmasında k parametresi ile ifade edilen küme sayısı önceden belirlenmeli ve çalıştırılmadan önce algoritmaya verilmelidir. Bu çalışmadaki verilerin OSB olup olmadığını gösterecek şekilde iki sınıfı vardır. Dolayısı ile kümeler oluşturulurken iki küme olması istendiğinden k parametresi iki olarak girilmiş ve veriler ikiye kümelendi. K-means algoritması ile kümelenen verilerin sınıfları, gerçek küme sınıfları ile karşılaştırıldığında toplamda 292 veriden 262'si doğru sınıflandırılmış ve sınıflandırmanın doğruluk oranı % 89,73 olmuştur. Gerçek sınıflar ve K-means ile elde edilen sınıflar Şekil 8 de verilmiştir.



Şekil 8. Gerçek sınıflar ve K-means ile elde edilen sınıflar

K-means algoritması ile elde edilen sonuçlar için çizdirilen hata matrisi Şekil 9'da verilmiştir.

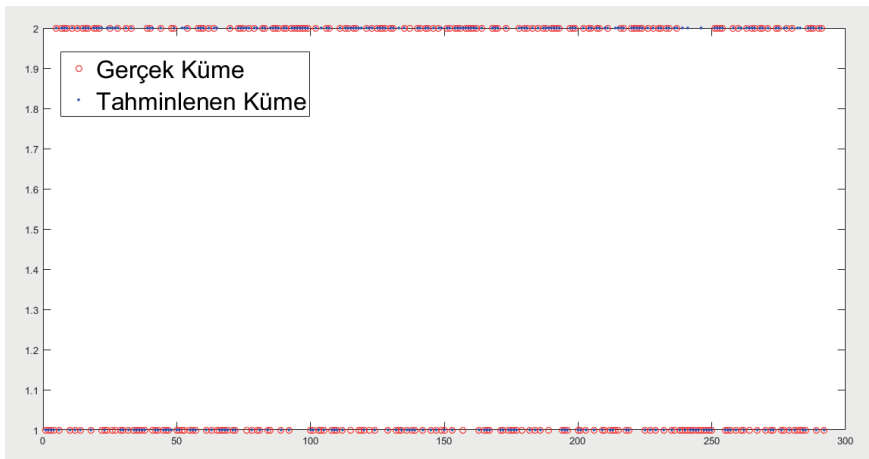


Şekil 9. K-means algoritması ile elde edilen sonuçlar için çizdirilen hata matrisi

Şekil 9’da görüldüğü gibi K-means algoritması ile kümelenen verilerin sınıfları çoğunlukla doğru tahminlenmiş olup, toplamda yanlış tahminlenen örnek sayısı 30 olmuştur. K-means algoritması ile kümelenen verilerin Doğruluk, Duyarlılık, Belirleyicilik ve F-Ölçütü değerleri sırası ile 0,897 0,821 0,979 0,892 şeklindedir.

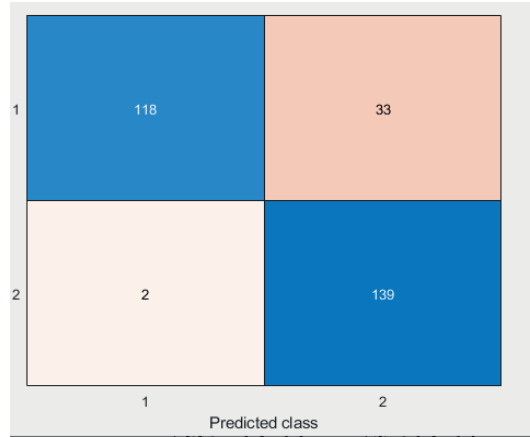
4.2.4. X-Means ile Elde Edilen Sonuçlar

X-means algoritması K-means algoritmasının geliştirilmesiyle oluşturulmuş daha güncel bir kümeleme yöntemidir. X-means, K-means’in çalışma yapısını kullanmasının yanında K-means’in geliştirilmiş halidir. Küme sayısının sabit bir değer olarak belirlenmesinin yerine bir aralık belirlenmesi ile çalışmaktadır. Böylece veri seti için en uygun olan küme sayısı algoritma tarafından belirlenebilmektedir. Küme sayısı belirlenirken, en iyi küme sayısını belirlemek için Bayesin bilgi kriteri kullanılmaktadır. Bu çalışmada, sınıf aralığı 2 ile 4 arasında girilerek çalıştırılan X-means algoritmasında küme sayısı algoritma tarafından 2 olarak belirlenmiştir. X-means algoritması ile ikiye kümelenen verilerin sınıfları, gerçek küme sınıfları ile karşılaştırıldığında toplamda 292 veriden 257’si doğru sınıflandırılmış ve sınıflandırmanın doğruluk oranı %88,02 olmuştur. Gerçek sınıflar ve X-means ile elde edilen sınıflar Şekil 10’da verilmiştir.



Şekil 10. Gerçek sınıflar ve X-means ile elde edilen sınıflar

X-means algoritması ile elde edilen sonuçlar için çizdirilen hata matrisi Şekil 11'de verilmiştir.



Şekil 11. X-means algoritması ile elde edilen sonuçlar için çizdirilen hata matrisi

Şekil 11'de görüldüğü gibi X-means algoritması ile kümelenen verilerin sınıfları çoğunlukla doğru tahminlenmiş olup, toplamda yanlış tahminlenen örnek sayısı 35 olmuştur. X-means algoritması ile kümelenen verilerin Doğruluk, Duyarlılık, Belirleyicilik ve F-Ölçütü değerleri sırası ile 0,880 0,781 0,986 0,871 şeklindedir.

4.2.5. Farklı Yöntemler ile Analiz Edilen Çocuklar İçin Otizm Spektrum Bozukluğu Verilerinin Tahminleme Başarılarının Karşılaştırılması

Bu çalışmada OSB veri setinin çocuklar için olan alt kümesi YSA ve DKSBS ile sınıflandırılmış, K-means ve X-means yöntemleri ile kümelendirilmiştir. Kullanılan yöntemlerin performans değerleri Tablo 2'de verilmiştir.

Tablo 2

Kullanılan yöntemlerin performans değerleri

Yöntem	Eğitim setinde doğruluk oranı	Test setinde doğruluk oranı
Yapay sinir ağları	%100	%98,85
Dilsel kuvvetli sinir-bulanık sınıflayıcı	%100	%100
K-Means	%89,73	Tüm veriler eğitim için kullanılmıştır
X-Means	%88,02	Tüm veriler eğitim için kullanılmıştır

Tablo 2'de görüldüğü gibi YSA ile sınıflandırma başarı oranı oldukça yüksek olmasına rağmen en yüksek başarı oranı DKSBS ile elde edilmiştir. DKSBS ile yapılan sınıflandırma da eğitim ve test verisi için %100 oranında doğruluk gerçekleştirilmiştir. Bu durum dilsel kuvvetli sinir-bulanık sınıflayıcının veri setindeki tüm örnekleri doğru sınıflandırdığı anlamını taşımaktadır. Kümeleme yöntemlerinin başarı oranları incelendiği zaman genel anlamda sınıflandırma yöntemlerine göre daha az başarılı olduğu ve K-means yöntemi ile elde edilen sonuçların, X-means yöntemine göre daha başarılı olduğu görülmektedir. Kullanılan Yöntemlerin Eğitim veri setleri için Doğruluk, Duyarlılık, Belirleyicilik ve F-Ölçütü değerleri Tablo 3'de verilmiştir.

Tablo 3

Kullanılan yöntemlerin eğitim veri setleri için Doğruluk, Duyarlılık, Belirleyicilik ve F-Ölçütü değerleri

Yöntem	Doğruluk	Duyarlılık	Belirleyicilik	F-Ölçütü
Yapay sinir ağları	1	1	1	1
Dilsel kuvvetli sinir-bulanık sınıflayıcı	1	1	1	1
K-Means	0,897	0,821	0,979	0,892
X-Means	0,880	0,781	0,986	0,871

Tablo 3'de sınıflandırma ve kümeleme yöntemlerinin eğitim verisi için hesaplanan Doğruluk, Duyarlılık, Belirleyicilik ve F-Ölçütü değerleri görülmektedir. Görüldüğü üzere sınıflandırma yöntemlerinde hesaplanan tüm değerler 1'dir. Kümeleme yöntemlerinde ise ortalama olarak K-means için hesaplanan değerler X-means'ten daha yüksektir.

Kullanılan Yöntemlerin test veri setleri için Doğruluk, Duyarlılık, Belirleyicilik ve F-Ölçütü değerleri Tablo 4’de verilmiştir.

Tablo 4

Kullanılan yöntemlerin test veri setleri için Doğruluk, Duyarlılık, Belirleyicilik ve F-Ölçütü değerleri

Yöntem	Doğruluk	Duyarlılık	Belirleyicilik	F-Ölçütü
Yapay sinir ağları	0.989	0.974	1	0.987
Dilsel kuvvetli sinir-bulanık sınıflayıcı	1	1	1	1
K-Means	Tüm veriler eğitim için kullanılmıştır	Tüm veriler eğitim için kullanılmıştır	Tüm veriler eğitim için kullanılmıştır	Tüm veriler eğitim için kullanılmıştır
X-Means	Tüm veriler eğitim için kullanılmıştır	Tüm veriler eğitim için kullanılmıştır	Tüm veriler eğitim için kullanılmıştır	Tüm veriler eğitim için kullanılmıştır

Tablo 4’de görüldüğü gibi test verisi için DKSBS ile yapılan sınıflandırmanın Doğruluk, Duyarlılık, Belirleyicilik ve F-Ölçütü değerleri 1 olarak hesaplanmıştır. Bu durum dilsel kuvvetli sinir-bulanık sınıflayıcının test verisinin tamamını doğru sınıflandırdığı anlamı taşımaktadır. Test verisi için YSA ile yapılan sınıflandırmanın Doğruluk, Duyarlılık, Belirleyicilik ve F-Ölçütü değerleri de 1’e oldukça yakın değerler çıkmıştır. Bu durum test verisi için dilsel kuvvetli sinir-bulanık sınıflayıcıdan daha başarısız olmakla birlikte kullanılabilir bir yöntem olduğunu ortaya koymaktadır. K-Means ve X-Means ile kümelene verilerin tamamı eğitim için kullanılmış, dolayısı ile test verisi için hesaplanan Doğruluk, Duyarlılık, Belirleyicilik ve F-Ölçütü değerleri bulunmamaktadır. Çalışmanın performans değerleri ve aynı veri setini kullanan diğer çalışmalardaki performans değerleri Tablo 5’de verilmiştir.

Tablo 5

Çalışmanın performans değerleri ve aynı veri setini kullanan diğer çalışmalardaki performans değerleri

Çalışma	Yöntem	Performans
Thabtah, 2018	Logistic regression	%97,94
Thabtah, 2018	Naive bayes	%92,80
Akyol ve Karaci, 2018	Lojistik regresyon	%92,00
Akyol ve Karaci, 2018	Bulanık kuralı lojistik regresyon kombinasyonu	%97,33
De Campos Souza ve Guimaraes, 2018	Bulanık sinir ağı mimarisi	%100
Hutchinson vd., 2019	J48 decision tree	%92,80
Hutchinson vd., 2019	Naive bayes	%96,20
Bu çalışma	Yapay sinir ağları	%98,85
Bu çalışma	Dilsel kuvvetli sinir-bulanık sınıflayıcı	%100
Bu çalışma	K-Means	%89,73
Bu çalışma	X-Means	%88,02

Tablo 5’de görüldüğü gibi aynı veri seti farklı çalışmalarda farklı yöntemlerle sınıflandırılmıştır ama bu çalışma dışında bir kümeleme işlemine rastlanmamıştır. Yapılan çalışmalardaki sonuçlar incelendiği zaman genel olarak başarı oranının %90’ın üzerinde olduğu görülmekle birlikte %100 olan bir çalışma da bulunmaktadır. Bu çalışma ile elde edilen sonuçlar ile diğer çalışmalarda elde edilen sonuçlar karşılaştırıldığı zaman ise verilerin DKSBS ve YSA ile sınıflandırıldığında sınıflandırma başarısının logistic regression, naive bayes, bulanık kuralı lojistik regresyon kombinasyonu ve j48 decision tree ile sınıflandırılan çalışmalardaki sınıflandırma başarısından daha yüksek olduğu görülmektedir. Bulanık sinir ağı mimarisi ile gerçekleştirilen sınıflandırma ve bu çalışma kapsamının da DKSBS ile gerçekleştirilen sınıflandırmanın başarı oranları ise aynıdır. İki yöntem ile sınıflandırılan verilerdeki doğruluk aynı orandadır ve bu oran %100 şeklindedir. Bu durum iki yöntemde tüm örnekleri doğru sınıflandırdığı anlamına gelmektedir. Bu çalışma kapsamında gerçekleştirilen kümeleme işlemi sonuçları incelendiği zaman ise kümeleme yöntemlerinden K-means yönteminin X-means yöntemine göre daha başarılı olduğu ve iki kümeleme yönteminin de başarı oranının sınıflandırma yöntemlerine göre daha düşük olduğu sonucuna ulaşılmaktadır.

5. TARTIŞMA, SONUÇ ve ÖNERİLER

Bu çalışma ile OSB konusunda veri madenciliği, yapay zeka gibi teknikleri kullanarak gerçekleştirilen çalışmalar incelenmiştir. Bu konuda oldukça başarılı sonuçlara ulaşan birçok çalışmanın olduğu görülmektedir. Bunun yanı sıra bilimsel bir yayın ile son dönemde gerçekleştirilen çalışmalarda güncelliğini yitiren veri setlerinin kullanıldığı eleştirisi dikkat çekmektedir. Yayının detayları incelendiği zaman güncel ve başarılı yöntemlerin kullanılarak teknolojik gelişmelerden faydalanıldığı

fakat deneysel çalışmalarda kullanılan verilerin güncelliğini yitirdiği vurgusu yapıldığı görülmektedir. Çalışmanın devamında ise eksiklik olarak görülen güncel olmayan veriler, geliştirilen son ölçeğe dayalı olarak toplanmış ve bir takım işlemler gerçekleştirildikten sonra paylaşılmıştır. Veri seti çocuk, genç, yetişkin şeklinde gruplandırılmak üzere toplamda 1100 örnek içermektedir. Bu çalışmada ise aynı veri setinin 292 örnek içeren çocuklar için olan alt küme kullanılmıştır. Toplamda 20 girdiden oluşan fakat tüm örneklerde aynı olduğu için çıkarılan bir öznelikten sonra 19'a düşürülen girdiler sonucunda bireyin otizmlili olma durumunu ifade eden çıktının tahminlenmesi için veri setine veri madenciliği yöntemlerinden olan sınıflandırma ve kümeleme işlemleri uygulanmıştır. Sınıflandırma ve kümeleme işlemleri uygulanmadan önce öncelikle tamamen numeric değerlere dönüştürülen veri seti üzerinde özneliklerin frekansına bakılarak eksik veriler tamamlanmış ve normalizasyon işlemleri gerçekleştirilmiştir. Yöntemlerin çalıştırılması için hazır hale getirilen veri seti son olarak %70 eğitim, %30 test verisi olmak üzere ayrılmıştır. Sınıflandırma işlemi için veriler test ve eğitim verisi olacak şekilde ayrılarak sınıflandırılmış, elde edilen sonuçlar birçok parametre ile ifade edilmiştir. Sınıflandırma işlemi için kullanılan yöntemler YSA ve DKSBS olmuştur. Kümeleme yöntemleri önöğreticisiz yöntemler olduğundan dolayı veri seti eğitim ve test olarak ayrılmamış tamamı eğitim verisi olacak şekilde kümeleneştir. Kümeleme işlemi için K-means ve X-means yöntemleri kullanılmıştır. Elde edilen sonuçlar birçok parametre ile ifade edilmiş olmakla birlikte tüm yöntemlerde ortak paydada karşılaştırma yapılabilmesi için temel olarak Doğruluk, Duyarlılık, Belirleyicilik, F-Ölçütü kullanılmıştır.

Sınıflandırma sonuçlarına bakıldığı zaman eğitim veri seti için YSA ve DKSBS olmak üzere iki yöntemde başarı oranının %100 olduğu görülmüştür. Dolayısı ile iki yöntemin eğitim veri seti için hesaplanan Doğruluk, Duyarlılık, Belirleyicilik, F-Ölçütü değerleri 1 olarak hesaplanmıştır. Bu durum iki yöntemde eğitim verilerinin tamamını doğru sınıflandırdığı anlamına gelmektedir. Test seti için elde edilen sonuçlara bakıldığı zaman başarı oranlarının birbirinden farklı olduğu görülmektedir. Dilsel kuvvetli sinir-bulanık sınıflayıcı'nın test verisi için gerçekleştirdiği başarı oranı eğitim verisindeki başarı oranı ile aynı olmuştur. Dolayısı ile test seti için hesaplanan Doğruluk, Duyarlılık, Belirleyicilik, F-Ölçütü değerleri 1 olarak hesaplanmıştır. Bu durum test verilerinin DKSBS tarafından tamamının doğru sınıflandırdığı anlamına gelmektedir. YSA'nın test verisi için gerçekleştirdiği başarı oranı ise %98.85 olmuştur. Test seti için hesaplanan Doğruluk, Duyarlılık, Belirleyicilik, F-Ölçütü değerleri sırası ile 0,989 0,974 1 0,987 olarak hesaplanmıştır. Bu sonuçlar YSA ile yapılan sınıflandırmanın test verisinde 87 örnekten yalnızca 1 tanesini yanlış sınıflandırdığı geriye kalan 86 örneği doğru sınıflandırdığını anlamını taşımaktadır. Sınıflandırma yöntemlerinin başarılarına bakıldığı zaman iki yöntemde oldukça başarılı olduğu ve veri setini sınıflandırmak için kullanılabilmesi görülmekle birlikte, tüm verileri doğru sınıflandırdığından dolayı dilsel kuvvetli sinir-bulanık sınıflayıcının daha başarılı olduğu söylenebilmektedir.

Kümeleme sonuçlarına bakıldığı zaman eğitim ve test verisi olarak ayrılmadığı tüm verilerin eğitim için kullanıldığı görülmektedir. Dolayısı ile K-means ve X-means yöntemi ile kümelenen verilerde yalnızca eğitim verisi için sonuçlar bulunmaktadır. K-means ile kümelenen verilerin başarı oranı %89,73 olmuştur ve verilerin hesaplanan Doğruluk, Duyarlılık, Belirleyicilik, F-Ölçütü değerleri sırası ile 0,897 0,821 0,979 0,892 şeklindedir. Bu sonuçlar K-means ile kümelenen verilerin 262 tanesinin doğru, 30 tanesinin yanlış sınıflandırıldığı anlamına gelmektedir. X-means ile kümelenen verilerin başarı oranı ise %88,02 olmuştur ve verilerin hesaplanan Doğruluk, Duyarlılık, Belirleyicilik, F-Ölçütü değerleri sırası ile 0,880 0,781 0,986 0,871 şeklindedir. Bu sonuçlar X-means ile kümelenen verilerin 257 tanesinin doğru, 35 tanesinin yanlış sınıflandırıldığı anlamına gelmektedir. Bu çalışma kapsamında elde edilen sonuçlara göre kümeleme yöntemlerinin başarılarına bakıldığı zaman sonuçların genel anlamda yüksek olduğu fakat sınıflandırma yöntemlerine göre daha düşük kaldığı görülmekle birlikte bu çalışmadaki veri seti için K-means yönteminin X-means yöntemine göre daha başarılı sonuçlar verdiği tespit edilmiştir.

Bu çalışmada elde edilen sonuçlar ile aynı veri seti ile yapılan diğer çalışmalardaki sonuçlar karşılaştırıldığı zaman Thabtah (2018), logistic regression ile yaptığı sınıflandırmanın başarı oranı %97,94 naive bayes ile yaptığı sınıflandırmanın başarı oranı %92,80 bulmuştur. Akyol ve Karaci (2018), lojistik regresyon ile yaptıkları sınıflandırmanın başarı oranı %92,00 bulanık kurallı lojistik regresyon kombinasyonu ile yaptıkları sınıflandırmanın başarı oranı %97,33 olarak elde etmişlerdir. De Campos Souza ve Guimaraes (2018), bulanık sinir ağı mimarisi ile yaptıkları sınıflandırmanın başarı oranı %100 olarak tespit etmişlerdir. Hutchinson ve arkadaşları (2019), j48 decision tree ile yaptıkları sınıflandırmanın başarı oranını %92,80

naive bayes ile yaptıkları sınıflandırmanın başarı oranını %96,20 olarak bulmuşlardır. Bu durum genel olarak veri setinin sınıflandırmaya uygun bir veri seti olduğunu göstermektedir. Aynı zamanda bu çalışma ile %100 oranında başarıya ulaşan DKSBS yönteminin kullanılan veri seti için literatürdeki birçok yöntemden daha başarılı sonuçlar elde ettiğini de göstermektedir. Elde edilen başarı oranlarının yüksek olması önemli olmakla birlikte gerçek hayat problemlerinde %100 oranında başarının var olması gerçekleşmesi zor bir ihtimaldir ve hata payının her zaman olabileceği ihtimalinin göz önünde bulundurulması gerekmektedir.

Bu çalışmanın literatüre olan katkıları ise şöyledir; öncelikle aynı veri seti ile yapılan çalışmalar incelendiği zaman logistic regression, naive bayes, bulanık kurallı lojistik regresyon kombinasyonu, bulanık sinir ağı mimarisi, j48 decision tree gibi yöntemlerin kullanılması ile yapılan sınıflandırma işlemleri olmakla birlikte çocuklar için olan alt kümede herhangi bir yöntem ile kümeleme işleminin gerçekleştirildiği bir çalışmaya rastlanmamış olmakla birlikte, YSA ve DKSBS ile yapılan bir sınıflandırmaya da rastlanmamıştır. Sonuçlara bakıldığında zaman tahminleme doğruluğu bakımından OSB'nin çocuklar için olan verilerinde, sınıflandırma yöntemlerinin kümeleme yöntemlerinden daha başarılı olduğu görülmektedir. Özellikle DKSBS yönteminin tüm verileri doğru sınıflandırarak literatürdeki birçok yöntemden daha yüksek başarı oranına sahip olduğu için tercih edilebilecek yöntemler arasında en iyi yöntemlerden bir tanesi olabileceği sonucuna varılmaktadır. Dolayısı ile bu çalışma ile kullanılan veri seti için farklı yöntemlerin denenmesine imkân tanınarak literatürdeki birçok çalışmadan daha başarılı sonuçlar elde ettiği görülen bir yöntem ile sınıflandırılma yapılmasına imkan tanımaktadır. Bunun yanı sıra elde edilen sonuçların değerlendirilmesi için kullanılan parametreler birçok çalışmadan daha detaylı olarak verilmiştir. Bu durum elde edilen sonuçların daha detaylı yorumlanmasına imkan tanımaktadır.

İleriki çalışmalarda ise veri setinin farklı alt kümeleri kullanılarak veya tüm alt kümeleri aynı anda kullanılarak bir takım analizler gerçekleştirilebilir. Yapılacak çalışmalarda karar destek sistemleri, uzman sistemler, görüntü işleme teknikleri gibi daha farklı teknolojiler ile desteklenen daha geniş kapsamlı veri setlerine uygulanan çalışmalar gerçekleştirilebilir. Verilerin yeniden güncel olarak toplanması için alanında uzman hekim veya hekimler ile birlikte ortak olarak yeni çalışmalar yürütülebilir. Devamında ise elde edilen bulgular ile sürekli öğrenmeye devam eden ve normal hayatta da kullanılabilen somut ürünler ortaya konulabilir.

Hakem Değerlendirmesi: Dış bağımsız.

Çıkar Çatışması: Yazarlar çıkar çatışması bildirmemiştir.

Finansal Destek: Yazarlar bu çalışma için finansal destek almadığını beyan etmiştir.

Yazar Katkıları: Çalışma Konsepti/Tasarım- S.Ç., M.Ş.Ç.; Veri Toplama- S.Ç., M.Ş.Ç.; Veri Analizi/Yorumlama- S.Ç., M.Ş.Ç.; Yazı Taslağı- S.Ç., M.Ş.Ç.; İçeriğin Eleştirel İncelemesi- S.Ç., M.Ş.Ç.; Son Onay ve Sorumluluk- S.Ç., M.Ş.Ç.

Peer-review: Externally peer-reviewed.

Conflict of Interest: The authors have no conflict of interest to declare.

Grant Support: The authors declared that this study has received no financial support.

Author Contributions: Conception/Design of Study- S.Ç., M.Ş.Ç.; Data Acquisition- S.Ç., M.Ş.Ç.; Data Analysis/Interpretation- S.Ç., M.Ş.Ç.; Drafting Manuscript- S.Ç., M.Ş.Ç.; Critical Revision of Manuscript- S.Ç., M.Ş.Ç.; Final Approval and Accountability- S.Ç., M.Ş.Ç.

Kaynaklar/References

- Akçapınar, G., Altun, A. & Aşkar, P. (2016). Çevrimiçi Öğrenme Ortamındaki Benzer Öğrenci Gruplarının Kümeleme Yöntemi İle Belirlenmesi. *Eğitim Teknolojisi Kuram ve Uygulama*, 6(2), 46-64. DOI: 10.17943/etku.91440.
- Akdemir, Ç. (2016). "Hilenin Veri Madenciliği İle Ortaya Çıkarılması ve Perakende Sektöründe Bir Uygulama", (Yayınlanmamış Doktora Tezi), İstanbul: Marmara Üniversitesi Sosyal Bilimler Enstitüsü.
- Akgöbek, Ö., ve Çakır, F. (2009). Veri Madenciliğinde Bir Uzman Sistem Tasarımı, *Akademik Bilişim Konferansı*, 9, 801-806.
- Akkaya, G. C., Demireli, E., Yakut, Ü. H., & Yakut, H. (2009). İşletmelerde Finansal Başarısızlık Tahminlemesi: Yapay Sinir Ağları Modeli İle İmkân Üzerine Bir Uygulama. *Eskişehir Osmangazi Üniversitesi Sosyal Bilimler Dergisi*, 10(2), 187-216.
- Akyılmaz, O., & Ayan, T. (2010). Esnek Hesaplama Yöntemlerinin Jeodezide Uygulamaları. *İTÜ Dergisi*, 5(1), 261-268.
- Akyol, K., & Karaci, A. (2018). A Study on Autistic Spectrum Disorder For Children Based On Feature Selection And Fuzzy Rule, *International Congress on Engineering and Life Science*, 804–807.
- Aydın, S. (2007). "Veri Madenciliği ve Anadolu Üniversitesi Uzaktan Eğitim Sisteminde Bir Uygulama", (Yayımlanmamış Doktora Tezi). Eskişehir: Anadolu Üniversitesi Sosyal Bilimler Enstitüsü.

- Bholowalia, P., & Kumar, A. (2014). EBK-Means: A Clustering Technique Based On Elbow Method And K-Means In WSN. *International Journal of Computer Applications*, 105(9), 17-24.
- Bradley, P. S., Fayyad, U. M., & Reina, C. (1998). Scaling Clustering Algorithms to Large Databases. *KDD*, 98, 9-15.
- Brossette, S. E., Sprague, A. P., Hardin, J. M., Waites, K. B., Jones, W. T., & Moser, S. A. (1998). Association Rules and Data Mining in Hospital Infection Control And Public Health Surveillance. *Journal of The American Medical Informatics Association*, 5(4), 373-381.
- Budak, E. Ç., & Bozkurt, M. R. (2013). Vertebra Lomber Disklerde Meydana Gelen Bozulmaların Manyetik Rezonans Görüntüleme (Mrg) İle Analizi. *AJIT-e: Online Academic Journal of Information Technology*, 4(11), 125-144.
- Çağlar, M. F., Çetişli, B., & Toprak, İ. B. (2010). Automatic Recognition of Parkinson's Disease From Sustained Phonation Tests Using ANN and Adaptive Neuro-Fuzzy Classifier. *Mühendislik Bilimleri ve Tasarım Dergisi*, 1(2), 59-64.
- Çalışkan, S. K., & Soğukpınar, İ. (2008). Kxknn: K-Means ve K En Yakın Komşu Yöntemleri İle Ağlarda Nüfuz Tespiti. *EMO Yayınları*, 120-24.
- Çelik, S., Bozkurt, Ö. Ç., & Çeşmeli, M. Ş. (2018). İnsan Omurgası Veri Setinin Sinir-Bulanık Sınıflayıcı İle Öznitelik Tespiti ve Sınıflandırılması. *Yönetim Bilişim Sistemleri Dergisi*, 4(1), 39-52.
- Çelik, S. (2020). Determination and Classification of Importance of Attributes Used in Diagnosing Pregnant Women's Birth Method. *Alphanumeric Journal*, 8(2), 261-274.
- Cetişli, B., & Kalkan, H. (2011). Classification of Multispectral Satellite Images By Using Adaptive Neuro-Fuzzy Classifier with Linguistic Hedges. *Signal Processing and Communications Applications Conference*, 19, 50-53.
- Çetişli B. (2006). Öznitelik Seçiminde Dilsel Kuvvetli Sinir Bulanık Sınıflayıcı Kullanımı. *Eskişehir Osmangazi Üniversitesi Mühendislik ve Mimarlık Fakültesi Dergisi*, 19(2), 109-130.
- Çınaroğlu, S., & Bulut, H. (2018). New Initialization Approaches for The K-Means And Particle Swarm Optimization Based Clustering Algorithms. *Journal of the Faculty of Engineering and Architecture of Gazi University*, 33(2), 413-423.
- Deb, C., Lee, S. E., & Santamouris, M. (2018). Using Artificial Neural Networks to Assess HVAC Related Energy Saving in Retrofitted Office Buildings. *Solar Energy*, 163, 32-44.
- Demiralay, M., & Çamurcu, A. Y. (2005). Cure, Agnes ve K-Means Algoritmalarındaki Kümeleme Yeteneklerinin Karşılaştırılması. *İstanbul Ticaret Üniversitesi Fen Bilimleri Dergisi*, 4(8), 1-18.
- De Campos Souza, P. V., & Guimaraes, A. J. (2018). Using Fuzzy Neural Networks for Improving the Prediction of Children with Autism Through Mobile Devices. *Symposium on Computers and Communications*, 01086-01089.
- Diler, S. (2016). "Veri Madenciliği Süreçleri ve Karar Ağaçları Algoritmaları ile Bir Uygulama", (Yayımlanmamış Yüksek Lisans Tezi). Van: Yüzüncü Yıl Üniversitesi Fen Bilimleri Enstitüsü.
- Doğan, B., Buldu, A., Demir, Ö., & Erol, B. (2018). Sigortacılık Sektöründe Müşteri İlişki Yönetimi İçin Kümeleme Analizi. *Karaelmas Fen ve Mühendislik Dergisi*, 8(1), 11-18.
- Feng, Y., & Hamerly, G. (2007). PG-Means: Learning The Number of Clusters in Data. In *Advances in neural information processing systems*, 393-400.
- Gülsöz, T., & Çıkılı, Y. (2018). Otizm Spektrum Bozukluğu Olan Öğrencilere Soğuk İçecek Hazırlama ve Sunma Becerisinin Video Model ile Öğretimin Etkililiği. *Abant İzzet Baysal Üniversitesi Eğitim Fakültesi Dergisi*, 18(1), 210-229.
- Güner, S., Codal, K. S., Geçer, H. S., & Coşkun, E. (2018). Trafik Kaza Desenlerinin Tanımlanmasında K-Means Kümeleme Algoritmasının Kullanılması: Sakarya İli Uygulaması. *İşletme Bilimi Dergisi*, 6(3), 89-105.
- Hamerly, G., & Elkan, C. (2004). Learning The K in K-Means. In *Advances in Neural Information Processing Systems*, 281-288.
- Hong, T. P., & Wu, C. W. (2011). Mining Rules From an Incomplete Dataset with a High Missing Rate. *Expert Systems with Applications*, 38(4), 3931-3936.
- Huang, D., & Chow, T. W. (2005). Efficiently Searching The Important Input Variables Using Bayesian Discriminant. *IEEE Transactions on Circuits and Systems I: Regular Papers*, 52(4), 8785-793.
- Hutchinson, J., Schauer, I., & Seetan, R. (2019). A Comparative Study of Data Mining Techniques Used to Test Predictive Accuracy of Autism Spectrum Disorder Screening Process. In *34th Annual Conference of The Pennsylvania Association of Computer and Information Science Educators*, 70-75.
- Işık, M., & Çamurcu, A. Y. (2007). K-Means, K-Medoids ve Bulanık C-Means Algoritmalarının Uygulamalı Olarak Performanslarının Tespiti. *İstanbul Ticaret Üniversitesi Fen Bilimleri Dergisi*, 6(13), 31-45.
- Işık M., & Çamurcu, A. Y. (2011). Document Clustering Using K-Means and Hyperspherical Fuzzy C-Means Algorithms. *Marmara Fen Bilimleri Dergisi*, 22(1), 1-18.
- Jain, A. K. (2010). Data Clustering: 50 Years Beyond K-Means. *Pattern Recognition Letters*, 31(8), 651-666.
- Jayalakshmi, T., & Santhakumaran, A. (2011). Statistical Normalization and Back Propagation for Classification. *International Journal of Computer Theory and Engineering*, 3(1), 1793-8201.
- Kalantar, B., Pradhan, B., Naghibi, S. A., Motevalli, A., & Mansor, S. (2018). Assessment of The Effects of Training Data Selection on The Landslide Susceptibility Mapping: A Comparison Between Support Vector Machine (SVM), Logistic Regression (LR) and Artificial Neural Networks (ANN), Geomatics, *Natural Hazards and Risk*, 9(1), 49-69.
- Kalogeratos, A., & Likas, A. (2012). Dip-Means: an Incremental Clustering Method For Estimating The Number Oo Clusters. In *Advances in Neural Information Processing Systems*, 2393-2401.
- Khaled, A., Sanjay, R., & Vineet, S. (1998). An Efficient K-Means Clustering Algorithm. In *IPPS: 11th International Parallel Processing Symposium*.
- Bilgin, T. T., & Çamurcu, Y. (2005). DbSCAN, Optics ve K-Means Kümeleme Algoritmalarının Uygulamalı Karşılaştırılması. *Politeknik Dergisi*, 8(2), 139-145.

- Kılınc, G. E., & Söğüt, M. Ü. (2018). Mikrobiyotaya Güncel Bir Yaklaşım: Otizm ve Mikrobiyotaya, *Türkiye Klinikleri Journal of Health Sciences*, 3(1), 88-94.
- Koçyiğit, Y., & Korürek, M. (2010). EMG İşaretlerini Dalgacık Dönüşümü ve Bulanık Mantık Sınıflayıcı Kullanarak Sınıflama. *İTÜDERGİSİ/d*, 4(3), 25-31.
- Küçükkoçaoğlu, G., Benli Keskin, Y., & Küçüksözen, C., (2005). Finansal Bilgi Manipülasyonunun Tespitinde Yapay Sinir Ağı Modelinin Kullanımı. *İMKB Dergisi*, 9(36), 1-23.
- Liu, H., Dougherty, E. R., Dy, J. G., Torkkola, K., Tuv, E., Peng, H., ... & Zhao, Z. (2005). Evolving Feature Selection, *IEEE Intelligent systems*, 20(6), 64-76.
- Mackinnon, M. J., & Glick, N. (1999). Applications: Data Mining and Knowledge Discovery in Databases—an Overview. *Australian & New Zealand Journal of Statistics*, 41(3), 255-275.
- MacQueen, J. (1967). Some Methods for Classification and Analysis of Multivariate Observations. In Proceedings of The Fifth Berkeley Symposium on Mathematical Statistics and Probability, 1(14), 281-297.
- Mehmet, D., & Akpınar, E. (2018). Sabit ve Hareketli Hava İle Isıtılan Güneş Kolektörlerinin Isıl Verimliliklerinin Karşılaştırılması ve Yapay Sinir Ağları İle Modellenmesi. *Fırat Üniversitesi Mühendislik Bilimleri Dergisi*, 30(1), 41-46.
- Moskov, D. (2016). "Knowledge Extraction From Published Papers in Literature For The Catalytic Methanol Production From Synthesis Gas Using Data Mining Tools", (Yayımlanmamış Yüksek Lisans Tezi). İstanbul: Boğaziçi Üniversitesi. Fen Bilimleri Enstitüsü.
- Muhr, M., & Granitzer, M. (2009). Automatic Cluster Number Selection Using A Split and Merge K-Means Approach, *In 20th International Workshop on Database and Expert Systems Application*, 363-367.
- Nilashi, M., Ibrahim, O., Dalvi, M., Ahmadi, H., & Shahmoradi, L. (2017). Accuracy Improvement For Diabetes Disease Classification: A Case on A Public Medical Dataset. *Fuzzy Information and Engineering*, 9(3), 345-357.
- Oğuzlar, A. (2003). Veri Ön İşleme. *Erciyes Üniversitesi İktisadi ve İdari Bilimler Fakültesi Dergisi*, 21, 67-76.
- Olçay Gül, S., & Tekin İftar, E. (2012). Otizm Spektrum Bozukluğu Tanısı Olan Bireyler İçin Sosyal Öykülerin Kullanımı. *Ankara Üniversitesi Eğitim Bilimleri Fakültesi Özel Eğitim Dergisi*, 13(2), 1-24.
- Özsoy, İ., & Fırat, M. (2004). Kirişsiz Döşemeli Betonarme Bir Binada Oluşan Yatay Deplasmanın Yapay Sinir Ağları İle Tahmini. *Deü Mühendislik Fakültesi Fen ve Mühendislik Dergisi*, 6(1), 51-63.
- Pelleg, D., & Moore, A. W. (2000). X-Means: Extending K-Means with Efficient Estimation of The Number of Clusters, *In Icml*, 1, 727-734.
- Pençe, İ., & Çetişli, B. (2013). El Yazı Karakterlerinin Kapalı Cebirsel Eğrilerle Modellenmesi ve Sınıflandırılması. *Sigma*, 5, 1-7.
- Reddy, S. K., Kodali, S. R., & Gundabathina, J. L. (2012). Classification of Vertebral Column Using Naïve Bayes Technique. *International Journal of Computer Applications*, 58(7), 38-42.
- Sabuncuoğlu, M., Cebeci, S., Rahbar, M. H., & Hessabi, M. (2015). Autism Spectrum Disorder and Attention Deficit Hyperactivity Disorder: Knowledge and Attitude of Family Medicine Residents in Turkey. *Turkish Journal of Family Medicine & Primary Care*, 9(2), 46-53.
- Sariman, G. (2011). Veri Madenciliğinde Kümeleme Teknikleri Üzerine Bir Çalışma: K-Means ve K-Medoids Kümeleme Algoritmalarının Karşılaştırılması. *Süleyman Demirel Üniversitesi Fen Bilimleri Enstitüsü Dergisi*, 15(3), 192-202.
- Sebik, N. B., & Bülbül, H. İ. (2018). Veri Madenciliği Modellerinin Akciğer Kanseri Veri Seti Üzerinde Başarılarının İncelenmesi. *TÜBAV Bilim Dergisi*, 11(3), 1-7.
- Sezgin, E., & Çelik, Y. (2013). Veri Madenciliğinde Kayıp Veriler İçin Kullanılan Yöntemlerin Karşılaştırılması. *Akademik Bilişim Konferansı*, 23-25.
- Steinley, D. (2006). K-Means Clustering: A Half-Century Synthesis, *British Journal of Mathematical and Statistical Psychology*, 59(1), 1-34.
- Tsumoto, S., Kimura, T., Iwata, H., & Hirano, S. (2017). Mining Text For Disease Diagnosis. *Procedia Computer Science*, 122, 1133-1140.
- Thabtah, F., Kamalov, F., & Rajab, K. (2018). A New Computational Intelligence Approach to Detect Autistic Features For Autism Screening. *International Journal of Medical Informatics*, 117, 112-124.
- Thabtah, F. (2017). Autism Spectrum Disorder Screening: Machine Learning Adaptation and Dsm-5 Fulfillment. *In Proceedings of the 1st International Conference on Medical and Health Informatics*, 1-6.
- Thabtah, F. (2018). An Accessible and Efficient Autism Screening Method For Behavioural Data and Predictive Analyses. *Health Informatics Journal*, 25(4), 1739-1755.
- Tekin, R., Kaya, Y., & Tağluk, M. E. (2011). K-Means ve Ysa Temelli Hibrit Bir Model İle Epileptik Eeg İşaretlerinin Sınıflandırılması. *Elektrik Elektronik Bilgisayar Sempozyumu*, 277-283.
- Torun, Y. T., & İşeri, E. (2018). Bebeklik Döneminde Önemli Bir Psikopatoloji: Otizm. *Türkiye Klinikleri Journal of Child Psychiatry-Special Topics*, 4(1), 48-55.
- Tsai, C. Y., & Chiu, C. C. (2008). Developing A Feature Weight Self-Adjustment Mechanism For A K-Means Clustering Algorithm. *Computational Statistics & Data Analysis*, 52(10), 4658-4672.
- Tiryaki, A. E., & Kazan, R. (2007). Bulaşık Makinesinin Bulanık Mantık İle Modellenmesi. *Mühendis ve Makine*, 48(565), 3-8.
- Tucker, C. S., Kim, H. M., Barker, D. E., & Zhang, Y. (2010). A Relief Attribute Weighting and X-Means Clustering Methodology For Top-Down Product Family Optimization. *Engineering Optimization*, 42(7), 593-616.
- Uğur, Ç., Sertçelik, M., Üneri, Ö., Dinç, G. Ş., Sekmen, E., & Solmaz, E. (2018). Evaluation of Serum Urotensin-II Levels of Children with ADHD and Autism Spectrum Disorder. *Anadolu Psikiyatri Dergisi*, 19(1), 80-86.
- Uğur, Ç., & Göker, Z. (2018). Bir Çocuk Psikiyatrisi Ünitesine Başvuran Olgularda Otizm Spektrum Bozukluğu: Demografik ve Klinik Bulgular, *Türkiye Çocuk Hastalıkları Dergisi*, 13(3), 177-183.
- Van den Bekerom, B. (2017). Using Machine Learning For Detection of Autism Spectrum Disorder. *In Proc. 20th Student Conf*, 1-7.
- Yavuz, Ü., Ekim, U. & Köklü, M. (2011). Üniversite Öğrencilerin Ortak Zorunlu Derslerdeki Başarılarının K-Means Algoritması İle İncelenmesi. *NWSA: Engineering Sciences*, 6(1), 342-347.
- Yavuz, S., & Deveci, M. (2012). İstatistiksel Normalizasyon Tekniklerinin Yapay Sinir Ağına Performansına Etkisi. *Erciyes Üniversitesi İktisadi ve İdari Bilimler Fakültesi Dergisi*, 40, 167-187.

Keşifsel Veri Analizi Kullanarak Vakum Şarj Test Süresinin Azaltılması

Reducing Vacuum Charge Test Time Using Exploratory Data Analysis

Tuğba Demircan Kutluay¹ , Süleyman Eken² 



¹(Yüksek lisans Öğrencisi), Kocaeli Üniversitesi, Teknoloji Fakültesi, Bilişim Sistemleri Mühendisliği, Kocaeli, Türkiye

²(Doç Dr.), Kocaeli Üniversitesi, Teknoloji Fakültesi, Bilişim Sistemleri Mühendisliği, Kocaeli, Türkiye

ORCID: T.D.K. 0000-0002-7350-6585;
S.E. 0000-0001-9488-908X

Corresponding author:

Süleyman EKEN
Kocaeli Üniversitesi, Teknoloji Fakültesi, Bilişim Sistemleri Mühendisliği, Kocaeli, Türkiye
E-mail address: suleyman.eken@kocaeli.edu.tr

Submitted: 07.03.2021

Revision Requested: 30.04.2021

Last Revision Received: 03.05.2021

Accepted: 04.05.2021

Published Online: 29.06.2021

Citation: Demircan Kutluay, T. ve Eken, S. (2021). Keşifsel veri analizi kullanarak vakum şarj test süresinin azaltılması. *Acta Infologica*, 5(1), 187-196. <https://doi.org/10.26650/acin.892586>

ÖZ

Günümüzde giderek artan rekabet koşulları ve pazarda kalıcı olabilmek gibi dış faktörler firmaların sistemlerini daha verimli kullanmaya, maliyetlerini azaltmaya, yüksek kalite düzeyinde ürün üretmeye yönlendirmektedir. Firmaların yaptıkları bu çalışmalarda proses iyileştirme adımları önem taşımaktadır. Bu çalışmada üretim süreçleri keşifsel yolla analiz edilerek darboğaz oluşturan istasyon tespit edilmiştir. Keşifsel veri analizi ile kalıpları keşfetmek, anormallikleri tespit etmek, özet istatistikler ve grafik gösterimler yardımıyla varsayımları kontrol etmek için veriler üzerinde ilk araştırmaları gerçekleştirmek mümkün olmuştur. Vakum test istasyonunun katma değersiz test süresi belirlenerek proses süresi azaltılmıştır. Belirlenen yeni test süresi ile üretim kapasite artışı ve hattın daha verimli çalışması sağlanmıştır. Bu sayede günlük üretim adedi 136'dan 188'e, hat verimliliği %55'ten %77'ye artmış ve istasyonların birbirini beklemesi %35'ten %10'a düşmüştür. Aynı zamanda hattın yıllık üretim kapasitesi %27 artarak 58K dan 79K adede ulaşmıştır.

Anahtar kelimeler: Keşifsel Veri Analizi, Grafikselsel Teknik, Tanımlayıcı İstatistik, Verimlilik, Kapasite Artışı, Plant Simulation

ABSTRACT

Nowadays, external factors such as increasing levels of competition and having a permanent place in the market are directing companies to use their systems more efficiently, to reduce their costs, and to produce higher quality products. Process improvement steps are important in studies of these companies. In this study, the vacuum test station that created a bottleneck was determined by analyzing the production process in an exploratory way. Through exploratory data analysis, it was possible to perform an initial data investigation to discover patterns, spot anomalies, test hypothesis, and check assumptions with the help of summary statistics and graphic representations. By determining the value-added test time of the vacuum test station, the process time was reduced. Using the new test period, production capacity increased and a more efficient operation of the line was achieved. The number of daily production increased from 136 to 188, line efficiency increased from 55% to 77%, and the stations waiting for each other decreased from 35% to 10%. At the same time, the annual production capacity of the line increased by 27%, reaching from 58K to 79K.

Keywords: Exploratory Data Analysis, Graphical Technique, Descriptive Statistics, Productivity, Capacity Increase, Plant Simulation

1. GİRİŞ

Keşifsel veri analizi, analitik endüstrisinde EDA olarak kısaltılan temel bir veri analizi tekniğidir ve analitik projesinin ilk aşamasında uygulanan çeşitli kavramlar ve en iyi uygulamalarla ilişkilendirilir. EDA, veri modellerini ve karşılaştırmalı veri analizini tanımlamak için grafik görselleştirme teknikleriyle ilişkili olup veri bilimi projeleri için özellik mühendisliği ve özellik seçim süreçlerinde tercih edilen bir tekniktir (Martinez vd., 2017; Sahoo vd., 2019). Yaygın olarak kullanılan EDA tekniklerinden bazıları tek değişkenli analiz, iki değişkenli analiz, çok değişkenli analiz, çubuk grafik, kutu grafiği, pasta karat, çizgi grafik, frekans tablosu, histogram ve saçılma grafiklerdir (Eken, 2020a; Eken, 2020b). EDA, makine öğrenimi modellerini tamamlayacak olan veri hazırlama aşaması için çok kullanışlıdır ve ekonomiden (Cleff, 2014) çevreye (Velleman, 1981), sağlıktan (Tummers vd., 2020) üretime (Chiarini vd., 2020) birçok alanda kullanılmıştır.

Mevcut durumda vakum test istasyonunda buzdolabına uygulanan vakum test süresi 7 dakikadır. Çevrim süresi en yüksek istasyon olan vakum test istasyonu, test süresi üretim kapasitesini belirlemekte ve dar boğaz oluşturan nokta olarak tespit edilmektedir. Dar boğaz oluşturan istasyonun test süreleri analiz edilerek optimum test süresinin kaç dakika olduğu belirlenmek istenmiştir. Bu çalışmada her bir cihaza uygulanan vakum test süreleri EDA ile incelenerek farklı modellerdeki dolabın istenen optimum basınç değerine kaçınıcı sürede ulaştığı araştırılmıştır. Test süresi verileri ile yapılan analiz sonucunda vakum test süresinin azaltılması amaçlanmıştır. Takip eden paragraflarda vakum teknolojilerini içeren literatürdeki bazı çalışmalara değinilmiştir.

Öztürk vd. (2011) vakumlu soğutma teknolojisinin; yapraklı sebzelerin, mantarların, fırınlanmış ürünlerin ve pişmiş yiyeceklerin önceden soğutulmasında yaygın olarak uygulanmasını anlatmışlardır. Vakumlu soğutma, nemli ve gözenekli ürünler için hızlı bir buharlaşmalı soğutma tekniğidir. Bu çalışmanın amacı, vakumlu soğutma tekniğini semizotunun soğutulmasında uygulamak ve basınç etkisini (0,7 kPa, 1 kPa, 1,5 kPa için) soğutma süresi ve sıcaklık düşüşü üzerindeki etkisini göstermektedir. Çalışmanın diğer bir amacı, geleneksel soğutmaya (buzdolabıyla soğutma) vakumlu soğutma ile karşılaştırmaktır. Geleneksel soğutmanın, 0,7 kPa basınçta vakumlu soğutma ile 6 °C'lik ortam sıcaklığı ile karşılaştırılması ve vakumlu soğutmanın, semizotu için geleneksel soğutmadan yaklaşık 10 kat daha hızlı olduğunu göstermektedir. Vakumlu soğutma için kütle kaybı oranı, geleneksel soğutmaya göre daha yüksektir. Bununla birlikte vakumlu soğutma için soğutma süresi, geleneksel soğutmaya göre daha kısadır.

Rennie vd. (2000) marulun vakumla soğutulması için ürün soğutma yükü ile vakum hızı arasındaki ilişkiyi anlatmışlardır. Basınç fonksiyonunun suyun buharlaşma sıcaklığına ve vakumlu soğutucudaki basıncın düştüğü orana dayalı teorik bir ilişki geliştirilmiştir. Farklı vakum oranlarından toplanan veriler kullanılarak, sıcaklık-zaman verilerinin maksimum eğimi belirlenmiş ve maksimum ürün soğutma yükünü hesaplamak için kullanılmıştır. Vakum uygulama hızı düştüğünde gecikme azalırken vakum hızı yüksek olduğunda ısı transfer gecikmesi olduğu görülmüştür.

Zhao vd. (2020) siklopentan yalıtımlı iki özdeş buzdolabı, biri iki yan duvarda vakum yalıtım paneli (VIP'ler) ve diğeri olmadan karşılaştırma yaparak incelemişlerdir. Bulgular VIP'lerin, daha düşük ortalama güç girişi ve daha uzun kompresör ile kapalı kalma süresi nedeniyle buzdolabının kararlı hal enerji tüketimini % 12,4 oranında azalttığını göstermektedir. Sonuç olarak, yalnızca dondurucu bölmesi süresi artar ve taze yiyecek bölmesi sönmüleme süresindeki azalmayı telafi eder. Bu da genel olarak benzer bir kompresörün açık kalma süresi ile sonuçlanır. Termal yük aktarım işleminin tam bir görünümünü elde etmek için her iki durumda da buharlaştırıcı aracılığıyla ısı transfer işleminin simülasyon sonuçları sağlanır.

Gang ve Xianping (2005) vakumlu termal izolasyon plakasının konfigürasyonunu ve kullanım özelliklerini analiz etmişlerdir. BCD-186CHS tipi buzdolaplarında kullanılan vakumlu ısı izolasyon plakası ve izolasyon tabakaları dağılımı anlatılmış, ayrıca poliüretan ve bileşik ısı izolasyonu, ısı kaybı ve uygulanan bu iki malzemenin enerji tüketimi karşılaştırılmıştır. Deneysel olarak, bileşik izolasyon kullanıldığında ısı kaybının %12,6 ve enerji tüketiminin %11,4 oranında tasarruf sağlandığı görülmüştür. Bu çalışmada, uygulama modları, uyarıları ve vakumlu ısı yalıtım levhasının buzdolabına etkisi ortaya koyulmuştur. Vakumlu ısı yalıtım levhası olasılığı açıklanmıştır.

Tao vd. (2004) vitrinli buzdolabının enerji verimliliğine ilişkin her zamankinden daha katı düzenlemelere yanıt olarak enerji kullanımını en aza indirmek için bir prototip tasarlamıştır. VIP'ler ve vakumlu cam (VG) kullanılarak yalıtım iyileştirmeleri

gibi seçenekler, vitrinli buzdolabında ısı kaybı ölçülerek incelenmiştir. Sonuçlar, vitrinli buzdolabına VG ve VIP'lerin aynı anda monte edilmesiyle enerji verimliliğinin %30'a kadar artırılabilceğini göstermektedir.

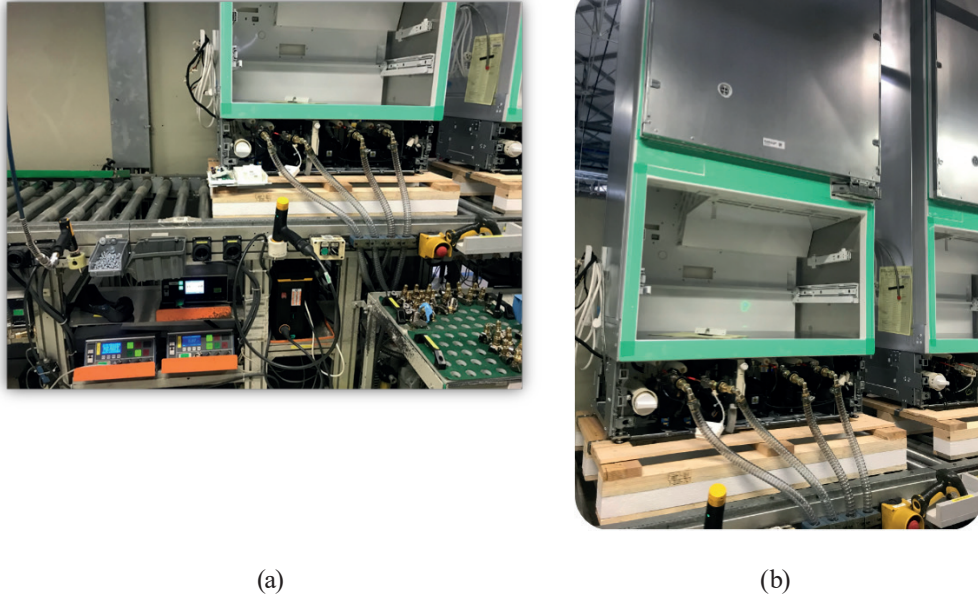
Chen vd. (2010) harici ısı transferli vakumlu termiyonik buzdolabı modeli incelenmiştir. Soğutma yükü ve performans katsayısı (COP) için genel ifadeler, sonlu zaman termodinamiği ve dengesiz termodinamiğin kombinasyonu kullanarak türetmişlerdir. Optimum soğutma yükü ve COP bölgeleri ile ısı rezervuar sıcaklığının cihazın optimum performansına etkisi detaylı sayısal örneklerle analiz edilmiştir. İş fonksiyonunun termiyonik cihazın soğutma yükü ve COP performansları üzerindeki etkileri de incelenmiştir. Burada elde edilen sonuçlar, harici ısı transferini dikkate almadan geleneksel analiz kullananlarla karşılaştırıldığında mevcut analizin gerçek vakumlu termiyonik cihazlar için daha pratik olduğu gösterilmiştir. Sonuçlar, pratik termiyonik buzdolaplarının tasarımı ve uygulaması için kılavuzlar sağlayabilecek şekilde açıklanmıştır.

Misra vd. (2013) mikroskobun, kanal yoğunluğu haritalarını alabilen ve seyreltik buzdolabı sıcaklıklarında ve yüksek (14 T) manyetik alanlarda atom altı uzaysal çözünürlük durumlarında uygulanabilen bir tarama formundaki yapısı ve performansını anlatmışlardır. Tamamen ultra yüksek vakum sistemi, alttan yüklemeli seyreltme buzdolabının sonunda iki numune mikroskobu aşamasına görsel erişim sağlar ve yerinde uçların ve numunelerin transferini kolaylaştırır. İki numuneli aşamanın incelenen numunenin en iyi alanını sağladığı ve testin ömrünü uzattığı görülmüştür. Ayrıntılı olarak anlatılan mikroskobun başarılı termal ankrajı, 250 mK'lık ölçülen elektron sıcaklığı ile 20 mK'lık bir temel sıcaklık okuması ile doğrulanmıştır. Tamamlayıcı titreşim ölçümleri ile birlikte atomik olarak çözülmüş görüntüler, bu cihazdaki titreşim izolasyon şemasının etkinliğini doğrulamıştır. Son olarak mikroskobun, uygulanan bir manyetik alanda hem sıfır alanında hem de taban sıcaklığında spektroskopik haritaları ölçülerek daha fazla soğutmalı tipik makinelerle aynı seviyede performans gösterebileceği açıklanmaktadır.

Sukumar vd. (2015) çalışmalarında günümüz soğutma ve iklimlendirme sistemlerinin insanlar için gerekli hale geldiği ve soğutma sistemlerinin soğutma etkilerini sağlamak için farklı soğutucu akışkanların kullanıldığını belirtmişlerdir. Dolaylı olarak bu soğutucular çevreye zararlı gazlar yayarak birçok etkiye neden olurlar. Kloroflorokarbon gibi gazların atmosfer ve ozon tabakası üzerindeki etkileri incelenmiştir. Bu nedenle çevre dostu soğutma sistemlerine ihtiyaç duyuldu. Geleneksel soğutma sistemlerinin neden olduğu kirlilik, buhar jetli soğutma sisteminin ve mevcut sistemin çalışmasını etkilemiştir. Bu nedenle vakum pompası flaş odası ve flaş odası içerisindeki toplam hava ile çıkarılabilir bir vakum sisteminin oluşturulduğu açıklanmıştır. Vakum sistemi sayesinde su, gizli ısıyı kendisinden alarak kaynama noktasını düşürdüğünde su buharlaşmaya başlar. Burada su düşük sıcaklıklara kadar soğutulur ve bu soğumuş suyun iklimlendirme için kullanılması anlatılır.

Thiessen vd. (2016) buzdolabının termodinamik davranışı deneysel olarak incelenmişlerdir. Buzdolabı içerisindeki vakum yalıtım paneli (VIP) kapsama alanı değiştirilerek kabinin etrafına yerleştirilmiştir. Çalışmalar sonunda farklı VIP düzenlemeleri ile aynı buzdolabı modelinden 16 örnek oluşturulmuş ve kapsamlı ters ısı kaçağı (RHL) ve enerji tüketimi testleri uygulanmıştır. Gözlenen sonuçlarla enerji tüketiminin hem kapsama alanına hem de konuma bağlı olduğu belirtildi. Toplanan veriler statik olarak analiz edilerek panellerin montajı için en uygun alanların belirlenmesi sağlanır. Kapılara ve arka duvara sırasıyla %6 ve %11 enerji tasarruflu paneller monte edilmiştir. Genel enerji tüketim analizinde VIP panellerin %56 kapsama alanının enerji tüketimini %21 azalttığı görülmüştür.

Buzdolabının soğutma çevriminde kullanılan elemanlar; kondansatör, çerçeve ısıtıcısı, arka duvar ısıtıcısı, kurutucu, stop valf, manyetik valf, kılcal boru, eşanjör ve kompresördür. Kompresör, kılcal boru içerisinde basınç farkı yaratmayı sağlar. Kompresörün içinde nem kaçağı olmaması amacıyla basınçlı azot bulunmaktadır (Şekil 1'e bakınız). Kompresörün içerisindeki nem kaçağını önlemek için kompresör ucundaki bakır kılcal boruya valf takılarak vakumlama işlemi yapılır. Mevcut çalışma Fridge ve Freezer iki soğutma kompartımanına sahip buzdolaplarının kompresöre uygulanan vakum test süresinin azaltma çalışmasını kapsamaktadır.



Şekil 1. Vakum şarj test cihazı (a) genel görünüm (b) daha yakın bir görünüm

Vakum test süresinin azaltılması ile vakum test istasyonunun bir vardiyada test edebileceği maksimum dolap sayısı artırılarak vardiyalık üretim kapasite artışının sağlanması amaçlanacaktır. Elde edilen analiz sonucu ile belirlenen yeni vakum test süresine göre vardiyalık test edilen toplam cihaz sayısı, Plant Simulation¹ programı ile simüle edilerek hesaplanmıştır. Yeni test süreleri ile simülasyonda belirlenen vardiyalık toplam cihaz sayısı, hattın üretebileceği maksimum adet sayısını belirleyecektir. Çalışmanın geri kalanı şu şekilde organize edilmiştir: 2. bölümde elde edilen veriseti hakkında bilgi verilmiş, kullanılan yöntemler ve verimlilik çıktıları sunulmuştur. 3. bölümde ise tartışmalara, sonuçlara ve gelecek çalışmalara yer verilmiştir.

2. MALZEME VE YÖNTEMLER

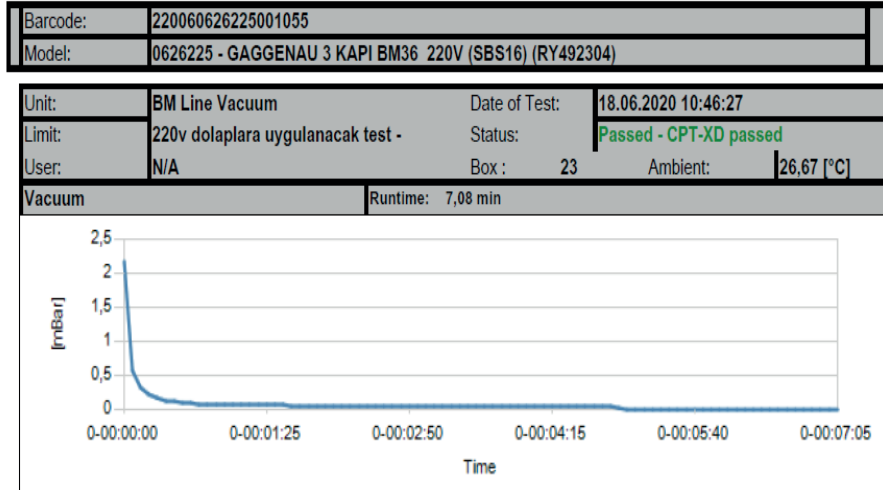
2.1. Veriseti

Veriseti, buzdolaplarının 12760 vakum testi verilerini içermektedir. Test verilerine ait bilgiler tarih, saat, cihaz modeli, hangi kompartımana ait olduğu, cihazın sıcaklığı, kılcal borudaki basınç değeri ve ortam koşuluna ait bilgileri içermektedir. Model sütunu inç bazlı Miele, Thermador, Gaganeu, Bosch, Siemens markalarını içermektedir. Ortam koşulu olarak yaz, kış mevsimi ve bakım öncesi, bakım sonrası test verileri değerlendirilmiştir. Şekil 2'de Gaganeu markasına ait bir test sonucunu yansıtan ölçüm bilgileri gösterilmiştir.

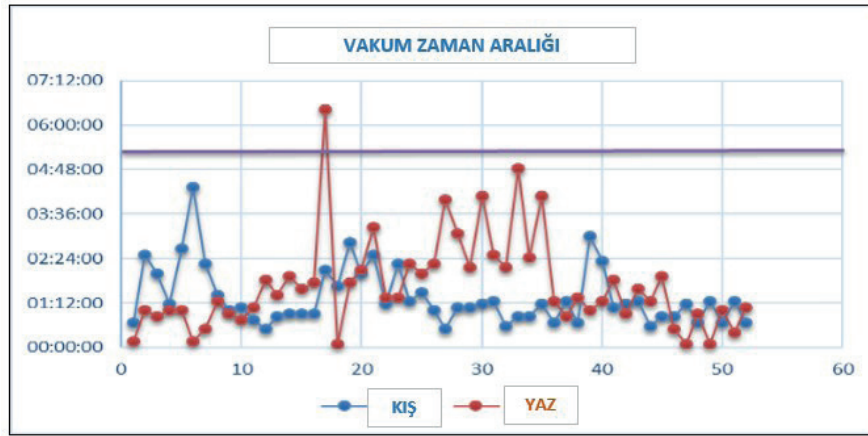
2.2. Vakum Şarj Test Süreleri için EDA

Bu kısımda, veri setinde hangi cihaz modellerinden ne kadar var olduğu, yaz ve kış mevsimleri bazında vakum şarj test sürelerinin araştırılması gerçekleştirilmiştir. Vakum testlerinin analiz sonucuna göre hedeflenen 0,06 bar basınç değerine Şekil 3'te belirtildiği gibi 0-5 dakika arasında ulaşıldığı tespit edilmiştir. Yaz ayına ait bir test verisinin sadece 5. dakikanın üzerinde istenilen bar değerine ulaştığı görülmüş ve bu veri sıra dışı durum olarak değerlendirilmiştir. Sonuç olarak maksimum vakum test süresi de 5 dakika olarak belirlenmiştir. Şekil 4 ise yaz ve kış aylarında 0,06 bar değerine kaç adet cihazın kaçınıcı dakikada ulaştığını göstermektedir.

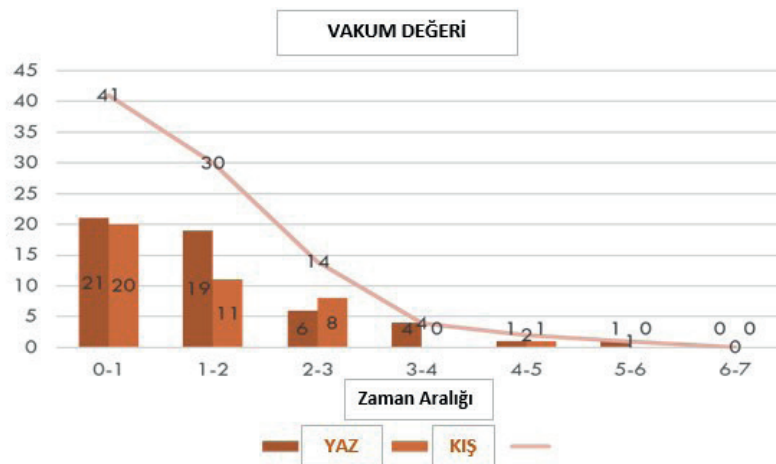
¹ <https://www.dex.siemens.com/plm/tecnomatix/plant-simulation>



Şekil 2. Gaggenau için test sonuçları



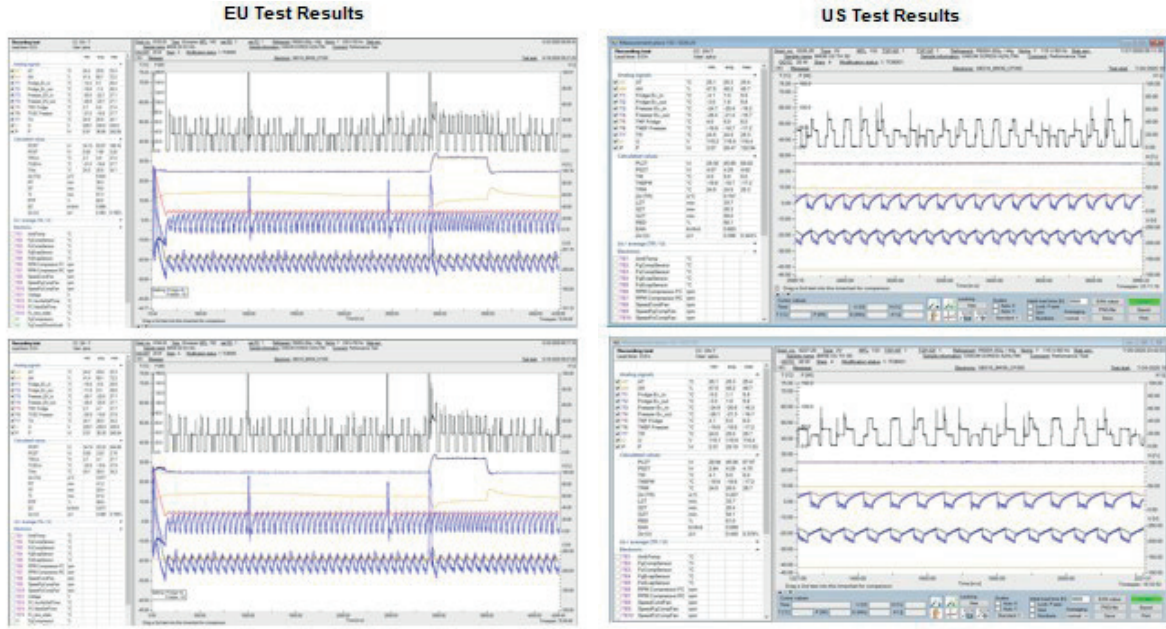
Şekil 3. 0,06 bar değerine göre zaman aralığı



Şekil 4. Vakum zaman aralığına göre 0,06 bar için toplam cihaz sayısı

Python programı kullanılarak yapılan vakum test süreleri EDA çalışmasına göre, yeni test süresi 5 dakika olarak belirlenmiştir. Buzdolabı cihazlarına 5 dakika vakum test süresi uygulandığı durumda test sonuçlarındaki değişim Şekil 5'te gösterilmiştir. Verisetine ve analiz dosyalarına şu linkten erişilebilmektedir:

<https://github.com/suleymaneken/Vacuum-Charge-Test-with-EDA>



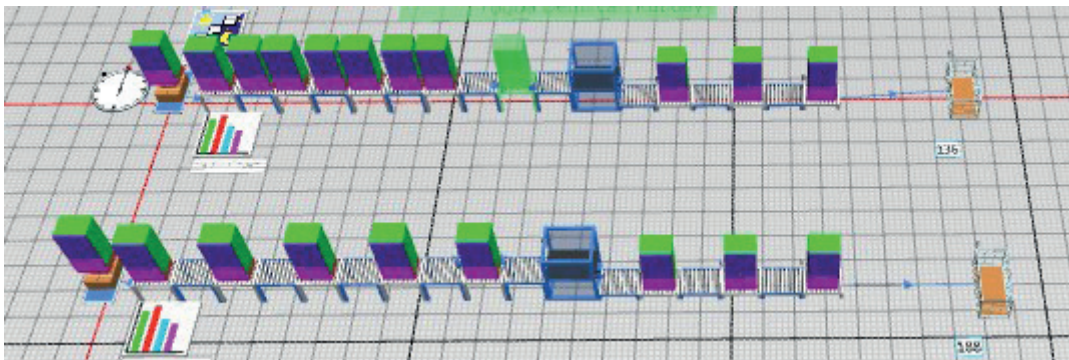
(a)

(b)

Şekil 5. (a) EU modeli (b) US modeli için 5 dakika test uygulama sonuçları

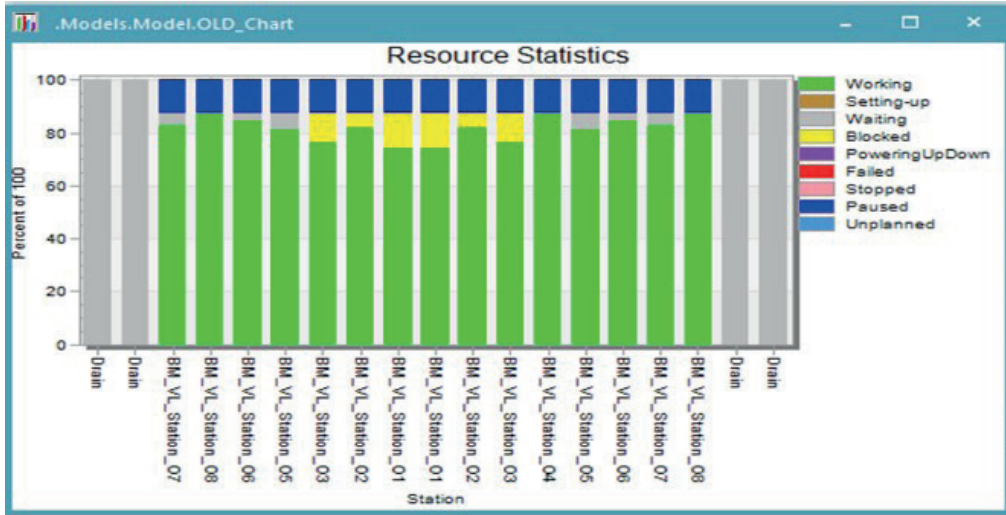
2.3. Plant Simulation ile Vardiya Verimlilik Analizi

Yeni test süresinin, 5 dakika olarak belirlenmesi ile vardiyalık üretim adedinin Şekil 6'da görüldüğü Plant Simulation programı kullanılarak simüle edildiğinde 68'den 94'e çıktığı gözlemlenmiştir.

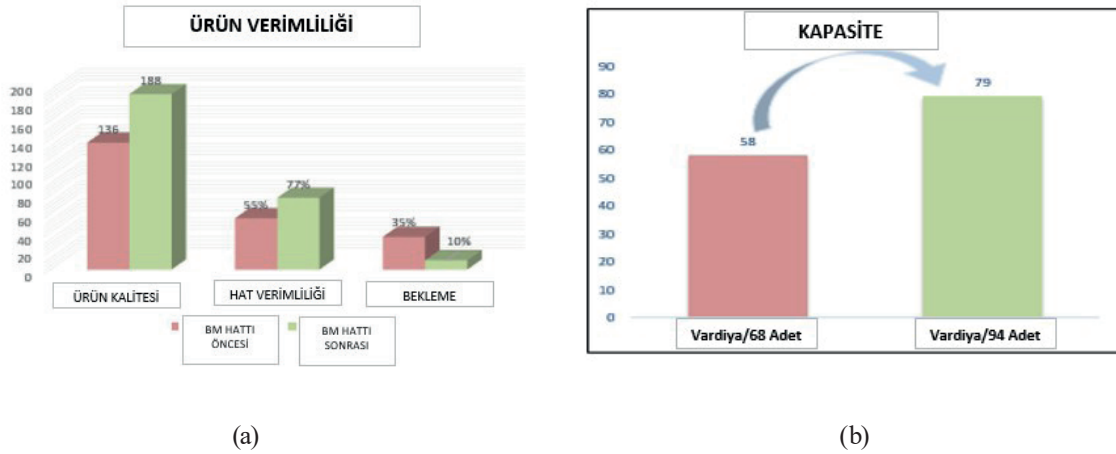


Şekil 6. Plant Simulation programı uygulama görseli

Vardiyalık üretim adedinin 94'e çıkması ile istasyonların birbirini beklemesi ve bloke etme durumunun azalması ile hat içi istasyon çalışma verimliliği artırılmıştır. Şekil 7'de simülasyon görselinde ve grafikte hat içi istasyon verimliliğindeki değişim anlatılmıştır. Şekil 8'de görüldüğü gibi günlük üretim adedi 136'dan 188'e, hat verimliliği %55'ten %77'ye artmış ve istasyonların birbirini beklemesi %35'ten %10'a düşmüştür. Aynı zamanda hattın yıllık üretim kapasitesi %27 artarak 58K dan 79K adede ulaşmıştır.



Şekil 7. Plant Simulation programı kaynak istatistiği



(a)

(b)

Şekil 8. Değişim grafikleri (a) ürün verimlilik (b) kapasite

3. TARTIŞMA VE SONUÇLAR

EDA, bir veri setine ilişkin içgörüyü en üst düzeye çıkarmak, altta yatan yapıyı ortaya çıkarmak, önemli değişkenleri çıkarmak, aykırı değerleri ve anormallikleri tespit etmek, temel varsayımları test etmek, açgözlü modeller geliştirmek ve optimum faktör ayarlarını belirlemek için çeşitli teknikler (çoğunlukla grafiksel) kullanan bir veri analizi yaklaşımı ve felsefesidir. Bu çalışmada yaygın olarak kullanılan EDA tekniklerinden çizgi grafik, histogram, ve özet istatistik (tanımlayıcı) bilgisinden yararlanılmıştır. Kaçınıcı saniyede kılcal borudaki basınç değeri istenilen basınç değerine ulaşıldığının takibi yapılmıştır. Bu çalışmada EDA yöntemleri kullanılarak vakum test süresi azaltılmış, hattın daha verimli çalışması sağlanmış ve üretim kapasitesi artırılmıştır. Vakum süresi azaltılmasaydı hattın talep edilen adette üretim yapabilmesi için 24K€ değerinde yatırım yapılarak ek bir istasyon kurulması ile sağlanacaktı. Bu çalışma sayesinde 24K€ yatırım yapılmadan hedeflenen üretim adedine ulaşılmıştır.

Tablo 1'de ise farklı kriterler altında önerilen çalışmanın diğer çalışmalarla karşılaştırılması verilmiştir. Enerji tüketimi açısından incelendiğinde önerilen çalışma enerji tüketimine odaklanmamıştır. Vakum şarj test süresini optimize ederek ürün verimliliği ve kapasiteye odaklanılmıştır. Dolayısıyla vakum test zamanında bir iyileşme ile odak sağlanmıştır. Vakum yalıtım paneli kullanım yüzeyi çalışmada yan, alt ve üst panellere yerleştirilmiş olup çevresi tamamen kapama veya sadece yan duvar panellerine yerleştirmenin vakum şarj süresine etkisi bakılmamıştır. Bu da test ortamındaki VIP kullanım yüzeyinin standart oluşundan kaynaklanmaktadır. Kompartıman olarak buzdolabı ve dondurucu kısmı ele alınmıştır. Isı kaybı ve kompresör eşanjör kaynak tipi açısından diğer çalışmalarla benzerlik göstermektedir.

Tablo 1

Farklı kriterler altında önerilen çalışmanın diğer çalışmalarla karşılaştırılması

Karşılaştırma kriteri	Önerilen çalışma	Thiessen vd. (2016)	Zhao vd. (2020)
Enerji tüketimi	Değişmedi	%21 azaltıldı	%12,4 azaltıldı
Vakum yalıtım paneli (VIP) kullanım yüzeyi	Yan, alt ve üst panellere yerleştirilmiş	Kabinin etrafına yerleştirilmiş	Yan duvar panellerine yerleştirilmiş
Vakum test zamanı	7 dk'dan 5 dk'ya düşürüldü	Değişmedi	Değişmedi
İncelenen soğutma bölgesi	Buzdolabı ve dondurucu bölmeleri	Buzdolabı, dondurucu ve tam esnek bölmeleri	Buzdolabı ve dondurucu bölmeleri
Isı kaybı	Azaltıldı	Azaltıldı	Azaltıldı
Kompresör ve eşanjör kaynak tipi	Lokring kaynak	Gaz altı kaynak	Lokring kaynak

Çalışmanın iyileştirilme adımlarında ise hatta devam eden %10'luk beklemenin sebebi araştırılacak ve hat içi %27'lik verimlilik kaybının nedenleri sorgulanacaktır. Böylece ürün verimliliği ve kapasitenin daha da artması beklenmektedir.

Hakem Değerlendirmesi: Dış bağımsız.

Çıkar Çatışması: Yazarlar çıkar çatışması bildirmemiştir.

Finansal Destek: Yazarlar bu çalışma için finansal destek almadığını beyan etmiştir.

Yazar Katkıları: Çalışma Konsepti/Tasarım- T.D.K.; Veri Toplama- T.D.K., S.E.; Veri Analizi/Yorumlama- T.D.K., S.E.; Yazı Taslağı- T.D.K., S.E.; İçeriğin Eleştirel İncelemesi- S.E.; Son Onay ve Sorumluluk- T.D.K., S.E.

Teşekkür: BSH Ev Aletleri Sanayi ve Ticaret A.Ş.'ye çalışmanın yapılmasındaki katkılarından dolayı teşekkür ederiz.

Peer-review: Externally peer-reviewed.

Conflict of Interest: The authors have no conflict of interest to declare.

Grant Support: The authors declared that this study has received no financial support.

Author Contributions: Conception/Design of Study- T.D.K.; Data Acquisition- T.D.K., S.E.; Data Analysis/Interpretation- T.D.K., S.E.; Drafting Manuscript- T.D.K., S.E.; Critical Revision of Manuscript- S.E.; Final Approval and Accountability- T.D.K., S.E.

Acknowledgement: We thank BSH Turkey for their contribution to the study.

Kaynaklar/References

- Chen, L., Ding, Z., & Sun, F. (2010). Performance analysis of a vacuum thermionic refrigerator with external heat transfer. *Journal of applied physics*, 107(10), 104507.
- Chiarini, A., Belvedere, V., & Grando, A. (2020). Industry 4.0 strategies and technological developments. An exploratory research from Italian manufacturing companies. *Production Planning & Control*, 31(16), 1385-1398.
- Cleff, T. (2014). Exploratory data analysis in business and economics. *Exploratory Data Analysis in Business and Economics*. <https://doi.org/10.1007/978-3-319-01517-0>.
- Eken, S. (2020a). An exploratory teaching program in big data analysis for undergraduate students. *Journal of Ambient Intelligence and Humanized Computing*, 11(10), 4285-4304.
- Eken, S. (2020b). Büyük Verinin İnteraktif Görselleştirilmesi: Tableau Üzerine Öğrenci Deneyimleri. *Avrupa Bilim ve Teknoloji Dergisi*, (18), 262-271.
- Gang, L., & Xianping, Z. (2005). Research on vacuum thermal isolation plate applied to refrigerator for isolation. *Cryogenics and Superconductivity*, 4.
- Misra, S., Zhou, B. B., Drozdov, I. K., Seo, J., Urban, L., Gyenis, A., ... & Yazdani, A. (2013). Design and performance of an ultra-high vacuum scanning tunneling microscope operating at dilution refrigerator temperatures and high magnetic fields. *Review of Scientific Instruments*, 84(10), 103903.
- Martinez, W. L., Martinez, A. R., & Solka, J. (2017). *Exploratory data analysis with MATLAB*. Florida, FL: Crc Press.
- Ozturk, H. M., Ozturk, H. K., & Kocar, G. (2011). Comparison of vacuum cooling with conventional cooling for purslane. *International Journal of Food Engineering*, 7(6), 1-18.
- Rennie, T. J., Raghavan, G. S. V., & Vigneault, C. (2000). Determination of the product refrigeration load in vacuum cooling of lettuce. *Applied Engineering in Agriculture*, 16(5), 459.
- Sahoo, K., Samal, A. K., Pramanik, J., & Pani, S. K. (2019). Exploratory data analysis using Python. *International Journal of Innovative Technology and Exploring Engineering (IJITEE)*, 8(12), 2019.
- Sukumar, R. S., Gopichand, A., Janardhana, U., Kumar, V. S., & Chanti, B. V. Design and Fabrication of Water Refrigeration System by Creating Vacuum. *International Journal of Innovative Research in Science, Engineering and Technology*, 4(4), 2428-2434.
- Tao, W. H., Huang, C. M., Hsu, C. L., & Lin, J. Y. (2004). Performance study of an energy-efficient display case refrigerator. *Chemical Engineering Communications*, 191(4), 550-565.
- Thiessen, S., Knabben, F. T., & Melo, C. (2016). An Experimental Study on the Use of Vacuum Insulation Panels in Household Refrigerators. *Proceedings of International Refrigeration and Air Conditioning Conference* (pp. 1-8). West Lafayette, IN: Purdue University.

- Tummers, J., Catal, C., Tobi, H., Tekinerdogan, B., & Leusink, G. (2020). Coronaviruses and people with intellectual disability: an exploratory data analysis. *Journal of Intellectual Disability Research*, 64(7), 475-481.
- Velleman, P. F., & Hoaglin, D. C. (1981). *Applications, basics, and computing of exploratory data analysis*. London, Duxbury Press.
- Zhao, R., Qiao, L., Gao, Z., & Huang, D. (2020). Effect of Vacuum Insulation Panels on Energy Consumption and Thermal Load Transfer between Compartments in a Three-Temperature Frost-Free Refrigerator. *Energies*, 13(7), 1559.

Evaluation of Healthcare Economics of OECD Countries: Multi-Objective Statistical Optimization Model

OECD Ülkelerine ait Sağlık Ekonomilerinin Değerlendirilmesi: Çok Amaçlı İstatistiksel Optimizasyon Modeli

Abdulkadir Atalan¹ , Cem Çağrı Dönmez² 



¹(Asst. Prof.), Gaziantep Islam Science and Technology University, Department of Industrial Engineering, Gaziantep, Turkey

²(Asst. Prof.), Marmara University, Faculty of Engineering, Industrial Engineering, Istanbul, Turkey

ORCID: A.A. 0000-0003-0924-3685;
C.Ç.D. 0000-0000-0000-0000

Corresponding author:

Abdulkadir ATALAN
Gaziantep Islam Science and Technology
University, Department of Industrial Engineering,
Gaziantep, Turkey
E-mail address: abdulcadiratalan@gmail.com

Submitted: 05.12.2020

Revision Requested: 10.03.2021

Last Revision Received: 10.03.2021

Accepted: 11.03.2021

Published Online: 27.04.2021

Citation: Atalan, A., & Donmez, C.C. (2021). Evaluation of healthcare economics of OECD countries: Multi-objective statistical optimization model. *Acta Infologica*, 5(1), 197-206. <https://doi.org/10.26650/acin.836372>

ABSTRACT

Healthcare systems comprise the most crucial factors affecting countries economically. In this study, the infrastructure and economic structure of healthcare systems belonging to the Organization for Economic Co-operation and Development (OECD) countries are discussed to evaluate the healthcare expenditures (HE) and healthcare expenditures per capita (pcHE) of OECD members. We have identified factors related to healthcare economics and analyzed the significance ratios of these factors to calculate the feasible values of the factors affecting the healthcare systems in the OECD members by statistical optimization analysis. Within the feasible values obtained by using a multi-objective optimization model (MOOM) because of statistical analysis, the HE of OECD countries was minimized, and an improvement of 29.13% was achieved. The second objective function aimed to maximize the pcHE that was estimated to be at least \$5,282.37 for the OECD members. Consequently, in countries that do not have a social healthcare system, it is perceived that HE amounts are excessive. The fundamental reason for this situation represents the healthcare sector's perception as a business in those countries.

Keywords: OECD Countries, Healthcare Expenditure, Healthcare Expenditure Per Capita, Multi-Objective Optimization Model, Statistical Analysis

ÖZ

Sağlık sistemleri, ülke ekonomilerini etkileyen en önemli faktörlerden biridir. Bu çalışmada, OECD üyelerinin sağlık harcamalarını (SH) ve kişi başına düşen sağlık harcamalarını (kbSH) değerlendirmek için Ekonomik İşbirliği ve Kalkınma Teşkilatı (OECD) ülkelerine ait sağlık sistemlerinin altyapısı ve ekonomik yapısı tartışılmıştır. İstatistiksel optimizasyon analizi ile OECD üyelerindeki sağlık sistemlerini etkileyen faktörlerin uygulanabilir değerlerini hesaplamak için sağlık ekonomisi ile ilgili faktörleri belirlenerek, bu faktörlerin önemlilik oranlarını analiz ettik. İstatistiksel analiz aracılığıyla çok amaçlı optimizasyon modeli (MOOM) kullanılarak elde edilen uygulanabilir değerler içerisinde OECD ülkelerinin SH'si minimize edilmiş ve % 29,13 oranında iyileşme sağlanmıştır. İkinci amaç fonksiyonu ise OECD üyeleri için kbSH'yi maksimize etmek amacıyla en az kbSH değeri 5.282,37 dolar olarak tahmin edilmiştir. Sonuç olarak sosyal sağlık sistemi olmayan ülkelerde SH miktarlarının aşırı olduğu algılanmaktadır. Bu durumun temel nedeni, sağlık sektörünün o ülkelerdeki bir işletme olarak algılanmaktadır.

Anahtar kelimeler: OECD Ülkeleri, Sağlık Harcamaları, Kişi Başına Düşen Sağlık Harcamaları, Çok Amaçlı Optimizasyon Modeli, İstatistiksel Analiz

1. INTRODUCTION

The OECD aims to develop policies to increase the economic and social well-being of people living in the world (OECD 2017e). The 35-member organization is a platform for cooperation between governments to seek solutions to common problems (OECD 2017d). The most critical issue among the issues of these countries is the healthcare-related challenge (Atalan 2018). In today's world, the importance of the healthcare field is increasing. The main reason for this is that people want to live in more developed economies. This shows that the economic development and progress in healthcare are directly proportional. OECD countries have spent about \$6.02 trillion in healthcare expenditures in 2017 (Atalan 2020).

Several studies have been conducted on the healthcare systems of OECD members. Relations between healthcare economics and economic dimensions of the countries have been sought in these studies. In 1992, a study of 19 OECD countries covered the aggregate gross domestic product (GDP) and socio-demographic factors affecting HE (Gerdtham et al. 1992). The OECD countries' PC and GDP data have been used to obtain unit root test results due to time series. In this study, the method used was a panel of OECD data (McCoskey and Selden 1998). In the panel data models, the correlation was developed in which the factors affecting the model are specified in the time-series dimension (Albouy, Davezies, and Debrand 2010). Another study directly examined the long-term economic relationship between HE and GDP as a panel data of the 20 members of OECD for the period 1971-2004 (Baltagi and Moscone 2010).

Using data from OECD countries, some studies have been carried out on the performance that showed the healthcare quality of those countries (Arah et al. 2006). Surveys or verbal dimensions examined in most of these studies were done without using any quality tools. However, to obtain quantitative and tangible findings in healthcare economics, some statistical and engineering techniques were applied in the mentioned studies related to healthcare costs. In particular, generalized linear models were used as the conventional method for reducing the skewness of the data (Malehi, Pourmotahari, and Angali 2015). Furthermore, regression methods were used to find the relationship between the factors and the response variables related to healthcare systems (Carides, Heyse, and Iglewicz 2000; Gregori et al. 2011; Jones, Lomas, and Rice 2015; Mihaylova et al. 2011). This study's goal was to determine the factors affecting healthcare economics and calculate these factors' optimum values using a statistical optimization approach.

Another method used in health studies was the optimization technique (Atalan 2019). Optimization techniques are the required methods used in management, engineering, and business applications. The purpose of using optimization or operation research techniques is to provide maximum benefits (such as revenue, production, etc.) and minimize loss (such as costs, waiting time, etc.). Optimization applications have been widely used in energy, automotive, manufacturing, transportation, and healthcare. It is common for optimization models related to waiting times to be applied in the healthcare field (Atalan and Donmez 2019). Likewise, in other health care issues, for example, the number of patients treated in hospitals, length of hospital stays, hospital resources utilization rates, etc. optimization methods have been tried to solve these areas of concern (Ajami, Ketabi, and MahmoodAbadi 2013; Iii, Hwang, and Nagarajan 2001).

Generally, there is one objective function in optimization models for applications. A model is created by creating a single objective function in MOOM and transforming the remaining objective functions into constraints (Ngatchou, Zarei, and El-Sharkawi 2005). For this reason, developed optimization models are not tempting to use the single-objective function of optimization models for such problems. Among the most important elements that complicate problems are the multi-objective functions (Mason, Duggan, and Howley 2017). It is inevitable to use MOOM to overcome economic complexity. The application of MOOM has not been addressed much in the field of healthcare. MOOM has been used to increase the population homogeneity and diversity of the small regions' medical services and minimize the distance between the micro-regions visited by patients (Steiner et al. 2015). In this study, two objective functions were formulated with the economic aspects of the healthcare systems of OECD members. The first objective function was to maximize the amount of pcHE in OECD countries (Atalan 2018). Secondly, attention was paid to minimizing the HE of OECD members.

In the research carried out so far, healthcare economics factors are considered to be demographic or social justice in nature (Dukhanin et al. 2018). However, the fact that the factors are quantitative and correct according to their values provides

convenience in statistical analysis. During the economic crisis in Europe in 2010, the healthcare systems belonging to European countries were affected (Atalan, Cinar, and Cinar 2020). European governments tended to reduce health expenditures during the crisis, but in some countries, which did not provide healthcare insurance policies, healthcare pocket payments increased. In support of this action, we advocate that healthcare economics should not be affected by the economic fluctuations of countries. Moreover, it has been stated that healthcare expenditures are increased by reducing the number of healthcare workers and hospitals in the healthcare systems in the time of economic crisis (Palasca and Jaba 2015).

This study examined the healthcare infrastructure of OECD members and applied the multi-objective statistical optimization method needed in order to have a better healthcare system. The five input variables (or factors) and two output variables (defined as objective functions of the multi-objective optimization model) that affect healthcare systems were expressed for this study. There were two stages of statistics and mathematical programming used in the present study. First, the variables of the healthcare system were statistically analyzed and then included in the optimization process. The reason for using MOOM in this study is that there are pcHE and HE variables defined as two different objective functions. The optimum values of decision variables and objective functions were calculated for OECD members to have a better healthcare system.

The present study consists of four sections. A literature review explaining about healthcare systems and healthcare economics is in the first section. The second part includes the factors affecting the healthcare systems of OECD members that form the methodology of this study. The results of the numeric example are considered in the third section. The conclusion of the study is discussed in the last section.

2. METHODOLOGY

2.1. Defining Variables of Independent and Responses

In terms of economy, many factors are influential in the healthcare system level headedly. Before determining the factors of healthcare economics, it is necessary to examine healthcare systems worldwide. There are two types of healthcare systems worldwide: private and state healthcare systems (Liu et al. 2018). Social countries have more health expenditures than non-social countries (Muyl, Dumas, and Herbert 2004). However, private healthcare systems have developed in non-social states such as the U.S. healthcare system. This is because most of the HE is covered by patients or those with health insurance in these countries.

This part of the research contains information on the data and methods to be used for healthcare economics. The healthcare economics of OECD members along with seven fundamentals: gross domestic product (GDP), gross domestic product per capita (pcGDP), life expectancy (LE), the total numbers of healthcare employments (E) of a member of OECD, and pharmaceutical expenditure (PE) were defined as a factor or independent variables and healthcare expenditure (HE) and healthcare expenditure per capita (pcHE) were assigned as a response or dependent variables in this study. The statistics optimization analysis calculates the significance ratings and feasible values of these factors. The responses and factor variables are examined one by one in Table 1.

Table 1
Indicators of OECD's healthcare systems

Variables	Status	Notations	Description
Healthcare Expenditure	Response (Output) Variable	y_h	HE (% of GDP) of OECD countries was considered (Bichay 2020; OECD 2016a). The general belief, as people's quality of life and their willingness to live increases, HE also increases. The optimization model is provided to minimize the amount of HE, which is the objective function.
Healthcare Expenditure Per Capita	Response (Output) Variable	y_{pc}	The second response, which is influenced by the factors considered, is the pcHE (OECD 2016b). The second optimization model stated that the pcHE should be maximum, and the feasible values of the factors constituting the constraints should be in this study.
Gross Domestic Product	Factor (Input) Variable	x_g	The GDP factor is assumed to affect an economy on HE and pcHE (WBG 2017). Countries with large GDP ratios can be thought of as determining the proportion of their budget for their healthcare systems.
Gross Domestic Product Per Capita	Factor (Input) Variable	x_{pc}	Among the factors that can affect healthcare economics, it is considered the pcGDP factor (WBG 2017).
Pharmaceutical Expenditures	Factor (Input) Variable	x_p	PE is thought to be particularly effective in HE (OECD 2017c). The average of OECD countries' PE is 1.4 % of GDP.
Life Expectancy	Factor (Input) Variable	x_l	LE data of OECD countries have been taken into consideration (OECD 2017a). The average LE is 80/65 in OECD countries (C. and B. 1977).
Total number of Health and Social Employment	Factor (Input) Variable	x_e	Total Health and Social Employment of countries are considered in OECD members (OECD 2017b).

The method used in this study is an improved linear regression equation with multiple independent variables. The healthcare economics involved are expressed by the following general eq. 1 to measure the contribution of the HE and pcHE for OECD countries.

$$y_i = \beta_0 + \beta_j x_{ij} + \epsilon_i \tag{1}$$

where; the value of y_i represents the response value, the constant values estimate the parameters β_0 , the fitted values estimate the parameters β_j , the parameter x_{ij} describes the factors, $i = \{\text{defined independent or input variables}\}$ and $j = \{\text{countries or OECD members}\}$.

2.2. Multi-Objective Optimization Models (MOOM)

The MOOM is often used with common constraints and applied to find the best balance between two different objective functions (To be minimized or maximized). Mathematical explanations (a general formulation) of MOOM are as follows (Deb and Kalyanmoy 2001);

$$\begin{aligned}
 &\text{Min or Max } f_j(x), j=1,2,\dots,J \\
 &g_a(x) \geq 0, a = 1,2,\dots,A \\
 &h_b(x) = 0, k=1,2,\dots,B \\
 &x_i^{(L)} \leq x_i^{(U)}, i=1,2,\dots,n
 \end{aligned} \tag{2}$$

where; L represents the lower bound, and U expresses the upper bound. In single-objective optimization models, the superiority ratios of a solution are compared to those of other solutions. That is, the values of the target function can be easily reached. However, in a multi-objective optimization problem, a solution is feasible determined by dominance. This optimization model is used to produce the Pareto boundary of the two objective functions most effectively. The feasible solution range is created by expressing the nonlinear line in which the max-max, min-min, or max-min in the Pareto graph is best described.

There are two different responses at the center of healthcare economics by examining the data in the study. In the mathematical optimization models created, there are two objective functions. For this reason, a MOOM was applied in this study. These objective functions are to maximize the HE and minimize the pcHE. The same constraints of factors are used for both objective functions.

As a result of the regression analysis, there are some constraints related to the derived objective functions. These constraints are composed of the factors affecting healthcare economics, and these factors have upper and lower limits (see equations 5, 6, and 7). The generated optimization models (see equations 3 and 4) are expressed as follows:

$$\text{Maximize}_{y_h} \beta_0 + \beta_1 x_{gj} + \beta_2 x_{pcj} + \beta_3 x_{lj} + \beta_4 x_{pj} + \beta_5 x_{ej} + \epsilon_i \quad (3)$$

$$\text{Minimize}_{y_{pc}} \beta_0 + \beta_1 x_{gj} + \beta_2 x_{pcj} + \beta_3 x_{lj} + \beta_4 x_{pj} + \beta_5 x_{ej} + \epsilon \quad (4)$$

subject to

$$l \leq x_{ij} \quad (5)$$

$$u \geq x_{ij} \quad (6)$$

$$0 \leq x_{ij} \quad (7)$$

3. RESULTS AND DISCUSSIONS

The correlation of the data used with each other was examined in Table 2. The correlation values demonstrate factors that strongly correlate with which factor or which responses. The results obtained from the correlation calculations need to have a value between -1.00 and 1.00—the closer to 0, the less correlated between the data. However, if the value obtained is close to 1.00 or -1.00, it is understood that there is a strong relationship between the data.

Table 2
Correlation relationship among variables

Variables	x_g	x_{pc}	x_l	x_e	x_p	y_{pc}
x_{pc}	0.162					
x_l	-0.023	0.598				
x_e	0.793	0.089	0.103			
x_p	0.376	-0.408	-0.252	0.365		
y_{pc}	0.528	0.883	0.500	0.376	-0.120	
y_h	0.681	0.480	0.431	0.540	0.244	0.794

The effect of data used in this analysis on the response was measured. The primary purpose of using this analysis was to determine which regression analysis was appropriate for the data used. Thirty-five countries are members of the OECD, i.e., fitted regression analysis was required to measure the data from five different factors and two response variables with appropriate analysis. The statistical analysis accuracy was 72.16 % of the R^2 value and 67.36 % of the adjusted R^2 value for HE and 93.95 % of the R^2 value and 92.90 % of the adjusted R^2 value for pcHE. The results obtained by the regression analysis were rated as necessary (see Table 3).

Table 3
Analysis of variance of regression analysis for HE and pcHE

Factors	HE		pcHE	
	Probability	Status	Probability	Status
x_g	0.001	Significant <0.05	0.002	Significant <0.05
x_{pc}	0.071	Provisional Significant <0.05	0.005	Significant <0.01
x_l	0.010	Significant <0.05	0.566	Insignificant <0.1
x_e	0.444	Insignificant <0.1	0.052	Insignificant <0.1
x_p	0.057	Significant <0.05	0.079	Provisional Significant <0.1

The effect of 3 out of 5 factors were compared to the others on the response variable HE. In short, LE and E factors were less affected by HE. Especially the effect of the GDP (M) and pcGDP factors of the OECD countries on the healthcare economy seemed to be excessive. Another factor influencing healthcare economics was the amount of expenditure for the pharmaceutical sector. In the regression analysis, except for the factors used for HE, except LE, the others affected the HE

variable. E was added to the model even though the effect is small. Although the p-value of PE is close to 0.05, the value of the model’s impact is 49225. Besides, as the factors are handled one by one, their interaction needs to be considered. In particular, the interaction of the factors GDP and E has been shown to affect response (p-value is 0.013) significantly. Likewise, the value of the interaction of the factors GDP and PE is close to 0.05, which is embedded in the model. However, the interaction of the pcGDP factor itself with the regression equation’s nonlinear model has been removed from the model due to its weakness (p-value is 0.071; we can accept this factor as conditional).

According to the statistical analysis, pcHE was defined as a second objective function, GDP, and pcGDP factors have a significant effect on the response variable. The E factor impact on pcHE is based on substantial evidence, but it is accepted that if the significance level ratio is close to 0.05. PE has a significant effect, even if it has weak evidence. Negatively, the LE factor has not counted any effect on the response variable. The regression coefficients of all factors except factor E have a positive value. The constant number, called the regression coefficient (known as the constant coefficient that gives the regression slope), has a negative value, which is -0.233.

The two regression equations obtained at the same time constitute the objective functions of this study. The first objective function was evaluated for HE (see equation 8). The second regression equation (defined as the second-objective function) belongs to pcHE (see equation 9). The equations were defined as follows:

$$\text{Maximize}_{y_h} [-2363+ 0.000250*x_{gj} + 0.07974 *x_{pcj} +30.1 *x_{lj} - 0.000024 *x_{ej} + 49225 *x_{pj}] \tag{8}$$

$$\text{Minimize}_{y_{pc}} [-0.233+ 0.000012 *x_{gj} + 0.000010 *x_{pcj} +0.00358 *x_{lj} - 0.00001 *x_{ej} + 1.334 *x_{pj}] \tag{9}$$

$$x_{ij} \geq 0.00; i=\{g,pc,l,e,p\}, j=\{\text{OECD members}\} \tag{10}$$

The same constraints were used for the objective functions (see equations Table 4). A limit was established by calculating the maximum (upper limit) and minimum (lower limit) values of the results of these constraints. Finally, these factors that make up decision variables that need to be greater than zero (see equation 10).

Table 4
The constraints of the objective functions

Variables	Lower Limits	Upper Limits
x_{gj}	\$17,180.0	\$18,624,475.00
x_{pcj}	\$8,209.00	\$100,573.00
x_{lj}	74.60	83.90
x_{ej}	5.00 (% of GDP _j)	22.00 (% of GDP _j)
x_{pj}	20944	20076000

The values to be taken by the factors must satisfy the objective functions. The most crucial factor to be considered for this is the desirability values of the objective functions. The desirability value ranges from 0 to 1. In this study, the desirability value obtained to satisfy both objective functions are 0.5552. This value is the average of the desirability values of the objective functions. The values show the data required for an OECD country to have a quality structure in the healthcare economy in Table 5. Countries with OECD members must have a minimum GDP of \$1,737,253,440.00 and a pcGDP of approximately \$54,000.00. Among findings, members of OECD with these values were also required to reduce pharmacy spending.

Table 5
Feasible values of objective functions and factors

Definitions	Status	Variables	Desirability	Feasible Values
Decision Variables or Input Factors	Constraints 1	x_{gj}	-	\$1,737,253,440
	Constraints 2	x_{pcj}	-	\$53,970.24
	Constraints 3	x_{lj}	-	79.46
	Constraints 4	x_{ej}	-	306207.90 ~ 306208
	Constraints 5	x_{pj}	-	1.190 %
Maximize	Objective Functions 1	y_{pc}	0.4768	\$5,282.38
Minimize	Objective Functions 2	y_h	0.6394	8.970 %

The OECD members must set a budget of at least \$5,282.38 to maximize the pcHE value. The objective function of HE, which is required to be drawn to the minimum level, requires members of OECD to allocate approximately 9.00 % of their GDP budget. For possible values for both objective functions and constraints, the feasible HE of OECD member countries for healthcare have been recalculated with the new method. Under this calculation, some countries spend unnecessarily excessive expenditures. Some countries have found that the percentage of HE in the amount of GDP they have is required to be increased. Seventeen out of 35 OECD member countries spend more on HE. The HE amounts in the U.S. are unusually excessive. Approximately 47.89 % of the HE amounts are seen to pay more. Turkey is experiencing just the opposite in this case. Turkey appears to be insufficient, and it was found that HE should increase at least 107.16 % (It is necessary to spend as much as the current HE).

The feasible HE PC value deducted by the new method developed is calculated as \$5,282.38 (threshold) on average. It is recommended that pcHE values be maintained or increased in countries (France, Ireland, Luxembourg, Netherlands, Norway, Sweden, Switzerland, and the U.S.) where HE PC values are above feasible. Because HE PC is the best measure of the quality of the health system. According to computational results, 27 out of the 35 OECD countries need to increase the pcHE amount significantly. Otherwise, they will not have a good health system in the future. Especially in two countries, Mexico and Turkey are required to provide an increase of approximately \$4,000.00.

The resulting feasible values and an improvement in the health economics of OECD countries have been achieved with this study. The budget allocated by members of the OECD for healthcare services is approximately \$6.02 trillion. The total amount that OECD members will spend on the healthcare sector is about \$4.3 trillion within the feasible results obtained with this developed application. Within the scope of the targeted values in health economics, HE has been minimized, and an improvement of approximately 29.13% has been achieved. This improvement will result in wasted expenditure on other investments in the healthcare field. The most crucial incentive factor for this situation is an overpayment in the pharmaceutical sector. The workplace created for the pharmaceutical industry causes people to turn from healthcare services to economic competition in developed and developing countries. The most active component in healthcare systems is the governments at the center of healthcare systems as a rule-maker in healthcare (Atalan and Donmez 2019). For this reason, governments should not allow the pharmaceutical industry to put people's lives at risk.

The second indicator of the healthcare economy is the pcHE. An increase in the amount of pcHE in world countries is requested. The main reason for this is that people want to live under health insurance. The size of the pcHE amount also indicates the quality of the healthcare systems that the countries have. It shows that countries with high HE PC amounts have an advanced healthcare system (technology, quality, adequate resources, etc.). With this method developed, the amount of pcHE in countries should be over \$5,000.00. It is a known fact that people in countries that do not have a social healthcare system spend too much money on treatment. For this reason, people want to get quality healthcare services if they have certain state or public healthcare insurance (the amount paid or the amount cut off by the government for the insurance premium). However, to provide quality healthcare, members of the OECD need to have the feasible values obtained.

4. CONCLUSIONS AND FUTURE PERSPECTIVES

Healthcare systems are one of the main elements that constitute countries' infrastructures, although the healthcare systems have a complex structure. The healthcare sector is among the largest sectors after the manufacturing industry in the world. In countries with economic developments, it is not surprising that developed healthcare systems exist. The fact that people have a high financial advantage in healthcare has opened the way to get quality healthcare services. There will inevitably be large investments economically to create quality healthcare services. However, to achieve this desired situation, the healthcare economy needs to be appropriately managed.

This study's central theme was to determine feasible results by selecting the effects of the factors that may be effective in healthcare economics on HE and pcHE. To manage healthcare economics, governments need to use the budgets they allocate efficiently. Consequently, in countries that do not have a social healthcare system, the HE is excessive. The main reason for this situation is the perception of the health sector as a business. With this study, approximately 29.13% improvement was

achieved in HE, thereby eliminating waste that occupies the healthcare budget. The other goal of this study was to ensure that the pcHE amount is high in countries that are members of OCED. On average, the pcHE amount was \$ 4,002.68. However, it was determined that the pcHE amount of OECD countries should be over \$5,200.00 with this advanced method. Thus, we have underlined that the quality and infrastructure of healthcare should be improved in these countries.

The applicability of optimization models with statistical analysis has been shown by contributing to the literature with this study. Additionally, this study emphasizes how to solve MOOM as opposed to single-objective optimization models. Researchers generally think that MOOMs are difficult to solve, and they come up with a solution by converting optimization models with multiple objective functions into a single objective function. As the future scope of this study, it is proposed to integrate statistical analysis with optimization models to find the optimal economic and social parameters for countries, cities, or institutions.

Hakem Değerlendirmesi: Dış bağımsız.

Çıkar Çatışması: Yazarlar çıkar çatışması bildirmemiştir.

Finansal Destek: Yazarlar bu çalışma için finansal destek almadığını beyan etmiştir.

Yazar Katkıları: Çalışma Konsepti/Tasarım- A.K.A., C.Ç.D.; Veri Toplama- A.K.A.; Veri Analizi/Yorumlama- A.K.A.; Yazı Taslağı- A.K.A., C.Ç.D.; İçeriğin Eleştirel İncelemesi- A.K.A., C.Ç.D.; Son Onay ve Sorumluluk- A.K.A., C.Ç.D.

Peer-review: Externally peer-reviewed.

Conflict of Interest: The authors have no conflict of interest to declare.

Grant Support: The authors declared that this study has received no financial support.

Author Contributions: Conception/Design of Study- A.K.A., C.Ç.D.; Data Acquisition- A.K.A.; Data Analysis/Interpretation- A.K.A.; Drafting Manuscript- A.K.A., C.Ç.D.; Critical Revision of Manuscript- A.K.A., C.Ç.D.; Final Approval and Accountability- A.K.A., C.Ç.D.

Kaynaklar/References

- Ajami, S., Ketabi, S., & MahmoodAbadi, H. B. (2013). Reducing Waiting Time in Emergency Department at Ayatollah-Kashani Hospital Using Simulation. *Journal of Health Administration*, 16(51), 84–94.
- Albouy, V., Davezies, L., & Debrand, T. (2010). Health expenditure models: A comparison using panel data. *Economic Modelling*, 27(4), 791–803. <https://doi.org/https://doi.org/10.1016/j.econmod.2010.02.006>
- Arah, O. A., Westert, G. P., Hurst, J., & Klazinga, N. S. (2006). A conceptual framework for the OECD Health Care Quality Indicators Project. *International Journal for Quality in Health Care*, 18(supply 1), 5–13.
- Atalan, A. (2018). Türkiye Sağlık Ekonomisi için İstatistiksel Çok Amaçlı Optimizasyon Modelinin Uygulanması. *İşletme Ekonomi ve Yönetim Araştırmaları Dergisi*, 1(1), 34–51. <http://dergipark.gov.tr/download/article-file/414076>
- Atalan, A. (2019). THE IMPACTS OF HEALTHCARE RESOURCES ON SERVICES OF EMERGENCY DEPARTMENT: DISCRETE EVENT SIMULATION WITH BOX-BEHNKEN DESIGN. *PONTE International Scientific Researchs Journal*, 75(6), 12–23. <https://doi.org/10.21506/j.ponte.2019.6.10>
- Atalan, A. (2020). Forecasting for Healthcare Expenditure of Turkey Covering the Years of 2018–2050. *Gümüşhane Üniversitesi Sağlık Bilimleri Dergisi*, 9(1), 8–16. <https://doi.org/10.37989/gumussagbil.538111>
- Atalan, A., Cinar, Z., & Cinar, M. (2020). A TRENDLINE ANALYSIS FOR HEALTHCARE EXPENDITURE PER CAPITA OF OECD MEMBERS. *Sigma Journal Of Engineering And Natural Sciences*, 10(3), 23–35.
- Atalan, A., & Donmez, C. (2019). Employment of Emergency Advanced Nurses of Turkey: A Discrete-Event Simulation Application. *Processes*, 7(1), 48. <https://doi.org/10.3390/pr7010048>
- Baltagi, B. H., & Moscone, F. (2010). Health care expenditure and income in the OECD reconsidered: Evidence from panel data. *Economic Modelling*, 27(4), 804–811. <https://doi.org/https://doi.org/10.1016/j.econmod.2009.12.001>
- Bichay, N. (2020). Health insurance as a state institution: The effect of single-payer insurance on expenditures in OECD countries. *Social Science & Medicine*, 113454. <https://doi.org/10.1016/j.socscimed.2020.113454>
- C., W. M., & B., S. W. (1977). Foundations of Cost-Effectiveness Analysis for Health and Medical Practices. *New England Journal of Medicine*, 296(13), 716–721. <https://doi.org/10.1056/NEJM197703312961304>
- Carides, G. W., Heyse, J. F., & Iglewicz, B. (2000). A regression-based method for estimating mean treatment cost in the presence of right-censoring. *Biostatistics*, 1(3), 299–313. <https://doi.org/10.1093/biostatistics/1.3.299>
- Deb, K., & Kalyanmoy, D. (2001). *Multi-Objective Optimization Using Evolutionary Algorithms*. John Wiley & Sons, Inc.
- Dukhanin, V., Searle, A., Zwerling, A., Dowdy, D. W., Taylor, H. A., & Merritt, M. W. (2018). Integrating social justice concerns into economic evaluation for healthcare and public health: A systematic review. *Social Science and Medicine*, 198, 27–35. <https://doi.org/10.1016/j.socscimed.2017.12.012>
- Gerdtham, U. G., Sogaard, J., Andersson, F., & Jönsson, B. (1992). An econometric analysis of health care expenditure: A cross-section study of the OECD countries. *Journal of Health Economics*, 11(1), 63–84. [https://doi.org/https://doi.org/10.1016/0167-6296\(92\)90025-V](https://doi.org/https://doi.org/10.1016/0167-6296(92)90025-V)

- Gregori, D., Petrinco, M., Bo, S., Desideri, A., Merletti, F., & Pagano, E. (2011). Regression models for analyzing costs and their determinants in health care: an introductory review. *International Journal for Quality in Health Care*, 23(3), 331–341. <https://doi.org/10.1093/intqhc/mzr010>
- Iii, J. H. E., Hwang, Y., & Nagarajan, N. J. (2001). Management control and hospital cost reduction: additional evidence. *Journal of Accounting and Public Polic*, 20(1), 73–88.
- Jones, A. M., Lomas, J., & Rice, N. (2015). Healthcare Cost Regressions: Going Beyond the Mean to Estimate the Full Distribution. *Health Economics*, 24(9), 1192–1212. <https://doi.org/10.1002/hec.3178>
- Liu, Y., Lei, H., Zhang, D., & Wu, Z. (2018). Robust optimization for relief logistics planning under uncertainties in demand and transportation time. *Applied Mathematical Modelling*, 55(Supplement C), 262–280. <https://doi.org/https://doi.org/10.1016/j.apm.2017.10.041>
- Malehi, A. S., Pourmohammadi, F., & Angali, K. A. (2015). Statistical models for the analysis of skewed healthcare cost data: a simulation study. *Health Economics Review*, 5, 11. <https://doi.org/10.1186/s13561-015-0045-7>
- Mason, K., Duggan, J., & Howley, E. (2017). Multi-objective dynamic economic emission dispatch using particle swarm optimisation variants. *Neurocomputing*, 270(Supplement C), 188–197. <https://doi.org/https://doi.org/10.1016/j.neucom.2017.03.086>
- McCoskey, S. K., & Selden, T. M. (1998). Health care expenditures and GDP: panel data unit root test results. *Journal of Health Economics*, 17(3), 369–376. [https://doi.org/https://doi.org/10.1016/S0167-6296\(97\)00040-4](https://doi.org/https://doi.org/10.1016/S0167-6296(97)00040-4)
- Mihaylova, B., Briggs, A., O'Hagan, A., & Thompson, S. G. (2011). Review of Statistical Methods for Analysing Healthcare Resources and Costs. *Health Economics*, 20(8), 897–916. <https://doi.org/10.1002/hec.1653>
- Muyl, F., Dumas, L., & Herbert, V. (2004). Hybrid method for aerodynamic shape optimization in automotive industry. *Computers and Fluids*, 33(5), 849–858. <https://doi.org/https://doi.org/10.1016/j.compfluid.2003.06.007>
- Ngatchou, P., Zarei, A., & El-Sharkawi, A. (2005). Pareto Multi Objective Optimization. *Proceedings of the 13th International Conference on, Intelligent Systems Application to Power Systems*, 84–91. <https://doi.org/10.1109/ISAP.2005.1599245>
- OECD. (2016a). *OECD Data Health Expenditure (Percent of GDP)*. <https://doi.org/http://stats.oecd.org/Index.aspx?DataSetCode=SHA>
- OECD. (2016b). *OECD Data Health Expenditure Per Capita*. <https://doi.org/https://data.oecd.org/healthres/health-spending.htm>
- OECD. (2017a). *Life expectancy at birth 2015*. <https://doi.org/https://data.oecd.org/healthstat/life-expectancy-at-birth.htm>
- OECD. (2017b). *OECD Data Health Care Resources 2015*. https://doi.org/http://stats.oecd.org/Index.aspx?DataSetCode=HEALTH_REAC
- OECD. (2017c). *Pharmaceutical spending Total, % of health spending, 2015*. <https://doi.org/https://data.oecd.org/healthres/pharmaceutical-spending.htm>
- OECD. (2017d). *The Organisation for Economic Co-operation and Development: Members and partners*. <http://www.oecd.org/about/membersandpartners/#d.en.194378>
- OECD. (2017e). *World Health Organization's Global Health Workforce Statistics, OECD, supplemented by country data. Physicians (per 1,000 people)*. <https://data.worldbank.org/indicator/SH.MED.PHYS.ZS>
- Palasca, S., & Jaba, E. (2015). Economic Crisis' Repercussions on European Healthcare Systems. *Procedia Economics and Finance*, 23(Supplement C), 525–533. [https://doi.org/https://doi.org/10.1016/S2212-5671\(15\)00568-7](https://doi.org/https://doi.org/10.1016/S2212-5671(15)00568-7)
- Steiner, M. T. A., Datta, D., Neto, P. J. S., Scarpin, C. T., & Figueira, J. R. (2015). Multi-objective optimization in partitioning the healthcare system of Parana State in Brazil. *Omega*, 52(Supplement C), 53–64. <https://doi.org/https://doi.org/10.1016/j.omega.2014.10.005>
- WBG. (2017). *The World Bank, GDP Per Capita, (current international \$) 2016*. <https://doi.org/https://data.worldbank.org/indicator/NY.GDP.PCAP.PP.CD>

Yapay Zekâ ve Tıbbi Görüntüleme Teknolojilerine Genel Bakış

An Overview of Artificial Intelligence and Medical Imaging Technologies

Furkan Atlan¹ , İhsan Pençe² 



ÖZ

Günümüzde yapay zekânın kullanıldığı alanlar her geçen gün artmakta olup, bu alanlardan biri de sağlık sektörüdür. Özellikle görüntü işlemede oldukça başarılı sonuçlar vermesi sebebi ile yapay zekânın bir alt dalı olan derin öğrenme, tıbbi görüntülerin işlenmesinde ve yorumlanmasında sıkça tercih edilmektedir. Her ne kadar tıbbi görüntüleme teknolojilerinin gelişmesi ile hastalık tanısı ve teşhisi gibi işlemlerdeki doğruluk oranı artsa da bu görüntülerin uzmanlar tarafından doğru bir şekilde yorumlanması zaman açısından maliyetli ve tedavi süreci açısından da olumsuz bir durum sergilemektedir. Bu sebeple, yapay zekâ kullanılarak otomatik tanı sistemleri oluşturulmakta ve bu sistemler gelişen teknoloji ve algoritmalar sayesinde her geçen gün ilerleme kat etmektedir. Çalışmanın amacı, tıbbi görüntüleme yapay zekâ kullanımı konusunda tüm bileşenlerin ele alınarak bilgi verilmesi ve bu alanda çalışma yapacak araştırmacılara bir temel teşkil edecek bir alt yapı oluşturmaktır. Bunun sağlanması için yapay zekâ ve tıbbi görüntüleme konusu öncelikle ayrı bir şekilde ele alınmış, tıbbi görüntüleme teknolojileri kapsamlı bir şekilde anlatılmış ve tıbbi görüntüleme yapay zekâ kullanımının mevcut durumu, geleceği, sorunları ve çözümleri açık bir şekilde belirtilmiştir. Son olarak yapay zekâ teknikleri ile tıbbi görüntülerin işlenmesine dair çalışmalar verilerek çalışmanın teorik anlam bütünlüğü sağlanmıştır.

Anahtar kelimeler: Tıbbi görüntüleme teknolojileri, Yapay zekâ, Bilgisayar destekli tanı sistemleri

ABSTRACT

Nowadays, the use of artificial intelligence is increasing steadily, particularly in the health sector. Deep learning, which is a sub-branch of artificial intelligence, is frequently preferred in the processing and interpretation of medical images, because it provides fruitful outcomes in image processing. Despite the development in medical imaging technologies and the increasing accuracy rate of disease diagnosis, accurate interpretation of these images by experts is time consuming, and unfavorable conditions may arise during treatment. For this reason, automated diagnostic systems are created using artificial intelligence, and these systems are improving gradually, owing to the evolution of several technologies and algorithms. This study aimed to provide information on the use of artificial intelligence in medical imaging with due consideration of all factors and create a base infrastructure for researchers in this field. To achieve this, previously artificial intelligence and medical imaging were discussed separately, placing more emphasis on medical imaging technologies. However, at present, potential problems and solutions in the use of artificial intelligence in medical imaging are clearly stated. In conclusion, by conducting more studies on the processing of medical images using artificial intelligence, the theoretical integrity of this field will become possible.

Keywords: Medical imaging technologies, Artificial intelligence, Computer-aided diagnosis systems

¹(Doktora Öğrencisi), Burdur Mehmet Akif Ersoy Üniversitesi, Sosyal Bilimler Enstitüsü, Yönetim Bilişim Sistemleri, Burdur, Türkiye

²(Dr. Öğr. Üyesi), Burdur Mehmet Akif Ersoy Üniversitesi Bucak Teknoloji Fakültesi, Yazılım Mühendisliği Bölümü, Burdur, Türkiye

ORCID: F.A. 0000-0003-1602-1941;
İ.P. 0000-0003-0734-3869

Corresponding author:

Furkan ATLAN
Burdur Mehmet Akif Ersoy Üniversitesi, Sosyal Bilimler Enstitüsü, Yönetim Bilişim Sistemleri, Burdur, Türkiye
E-mail address: furkanatlan@gmail.com

Submitted: 20.10.2020
Revision Requested: 21.12.2020
Last Revision Received: 15.01.2021
Accepted: 16.01.2021
Published Online: 01.03.2021

Citation: Atlan, F. ve Pence, I. (2021). Yapay zekâ ve tıbbi görüntüleme teknolojilerine genel bakış. *Acta Infologica*, 5(1), 207-230.
<https://doi.org/10.26650/acin.813736>

1. GİRİŞ

Günümüzde, yapay zekânın kullanım alanları gittikçe genişlemekte ve bu alanlarda yapılan çalışmalar da artmaktadır. Özellikle görüntü işleme alanında yapay zekânın alt disiplinlerinden biri olarak karşılaşılan derin öğrenme ve tıbbi görüntü formatlarının dijitalleşmesi sayesinde yapay zekânın sağlık sektöründe kullanımı oldukça dikkat çekici konular arasında yer almaktadır.

Her yıl milyonlarca insanın ölümüne neden olan kanser (Bray ve ark., 2018) vb. hastalıkların önceden tespiti, sınıflandırılması ve bunların otomatikleştirilmesi ile ilgili çalışmalar (Kiani ve ark., 2020; Ghoneim, Muhammad ve Hossain, 2020; Dandil ve Serin, 2020; Öztad, 2020), tıbbi görüntüleme yapay zekâ kullanımının ihtiyacını ve önemini ortaya koymaktadır. Yapay zekânın bu alandaki kullanımı tüm tarafların yani hastaların, doktorların, klinisyenlerin vd. sağlık çalışanları ile yapay zekâ geliştiricilerinin katılımı ile ilerleyen bir süreçten ibaret olarak görülebilir (Al-shamasneh ve Obaidallah, 2017; Greenspan, Van Ginneken ve Summers, 2016). Tıbbi görüntüleme yapay zekâ kullanımını sadece dijital tıbbi görüntülerin yapay zekâ teknikleri ile ön işlenmesinden ibaret değildir. Burada asıl amaç, elle işlenmesi zaman ve maliyet açısından oldukça pahalıya mâl olan bir süreci hastaların, sağlık çalışanlarının ve araştırma görevini yürüten üçüncü taraf kişi ya da kurumların lehine dönüştürmektir. Geliştirilecek uygulama ve/veya modellerin, radyolog, patoloğ gibi işin uzmanlarına danışılarak tüm süreç boyunca karşılıklı bilgi alışverişinde bulunularak hem sağlık çalışanlarının bilgilerinden faydalanılıp onların tecrübesinin oluşturulacak modele yansıtılarak en başarılı modelin ortaya konulması hem de elde edilecek verilere erişimde kolaylık sağlanması ve bunlar yapılırken kişisel mahremiyetin ihlal edilmemesi, günümüzde tıbbi görüntüleme yapay zekâ kullanımında en sık karşılaşılan sorunlar arasındadır (Litjens ve ark., 2017; Fotin, Yin, Haldankar, Hoffmeister ve Periaswamy, 2016).

Bu çalışmada yapay zekâ ve onun alt disiplinleri ile tıbbi görüntüleme teknolojilerine ait kavramlar açıklanmakta, yapay zekânın tıbbi görüntüleme alanında kullanımı ve bu alanda yapılan çalışmaların başarı durumları, sonuçları ve geliştirilen modeller hakkında bilgi verilmektedir. Bu çalışmanın amacı, yapay zekânın tıbbi görüntüleme kullanılması ile ilgili yapılacak çalışmalara temel teşkil edecek bilgilerin verilmesi, literatürde açık bir şekilde ele alınmayan tıbbi görüntüleme yapay zekâ kullanımının mevcut durumu, farklı enerji türleri kullanan tıbbi görüntüleme teknolojileri alanında yapılmış başarılı çalışmalara ait bilgi ve sonuçların sunulmasıdır. Çalışmada ele alınan konular kapsamlı ve açıklayıcı bir şekilde birlikte ele alan bir çalışmanın bulunmaması ve özellikle nükleer tıp görüntüleme konusundaki çalışmaların yetersiz olması bu çalışmanın önemini ve özgünlüğü ortaya koymaktadır.

2. GENEL KAVRAMLAR

Bu bölümde, çalışmanın amacının anlaşılması ve misyonunun yerine gelmesini sağlamak amacıyla yapay zekâ ve tıbbi görüntülemeye dair temel ve ilişkili kavramlar açıklanmaktadır.

2.1. Yapay Zekâ

Zekâ, Arapça bir kelimedir ve Türk Dil Kurumu'na (TDK) göre zekâ: "İnsanın düşünme, akıl yürütme, objektif gerçekleri algılama, yargılama ve sonuç çıkarma yeteneklerinin tamamı, anlayış, dirayet, zeyreklik, feraset" olarak tanımlanmaktadır (<https://www.sozluk.gov.tr>).

Yapay zekâ ve ondan türemiş diğer kollara ait kavramların tek bir tanımı yoktur. Ancak yapılan tanımlamaların özeti olarak aşağıdaki gibi bir kavramsal çerçeve çizilebilir.

Yapay Zekâ, farklı eylemleri ve görevleri zeki bir canlı gibi yerine getiren makine ve sistemlerin meydana getirilmesi ile ilgilenen genel bir teknolojidir. Bir terim olarak yapay zekâyı tanımlayan ilk kişi John McCarthy'dir. McCarthy'e (1955) göre yapay zekânın amacı, "zeki gibi davranan makineler geliştirmektir." (Ertel, Black ve Mast, 2017).

Makine Öğrenmesi, bilgisayarların veriden (data) kurallar çıkararak kendi kendilerine öğrenmelerini ve bu sürecin gelişmesini sağlayan bir yaklaşımdır (Kubat, 2018).

Derin Öğrenme, 2012 yılında ImageNet adı verilen çok büyük sayıdaki etiketli bir veri setini içerisinde yer alan görsellerin içeriğini (sınıflandırılmasını) tahmin etmek için ImageNet 2012 yarışmasında (challenge) (Deng, ve ark., 2012; <https://www.kdnuggets.com/>) AlexNet (Krizhevsky, Sutskever ve Hinton, 2012) adı verilen bir derin öğrenme modelinin %15.3 hata oranı elde ederek görselleri doğru bir şekilde sınıflandırması ile tekrar gündeme gelmiş bir kavramdır (<https://www.economist.com>).

Derin Öğrenme, -ilk ortaya çıkışı ve geliştirilmesi itibari ile- özellikle görseller arasındaki örüntülerin tanınmasında, özniteliklerin çıkarımında, sınıflandırılmasında vb. işlemleri için geliştirilmiş bir yaklaşımdır ve yapay zekânın bir alt dalıdır (Gulli, Kapoor ve Pal, 2019). Ancak son yıllarda derin öğrenmenin kullanım alanları, konuşma sinyallerinin yok edilmesinden genlerin kümelenme modellerinin keşfine kadar farklı alanlara yayılmıştır (Wang ve Raj, 2017).

Günümüzde ImageNet veri seti ile yapılan çalışmalarda, geliştirilen yeni modeller sayesinde Top-1 doğruluk (accuracy) oranı (sınıflandırma çıktısının en yüksek olasılığa sahip tek çıktı ile eşleşmesi) %88.5 ve Top-5 doğruluk (accuracy) oranı (sınıflandırma çıktısının en yüksek olasılığa sahip 5 çıktıdan biri ile eşleşmesi) %98.7 oranına ulaşmıştır (<https://paperswithcode.com>).

2.2. Tıbbi Görüntüleme

Tıbbi görüntüleme (medical imaging), vücudun içyapılarının (doku, organ) analiz ve müdahale için farklı enerji türleri ile görüntülenmesini sağlamak amacıyla kullanılan teknikler ve yapılan işlemler bütünüdür (Bankman, 2009).

İnsan vücudunun tıbbi açıdan görüntülenmesi için enerjiye ihtiyaç duyulmaktadır. Farklı görüntüleme tekniklerinde farklı enerji türleri kullanılmaktadır. Örneğin; endoskopi, gastroenteroloji ve ışık mikroskopisi olarak da adlandırılan patolojide enerji türü olarak görünür ışık (visible light) kullanılmaktadır. Radyoloji tabanlı görüntüleme teknikleri olan mamografi, bilgisayarlı tomografide (BT) enerji türü olarak X-ışınları kullanılmaktadır. Manyetik rezonans görüntülemede (MRI) enerji türü olarak radyo frekansları kullanılmaktadır. Nükleer tıp görüntülemede enerji türü olarak gama ışınları kullanılmaktadır. Ultrason görüntülerinde ise enerji türü olarak yüksek frekanslı ses dalgaları kullanılmaktadır (Bushberg, Seibert, Leidholdt Jr, Boone ve Goldschmidt Jr, 2011).

Nükleer tıp görüntülemenin aksine diğer bütün tıbbi görüntüleme tekniklerinde görüntüleme işleminin gerçekleşmesi, enerjinin vücuda nüfuz etmesi ile mümkündür. Vücuda gönderilen madde, vücut tarafından emilim ya da saçılma yoluyla kabul edilir. Vücudun içerisine giren bu madde ile vücudun içyapısı görüntülenebilir. Nükleer tıp görüntüleme de ise ajan olarak adlandırılan bir radyo aktif madde vücuda enjekte edilir veya ağız yolu ile yutulur. Daha sonra görüntüleme işlemi gerçekleştirilir. Görüntü elde edildikten sonra, tıbbi görüntünün tanınabilir bir fayda sağlaması yani hastalık veya kişinin mevcut durumu hakkında nitelikli bilgi vermesi görüntünün teknik kalitesi ve görüntünün elde edilme süreci ile doğrudan ilişkilidir. Daha uzun süre ve hastanın daha çok radyasyona maruz kalması ile daha kaliteli bir tıbbi görüntü elde edilebilir ancak bu durum hem uzmanlar hem de hasta açısından pek de kabul görmeyen bir durumdur. Bu sebeple tıbbi görüntülemede eğer enerji türü radyasyon ise radyasyona maruz kalma süresi ile tıbbi görüntünün kalitesi arasında bir denge gözetilmektedir (Bushberg, Seibert, Leidholdt Jr, Boone ve Goldschmidt Jr, 2011; Webb ve Flower, 2016).

2.2.1. Tıbbi Görüntüleme Teknikleri

Tıbbi görüntüleme, birçok hastalığın tanı, tedavi yöntemi ve teşhis gibi aşamalarında kullanılmaktadır. Tıbbi görüntüleme, yıllardır kanserin erken teşhis, tanı ve tedavi aşamalarının çok önemli bir parçası haline gelmiştir. Tıbbi görüntüleme sadece kanserin teşhisinde değil tedavi aşamasında hastalığın ne denli ilerlediği ve nasıl bir yayılım izlediği ile de ilgili bilgiler vermektedir.

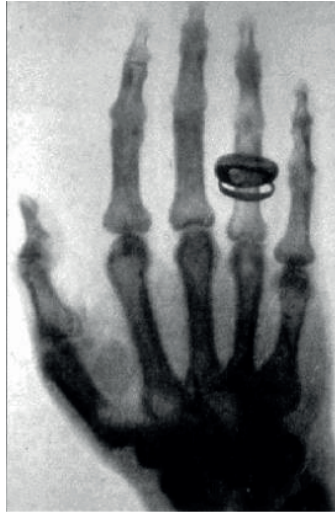
Enerjilerin türleri ve elde etme teknolojileri değiştirilerek farklı tıbbi görüntüler elde edilebilir. Farklı görüntü yapma yöntemleri literatürde “yöntem (modalities)” olarak adlandırılmaktadır (Bushberg, Seibert, Leidholdt Jr, Boone ve Goldschmidt Jr, 2011).

2.1.1.1. Radyografi

Radyografi, Alman fizikçi Wilhelm Conrad Röntgen tarafından 1895 yılında icat edilen ilk tıbbi görüntüleme teknolojisidir.

Röntgen, bir laboratuvar çalışmasında yüksek gerilimli elektrik akımını “Crooks” adı verilen tüpten geçirdiği zaman bir cam kavanoz içerisindeki kristallerde parıltılar gözlemlemiştir. O zamana kadar keşfedilmemiş bu parıltılara sebep olan ışınlara ise bilinmeyen ışın anlamına gelen “X-ışınları” adı verilmiştir. Crooks tüplerinden yüksek gerilimli elektrik akımları geçtiğinde bir ekranda parıltılar meydana getiren X ışınlarının, farklı cisimleri farklı derecelerde geçme özelliği olduğunu ve bu parıltıların (görüntülerin) kurşun levhalar ile depolanabildiğini fark eden Röntgen eliyle tuttuğu kurşun levhalardan ekrana yansıyan gölge de kendi parmaklarının görüntüsü görmüştür. Daha sonra eşinin elini fotoğraf plağı bulunan bir kasetin üzerine yerleştirerek üzerinde eşinin yüzüğünün de bulunduğu ilk röntgen görüntüsünü elde etmiştir (Ehrlich ve Daly, 2008; Arslan, 2010). Röntgen sayesinde insan anatomisi ilk defa radyogram olarak görüntülenmiştir. Şekil 1’de ilk radyogram görüntü olan, Wilhelm Conrad Röntgen’in eşinin elinin radyografik görüntüsü gösterilmektedir.

Radyografi teknolojisi sayesinde radyoloji alanı oluşmuş ve bu alanda uzman olan radyologlar yetişmiştir. Radyografi, hastanın bir tarafında X-ışını kaynağı ve hastanın diğer tarafında bir X-ışını detektörü ile yapılır. X-ışını tüpü tarafından çok kısa süreli (0.5 saniyeden az) bir nabız atımı, X-ışınlarının çok büyük bir bölümü hasta üzerinde etkileşir ve röntgenlerin bir kısmı hastadan geçerek radyogram görüntünün oluştuğu detektöre ulaşır. Hastaya isabet eden X-ışınlarının homojen dağılımı, X-ışınlarının dokulardaki saçılma ve soğurma yoluyla ışıktan çıkarılma derecesi ile değiştirilir.



Şekil 1. Wilhelm Conrad Röntgen’in eşinin eline ait radyogram görüntüsü (<https://www.winally.com>)

Hastanın vücudunun içerisindeki yumuşak doku, kemik ve hava gibi özellikleri de birbirlerinden farklı oldukları için hastadan çıkan X-ışınlarının heterojen dağılımda farklılık görülür. Radyogram görüntü bu röntgen dağılımının bir nevi bir resmi olarak ifade edilebilir. Radyografide kullanılan detektör, bir fotoğraf filmi olabilir. Radyogram görüntüler, akciğer kanseri, kemik kırılmalarında teşhis konulması, kardiyovasküler bozukluklar vb. alanlarda sıklıkla kullanılmakta ve çeşitli tıbbi endikasyonlarda da oldukça faydalıdır (Hou, Zhang, Wei, Guo ve Zhang, 2020).

2.2.1.2. Floroskopi

Floroskopi, zamana bağlı olarak bir dizi X-ışını görüntüsünün, aslında hastanın gerçek zamanlı bir X-ışını filminin sürekli bir şekilde elde edilmesidir. Floroskopi, gerçek zamanlı bir radyografi türüdür. Floroskopi kalp veya yemek borusu (özafagus) gibi anatomik hareketlerin röntgen filmlerinin yapımında kullanılmaktadır (Bushberg, Seibert, Leidholdt Jr, Boone ve Goldschmidt Jr, 2011).

Floroskopi ile X-ışınları tıpkı bir film gibi (zamana bağlı anlık görüntülerin serileştirilmesi) ekrana yansıtılmaktadır. Floroskopi işlemi esnasında vücudun görüntülenmek ve incelenmek istenen özel bir bölümü (organ) içinden sürekli bir

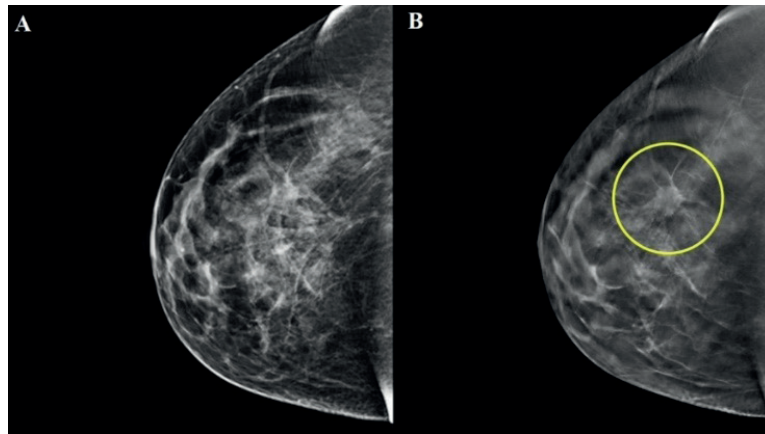
biçimde X-ışını demetleri geçirilmektedir. Bu görüntüler daha sonra ilgili organın durum ve hareketinin uzmanlar tarafından değerlendirilmesi için ekrana yansıtılmaktadır (Larson, Lionhart, Roh ve Colglazier, 2018). Şekil 2’de bağırsak organının floroskopi yöntemi ile görüntülenmesi gösterilmektedir.



Şekil 2. Hastanın bağırsaklarının floroskopi ile görüntülenmesi (<https://www.news-medical.net>)

2.2.1.3. Mamografi

Mamografi, meme kanserinde kullanılan bir görüntüleme tekniği ve memenin bir radyografisidir. Mamografide, memedeki kontrastı vurgulamak için klasik radyografiden çok daha düşük seviyede X-ışını enerjisi kullanılmaktadır. Mamografi görüntüleme teknolojisinde kullanılan X-ışını ve detektör sistemleri sadece meme görüntülerine özel olarak tasarlanmış spesifik bir tekniktir. Mamografinin başlıca 2 kullanım amacı vardır. Bunlar, meme kanseri ile ilgili herhangi bir belirti göstermeyen kadınların kanser taramasında (tarama mamografisi) ve yumru varlığı gibi meme kanseri belirtisi gösteren kadınlarda teşhise (tanı tomografisi) yardımcı olması amacıyla kullanılmasıdır (Pisano ve Yaffe, 2005). Şekil 3’te meme kanseri tanısı için bir memenin mamografi yöntemi ile 2 boyutlu ve 3 boyutlu görüntülenmesi gösterilmektedir.



Şekil 3. Kanserli bir memeye ait Mamografik görüntü A) 2 Boyutlu, B) 3 Boyutlu (<https://www.wakerad.com/>)

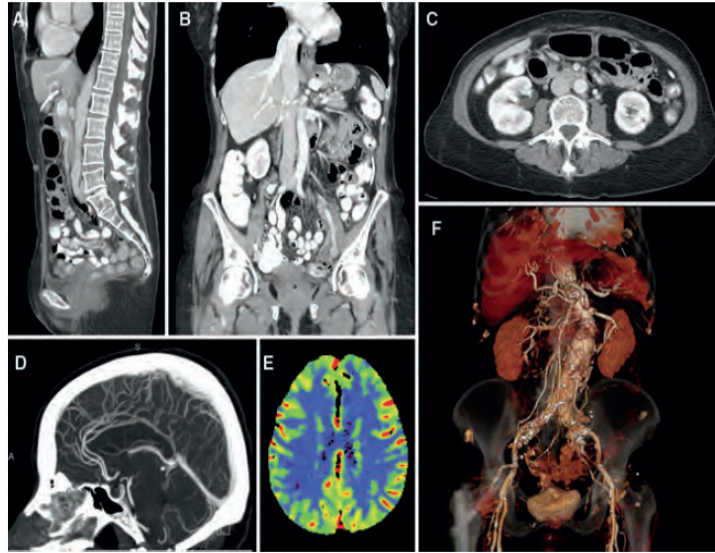
Şekil 3’te gösterilen mamografi görüntüleri aynı hastaya ait aynı görüntülerdir. Sol tarafta gösterilen (A) görüntü, memeye ait 2 boyutlu mamografi görüntüsü ve sağ tarafta gösterilen (B) görüntüsü ise memeye ait 3 boyutlu mamografi görüntüsüdür. Şekil 3’te (B)’de yer alan 3 boyutlu mamografi görüntüsünde işaretlenen kısım, memedeki kanserli kısmı göstermektedir.

2.2.1.4. Bilgisayarlı tomografi

BT, 1970'lerin başında kullanımını mümkün hâle gelen ve bir bilgisayar tarafından tıbbi görüntülemenin sağlandığı ilk yöntemdir. BT görüntüleri, X-ışınlarını vücuttan çok sayıda geçirmek suretiyle X-ışını tüpünü vücut etrafında döndürerek üretilmektedir. Röntgen kaynağının karşısında bir detektör dizisi iletim projeksiyon verilerini toplar. Bu biçimde toplanan çok sayıda veri noktası bir bilgisayar yardımı ile hastanın tomografik görüntülerine sentezlenir. Tomografi terimi iki kelimenin birleşiminden oluşmaktadır. Tomo, dilim ve graphy ise resim olarak çevrilmektedir. Yani tomografi, bir dilimin bir resmini gösterir. BT, hastaya ait münferit doku levhalarının görüntülerini veren bir iletim tekniği olarak düşünülebilir. Günümüz teknolojisinin getirdiği yenilikler düşünüldüğünde BT'nin radyografiye göre üstünlüğü, ilgili anatominin 3 boyutlu dilimlerini gösterebilme, anatomik yapıların birbirlerinin üstüne binme sorununu ortadan kaldırma ve bu sayede doktorlara anatominin ayrıntılı ve kusursuz bir görünümünü sunma özelliğidir (Bushberg, Seibert, Leidholdt Jr, Boone ve Goldschmidt Jr, 2011; Szabo, 2004; Sluimer, Schilham, Prokop ve Van Ginneken, 2006).

BT teknolojisi, keşif cerrahisi ihtiyacını önemli ölçüde azaltarak tıp dünyasında önemli bir çığır açmıştır. Modern bir BT tarayıcısı 5 saniye gibi çok kısa bir zaman diliminde bir hastaya ait 50 cm uzunluğunda (800 görüntüye eşittir) ve 0.50 ila 0.62 mm kalınlığında tomografik görüntüleri elde edebilecek kapasiteye sahiptir (Szabo, 2004).

Şekil 4'te BT'nin farklı organlarda kullanımına ait bir örnek içeren, farklı organlara ait 6 farklı BT görüntüsü gösterilmektedir. Her bir görüntü bir harf ile temsil edilmektedir.



Şekil 4. Farklı organlara ait BT görüntüleri, Karın-pelvis BT Taraması Sonucunda Sagittal (A), Koronal (B) ve Eksenel (C) görüntüleri ile İyotlu Madde Enjeksiyonu Sonucunda Baş (D) Bölgesi, İnme (E) Değerlendirmesi Yapılan bir Hastaya ait Renk Kodlu Kan Hacim Haritası ve BT Verilerinden Görüntü İşleme ile Anatominin Sahte Renkli bir 3 Boyutlu Temsili (F) (Bushberg, Seibert, Leidholdt Jr, Boone ve Goldschmidt Jr, 2011)

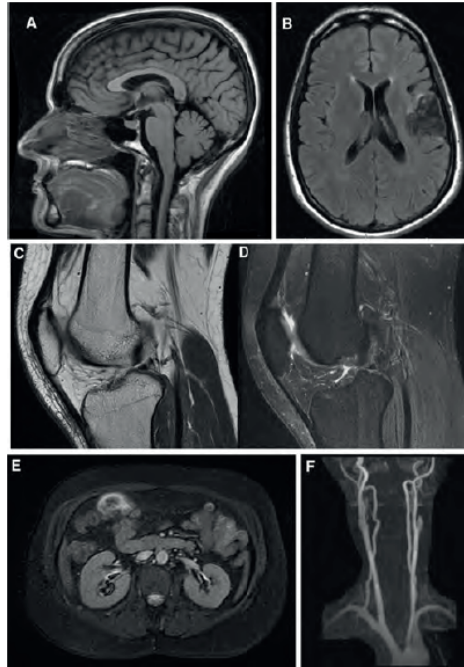
2.2.1.5. Manyetik rezonans görüntüleme

MRG tekniğinde kullanılan tarayıcı cihazların yararlandıkları manyetik alan, dünyanın manyetik alanından yaklaşık 10.000 ila 60.000 kat daha güçlüdür. MRG tekniğinde ortam olarak genellikle protonun nükleer manyetik rezonansları tercih edilmektedir. Yani biyolojik dokularda oldukça fazla sayıda bulunan hidrojen atomunun çekirdeğini kullanırlar. Proton, manyetik bir momente sahip olup 1.5 T (1.5 T = 10^5 Gauss) manyetik alana yerleştirildiği zaman proton eksenini etrafında ilerler ve radyo dalgası enerjisini saniyede yaklaşık 64 milyon devir rezonans frekansında (megahertz-MHz) emer (Huettel, Song ve McCarthy, 2004; Arslan, 2010).

MRG'de görüntüleme prosedürünün başlatılması için hasta manyetik bir alan içerisine yerleştirilir. Daha sonra hastanın etrafına dizilmiş olan radyo antenleri aracılığı ile radyo dalgaları üretilir. Vücutta bulunan protonlar radyo dalgalarını absorbe eder. Sonrasında ise hasta, bu radyo enerjisini dokunun manyetik özelliklerine bağlı olarak bir süre sonra tekrar salgılar. Hastadaki protonların yaydığı radyo dalgaları ise hastanın etrafına yerleştirilmiş olan antenler tarafından tespit edilir. Hastanın bulunduğu manyetik alanın gücü çok hafif bir şekilde değiştirilerek proton rezonans fonksiyonu değiştirilir. Bu, frekans ve manyetik alan kuvvetinin doğru orantılı olmasından kaynaklanmaktadır. MRI'da hastadan MRI cihazına gönderilen her bir sinyalin konumunun tespit edilebilmesi, geri dönen radyo dalgalarına ait frekans ve faz bilgileri ile mümkün olmaktadır. MRG'nin çok sık kullanılan çalışma yöntemlerinden biri "spin eko (spin echo)" görüntüleme olarak isimlendirilmektedir [(Bushberg, Seibert, Leidholdt Jr, Boone ve Goldschmidt Jr, 2011; Huettel, Song ve McCarthy, 2004; Buxton, 2009).

MRG, görüntüdeki her noktanın o noktaya karşılık gelen dokunun mikro manyetik özelliklerine bağlı bir dizi tomografik görüntü üretir. Beyinde yer alan yağ, gri ve beyaz madde, beyin omurilik sıvısı ve kanser gibi farklı doku türlerinin hepsi birbirlerinden farklı yerel manyetik özelliklere sahip olduklarından dolayı MRG ile üretilen görüntüler anatomik varyasyonlara karşı çok büyük bir hassasiyet gösterir. Bundan dolayı da görüntülerin kontrastı oldukça yüksektir. MRG, nörolojik görüntüleme (baş-omurga gibi), yaralanma ve sakatlanma gibi durumlarda diz görüntüleme gibi kas-iskelet yapılarında olağanüstü bir fayda göstermiştir (Huettel, Song ve McCarthy, 2004; Buxton, 2009).

MRG'de farklı nabız dizileri kullanılması sayesinde, MRG cihazı vücuttaki farklı doku özelliklerini ortaya çıkarabilmekte ve doku türlerini de ayırt edebilmektedir. Bir futbol müsabakasında dizinden sakatlanan bir futbolcu düşünüldüğünde, dize ait MRG sayesinde futbolcunun çapraz bağlarının yırtılıp yırtılmadığı ya da ne kadar hasar aldığı tespit edilebilmektedir. Eğer hasar ciddi ise futbolcunun sahalardan ne kadar uzak kalması gerektiği bilinebilir ya da normal bir hastanın beyin MRG'si sayesinde beyindeki gri madde ile beyaz madde arasındaki farkın ölçülmesi mümkündür. Ayrıca; tümörler, kemik hasarları, damarlarda meydana gelen anormal durumlar için özel hassas görüntüler elde edilmesini sağlayan farklı nabız dizilerinin oluşturulması da MRG sayesinde gerçekleştirilebilmektedir (Huettel, Song ve McCarthy, 2004). Şekil 5'te MRI ile insan vücudunun görüntülenmiş farklı organları ve bunlara dair açıklamalar verilmektedir.



Şekil 5. MRI yöntemi ile görüntülenmiş organlar, (A) Beynin T1 ağırlıklı kontrast; (B) beyin enfarktüsü bölgesinin FLAIR görüntüsü; (C) dizlerin T1 ağırlıklı kontrast; (D) dizlerin T1 ağırlıklı kontrast ile yağ doyunluğu; (E) maksimum yoğunluk projeksiyonu; (F) Gadolinyum kontrastlı karın görüntüsü (Bushberg, Seibert, Leidholdt Jr, Boone ve Goldschmidt Jr, 2011)

Şekil 5’te verilen farklı organlara ait MRG görüntüleri bir harf ile temsil edilmekte ve açıklanmaktadır. Buna göre; (A) Beynin Sagittal T1 ağırlıklı kontrast görüntüsünü, (B) beyin enfarktüsü bölgesini gösteren aksel sıvı ile zayıflatılmış inversiyon geri kazanımı (FLAIR) görüntüsünü, dizlerin sagittal görüntüsü olarak; (C) T1 ağırlıklı kontrast ve (D) T1 ağırlıklı kontrast ile “yağ doyunluğu”, (E) uçuş zamanı MR anjiyo gramının aksel tomografik görüntülerinden üretilen maksimum yoğunluk projeksiyonunu, (F) ise Gadolinyum kontrastlı karın görüntüsüne karşılık gelmektedir.

MRG, vücut içerisindeki doku ya da organları radyo frekansı kullanarak görüntülemektedir. MRG tekniği radyasyon kullanmamakta ve yaymamaktadır.

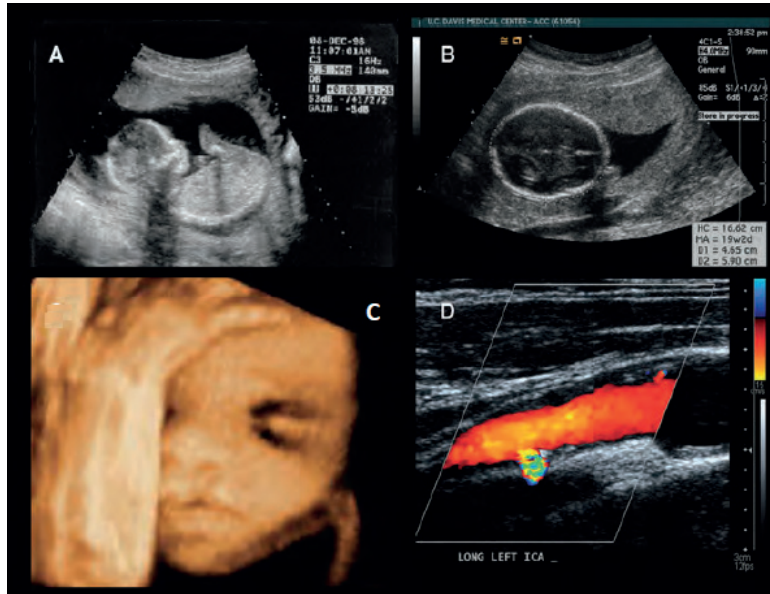
2.2.1.6. Ultrason

Bir cisim yukarıdan aşağıya bir nesnenin üzerine düştü zaman etki basınç dalgalarının (ses) havada belirli bir mesafeden duyulacak şekilde yayılmasını sağlar. Yüksek frekanslı (ultra) ses formundaki mekanik enerji, hastaya ait anatomik görüntüleri oluşturmak için kullanılabilir. Kısa süreli bir ses darbesi, görüntülen dokular ile doğrudan fiziksel temas halinde olan bir ultrason dönüştürücü aracılığı ile üretilmektedir. Ses dalgaları doku içerisine girerek vücutta yer alan iç yapılar sayesinde yansıtılır ve bu sayede yankı oluşturulur. Vücut içerisindeki iç yapılar ile yansıtılan ses dalgaları bir dönüştürücüye ulaşır. Dönüştürücü, yansıtılıp geri dönen sesi kaydeden bir mekanizmadır. Ultrason cihazının bu şekilde çalışan modu “darbe yankı görüntüleme (pulse echo imaging)” olarak adlandırılmaktadır (Leighton, 2007).

Ses dalgası, dikdörtgen şeklinde taranan bir alan oluşturmak için bir doğrusal dizi çok elemanlı dönüştürücü kullanılarak ya da bir sektör kullanılarak taranan alan üretmek için aşamalı bir dizi çok elemanlı dönüştürücü ile artan açılar yoluyla hastanın hizası üzerinde satır satır kaydırılır. Ultrason hatlarından alınan her bir yankı kaydedilir ve görüntülenmek istenen dokuların tomografik bir parçayı simgeleyen gri seviyede kodlanmış olan ses sinyalleri aracılığı ile bir parlaklık modu görüntüsünün elde edilmesi amacıyla kullanılmaktadır (Szabo, 2004; Shung, 2015).

Özellikle karın bölgesindeki organların yüzeyleri ve iç yapıları ultrason yöntemi ile çok güçlü bir şekilde yansıtılır ve görüntülenir. Ultrason, özellikle iyonlaştırıcı radyasyondan daha az zararlı olduğu için bu durum büyüyen bir fetüsün görüntülenmesinde ve gebe kadın (obstetrik) hastaların görüntülenmesinde Şekil 6’daki gibi kullanılmaktadır.

Şekil 6’da verilen dört farklı (A, B, C, D) görüntü arasında ilk üç görüntü (A, B, C) anne karnındaki bir bebeğe ait görüntülerdir. Zira bu görüntüler obstetrik hastalara ait görüntülerdir. Şekil 6’da verilen farklı görüntüler arasındaki son görüntü (D) ise ultrason görüntülemenin farklı bir versiyonu olan Doppler Ultrason görüntüsüdür.

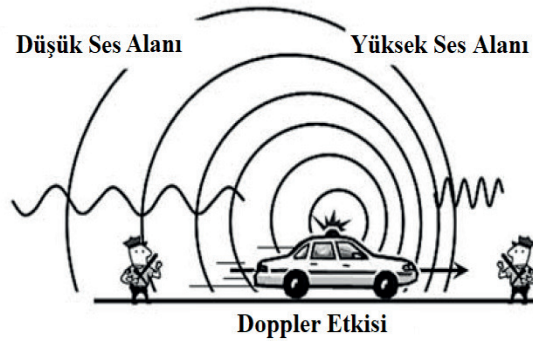


Şekil 6. Anne karnındaki bebeğin görüntüsü, (A) 2.5 aylık bir bebek; (B) bebek yaşı tahmini için bir fetüsün kesitsel beyin ultrason görüntüsü; (C) Ses görüntü verilerinin 3 boyutlu görüntüsü; (D) Doppler renk akışı görüntüleme yöntemi ile vasküler (damar) değerlendirme (Bushberg, Seibert, Leidholdt Jr, Boone ve Goldschmidt Jr, 2011)

2.2.1.7. Doppler ultrason

Doppler ultrason, Doppler etkisinin kullanıldığı özel ve geliştirilmiş bir ultrason çeşididir. Doppler etkisi açıklanırken genellikle çok bilindik bir fenomen kullanılmaktadır. Örneğin; çok hızlı bir şekilde yol alan ve istasyonda bekleyen birine yaklaşan bir tren düşünülün. Tren istasyona yaklaştıkça çıkardığı sesin yükselmesi, istasyondan geçtikten sonra da bu sesin azalmasının sebebi yani sesin frekansında meydana gelen değişiklik, Doppler etkisinin bir sonucudur (Maulik ve Zalud, 2005).

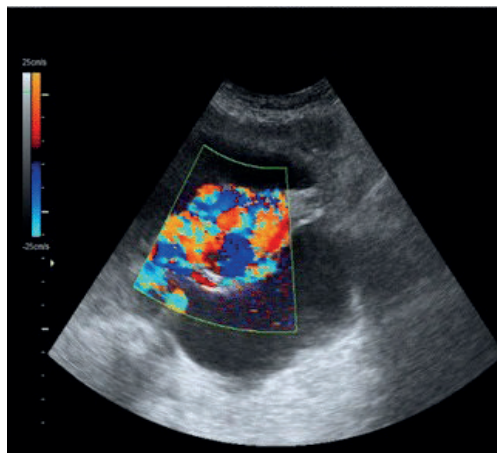
Doppler etkisi, ses dalgasının kaynağı ile bunu gözlemleyen veya maruz kalan kişi arasında bir hareket gerçekleştiğinde, enerjinin bir dalga olarak iletiminde frekansta meydana gelen değişiklikleri kapsayan bir süreçtir (Maulik ve Zalud, 2005). Şekil 7’de Doppler etkisinin gerçek hayatta görülebileceği bir örnek görülmektedir.



Şekil 7. Polis sireni ile Doppler etkisinin açıklanması (Maulik ve Zalud, 2005)

Şekil 7’de de görüldüğü gibi bir kaynak (polis sireni) ve iki gözlemci (polisler) vardır. Polis sireni Şekil 7’de sol tarafta bulunan polisten uzaklaşıp sağ taraftaki polise doğru yaklaşmaktadır. Bu sırada Doppler kayma sesi (siren sesi) ile sol taraftaki polis arasındaki dalga boyu artmıştır; yani frekans düşük seviyededir. Ancak sağ taraftaki polis ile siren arasındaki dalga boyu azalmıştır; frekans yüksek seviyededir. Siren sesi, hangi gözlemciye doğru hareket ederse, frekans yüksek seviyeye ulaşacaktır, arkasında bıraktığı gözlemci de ise frekans düşük seviyede olacaktır.

Normal bir ultrason ses dalgaları ile organların durumunu gösterir ve radyasyon yaymaz. Doppler Ultrasonun normal ultrasondan farkı ise damarlardaki kan akışını kontrol etmek amacıyla yüksek frekansa sahip ses dalgaları kullanmasıdır. Doppler Ultrason sayesinde kan hücreleri tespit edilebilmekte ve renkli bir şekilde görüntülenebilmektedir. Doppler Ultrason ile vücudun farklı organları ve dokuları görüntülenebilmektedir. Şekil 8’de Doppler Ultrason tekniği ile rahim ağzındaki kist bölgesindeki kan akışı görülmektedir.



Şekil 8. Doppler Ultrason ile rahim ağzındaki kistlerde kan akışının görüntülenmesi (Mettler ve Guiberteau, 2012)

Şekil 8’de verilen Doppler Ultrason tekniği ile görüntülenen bölgedeki farklı renkler, o bölgedeki kan akışı seviyeni göstermektedir.

2.2.2. Nükleer tıp görüntüleme

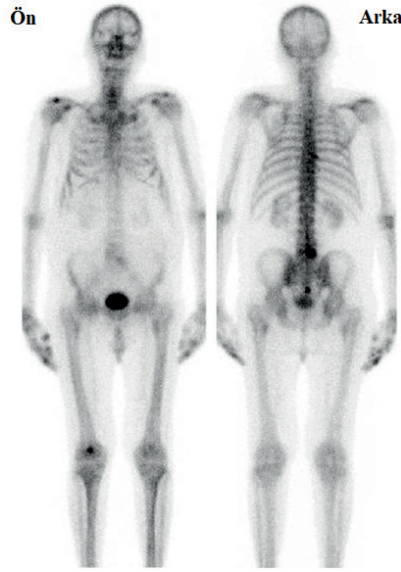
Nükleer tıp, radyoaktif izotop barındıran bir kimyasal ya da bir başka maddenin hastaya oral, enjeksiyon ya da havayı içine çekme (inhalasyon) yoluyla verilen radyoloji dalına verilen isimdir. Nükleer tıp görüntülemede (NTG) öncelikle materyal, hastanın fizyolojik durumuna göre dağıtılır. Burada materyal ile kastedilen, düşük seviyede bir radyoaktif maddedir. Bu, radyoaktif madde “ajan” olarak da isimlendirilmektedir. Ajan, incelenmek istenen dokuya ulaşır ve orada yerleşir. Ajanın yerleştiği dokunun görüntülenmesi için bir radyasyon dedektörü kullanılmaktadır. Daha sonra ajanın radyoaktif bozulma esnasında yayılan X-ışını ya da gama ışınlarından projeksiyon görüntüleri oluşturmak için bir radyasyon detektörü kullanılır. Nükleer tıp, emisyon görüntülerini üretir (Bushberg, Seibert, Leidholdt Jr, Boone ve Goldschmidt Jr, 2011).

NTG, bir fonksiyonel görüntüleme biçimidir. Nükleer tıp görüntüleri sadece hastaya ait anatomik yapı hakkında bilgi vermekle kalmaz bunun yanında hasta ile ilgili fizyolojik durumlarla alakalı bilgi de verir. Örnek verilecek olursa talyum, normal kalp kasında konsantrasyon olma eğiliminde olmasına rağmen infarkte (dolaşım yetmezliği sebebiyle gerçekleşen doku ölümü) olan ya da iskemi (kan akışının azaldığı ya da kesildiği) olan bölgelerde talyum da konsantrasyon olmayacaktır. Söz konusu bölgeler bir nükleer tıp görüntüsünde “soğuk noktalar (cold spots)” biçiminde görüntülenmektedir. Öte yandan nükleer tıp görüntüleri kalbin fonksiyonel durumunun görüntülenmesinde kullanılmaktadır. Tiroid dokusu iyot için büyük bir çekim (afinite) gücüne sahip olduğu için tiroid dokusuna radyoaktif iyot uygulandığında bir nükleer tıp görüntüsü olarak tiroid görüntülenebilmektedir. Eğer tiroid kanseri, hastada metastaz yapmışsa (yayılmışsa) işte o zaman nükleer tıp görüntülerinde yerlerini gösteren “sıcak noktalar (hot spots)” olabilir. Bu özellikleri düşünüldüğünde fonksiyonel görüntüleme nükleer tıbbın kalbi sayılabilmektedir (Mettler ve Guiberteau, 2012; Bushberg, Seibert, Leidholdt Jr, Boone ve Goldschmidt Jr, 2011).

2.2.2.1. Nükleer tıp düzlemsel görüntüleme

NTG yöntemi ile incelenmek istenen hastanın vücuduna bir ajan olarak radyoaktif madde enjekte edilir. Nükleer tıp düzlemsel görüntüleme (NTDG) sayesinde vücuda enjekte edilen bu ajanın ilgili doku ya da organ üzerindeki dağılımının görüntülenmesi düzlemsel (planar) bir X-ışını taraması gibi düşünülebilecek bir 2 boyutlu görüntü ile görülebilmektedir. NTDG genellikle tümör tespiti amacıyla kullanılmaktadır. NTDG’nin özelleştirilmiş bir kullanımı ise kemik ve yayılmacı (metastatik) tümörlerin tespitinde vücudun tamamının taraması ile sağlanmaktadır (Smith ve Webb, 2010).

NTDG yöntemi ile oluşturulan görüntüler birer projeksiyon görüntüleri olarak düşünülebilir. Çünkü görüntüde yer alan her bir nokta, hastanın üzerine yansıtılan bir çizgi boyunca radyo izotop aktivitesini temsil etmektedir. Düzlemsel nükleer görüntüler bir nevi 3 boyutlu radyo izotop dağılımına ait 2 boyutlu haritalar olarak düşünülebilir. Bu durum birçok bozukluğun değerlendirilmesinde yardımcı olmaktadır (Buck ve ark., 2008). Şekil 9’da NTDG ile kemik taraması görüntüleri görülmektedir.



Şekil 9. NTDG ile kemik taraması (Buck ve ark., 2008)

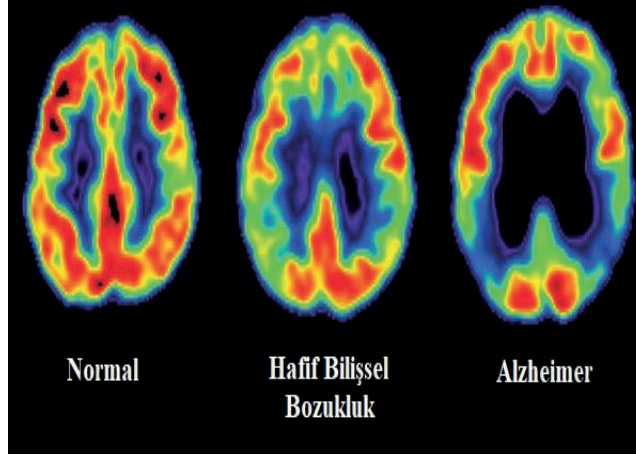
Şekil 9’da kemik taraması için hastanın tüm vücudu NTDG yöntemi ile taranmaktadır. Elde edilen NTDG görüntüsü aynı hastaya ait olup önünden ve arkasından çekilmiştir.

2.2.2.2. Tek foton emisyonlu bilgisayarlı tomografi

Tek foton emisyonlu bilgisayarlı tomografi (SPECT), fizyolojik aktivitenin kesit ve 3 boyutlu görüntülerini elde etmeye yönelik düzlemsel sintigrafinin mantıksal yolu olarak tanımlanabilir. Böylelikle SPECT, BT ve MRI ile üretilen yapısal görüntüleri tamamlama imkânına sahip olur. “Anger” olarak adlandırılan bir kamera, herhangi bir anda radyoaktif bozulmaya ait projeksiyon görüntüsünü kameranın açısı boyunca çeker. Kamera döndüğünde projeksiyonlar BT’de kullanılan aynı rekonstrüksiyonun elde edilmesini sağlayan farklı açılardan toplanabilir. Tüm bunlara rağmen, BT ile SPECT karşılaştırıldığı zaman SPECT çok daha yüksek bir gürültü seviyesine ve daha düşük bir uzaysal çözünürlüğe sahiptir (Haidekker, 2013).

X-ışınları ile görüntüleme bir organın vücudun iç yapılarının ne durumda olduğu yani nasıl görüldüğü görüntülenirken, SPECT ile görüntüleme ise o organın nasıl çalıştığı görüntülenebilmektedir. Örneğin, beynin hangi bölgesinin aktif hangi bölgesinin daha az aktif olduğu SPECT ile görüntülenebilmektedir. Şekil 10’da beynin SPECT ile görüntülenmesi sayesinde Alzheimer Hastalığı teşhisinin yapılması görülmektedir.

Şekil 10’da SPECT yöntemi kullanılarak 3 farklı beyin görüntüsü gösterilmektedir. Soldan sağa doğru ilk beyin görüntüsü Normal, ikinci beyin görüntüsü Hafif Bilişsel Bozukluk ve üçüncü beyin görüntüsü ise Alzheimer Hastası olan üç farklı hastaya aittir. SPECT yöntemi sayesinde beynin sadece yapısı değil aynı zamanda nasıl çalıştığı da görüntülenebildiği için beyin görüntülenmesinde beyin ile ilgili hastalıklar da teşhis edilebilmektedir.

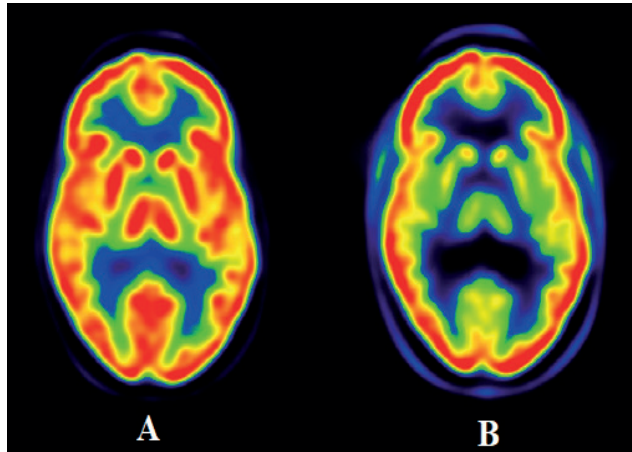


Şekil 10. SPECT ile beynin aktif bölgelerinin görüntülenmesi (<https://www.pinterest.co.uk>)

2.2.2.3. Pozitron emisyon tomografi

Pozitron emisyon tomografi (PET) görüntüleme, nükleer tıbbın neredeyse tüm etkisini ve rolünü değiştirecek şekilde geliştirilmiştir. Çünkü SPECT gibi geleneksel tek foton görüntülemeye ait tüm işlevleri daha iyi yapmanın yanında hızlı bir şekilde büyüyen teşhis uzmanlığı ve halkla ilişkiler özelliğinin etkisini de barındırmaktadır. PET yöntemi şu ana kadar en iyi radyo farmasötik kullanılan güçlü bir metabolik görüntüleme tekniğidir. Ayrıca PET, nükleer olmayan tıp klinisyenleri tarafından önemi takdir edilecek birçok çalışmada gösterdiği mükemmel kalitede görüntüler vermektedir (Bailey, Maisey, Townsend ve Valk, 2005).

PET, damar yolu ile vücuda enjekte edilen metabolik radyoaktif ajanların biriktiği normal veya patolojik dokuları görüntüleme imkânı sağlayan bir cihazdır. Şekerin bir türevi olan ve pozitron ışıması yapabilen flor-18 ile işaretlenmiş “fdg” molekülü damar yolu ile enjekte edilmek suretiyle hastaya verilir (Buck ve ark., 2008). PET görüntülerinin işlenmesinde zayıflama düzeltmesi kullanılmaktadır. Zayıflama düzeltmesi ile beyin merkezinden salınan fotonların beyin çevresinden salınan fotonlardan daha fazla azaltılmasının önüne geçilmektedir (Lameka, Farwell ve Ichise, 2016). Şekil 11’de flor-18 ile işaretlenmiş ve zayıflama düzeltmesi uygulanan ve uygulanmayan bir beynin PET ile görüntülenmesi görülmektedir.



Şekil 11. Bir beynin PET ile görüntülenmesi, A) Zayıflama düzeltmesi uygulanan, B) Zayıflama düzeltmesi uygulanmayan (<https://www.newyorker.com>)

Şekil 11’de görüldüğü gibi zayıflama düzeltmesi uygulanmayan bir beynin PET görüntülenmesinde yoğunluğu yaklaşık %50 oranında azalmakta olup bu durum uzmanların ve geliştirilen sistemlerin daha başarılı değerlendirmesinde yararlı olmaktadır.

3. TIBBİ GÖRÜNTÜLEMEDE YAPAY ZEKÂ KULLANIMI

Son yıllarda özellikle derin öğrenmenin popülerliği ve gelişimi sayesinde yapay zekâ teknolojisinde çok ciddi ilerlemeler görülmektedir. Yapay Zekâ teknolojisindeki gelişmelerden faydalanan konulardan biri de tıbbi görüntülemedir. Tıbbi görüntüleme alanında radyolojide görülen gelişmeler ve yapay zekânın bu alandaki kullanımından elde edilen başarılı sonuçlar, bu alandaki radyolog ve teknisyenlerin bu teknolojiyi ne denli benimsediklerinin iyi bir örneğidir (The, 2018).

2019 yılı itibari ile yapay zekâ destekli birçok sağlık çalışması araştırma ve geliştirme aşamasında olup, tıbbi görüntüleme alanında yapay zekâ kullanımı ile ilgili çalışmaların klinik değeri büyük oranda spekülasyona açık ve hâlâ çok güçlü klinik bulgularla desteklenmemektedir (The, 2018; Morra, Delsanto ve Correale, 2019).

Tıbbi görüntülemede yapay zekâ tabanlı uygulamalar geliştirmenin en zorlu yanlarından ve hatta en büyük dezavantajlarından biri veri kümesine erişim ve veri kalitesinin istenen düzeyde olmasıdır. Son yıllarda tıbbi görüntülemenin tamamen dijitalleştirilmesinin beraberinde getirdiği sorunlardan biridir veriye erişmek. Başarılı sonuç veren bir yapay zekâ modelinin ortaya konulması için eğitim ve test verileri olmazsa olmazlar arasındadır. Bu nedenle bu sorunlar yakın gelecekte acilen çözüm bulunması gereken en önemli konuyu teşkil etmektedir yani veriye erişim konusu (Morra, Delsanto ve Correale, 2019).

Veriye erişim sorununun kısmen çözülmesi için tıbbi görüntüleme alanında yapay zekâ kullanımının geliştirilmesini amaçlayan araştırmacılar tarafından önerilen yollardan biri de, yapay zekâ içerisindeki öğrenme stratejilerinin geliştirilmesi ve veri artırma (data augmentation), transfer öğrenme (transfer learning) ve aktif öğrenme (active learning) gibi konulara yatırım yapılarak veri bağımlılığının en aza indirgenmesidir (Sahiner ve ark., 2019).

Yine veriye erişim ile ilgili sorunların çözümü noktasında önerilen bir başka yol daha mevcuttur. Bu da, kişisel verilere erişimin düzenlenmesine yönelik yeni bir etik anlayışıdır. Kişisel verilere erişimde mahremiyet kurlarının ihlal edilmemesi sadece tıbbi alanda araştırma yapan kişi ya da toplulukların değil aynı zamanda teknoloji devi şirketlerin de karşılaştığı ve hatta yargılandığı bir sorundur (Hutton ve Henderson, 2017). Bu sorun, özellikle Avrupa'daki araştırmacıların, Çin gibi kişisel verilere erişimde daha geniş haklar tanıyan ülkelerin bu alanda yapacakları çalışmalar ile büyük bir avantaj sağlayacaklarını düşünmelerine neden olmaktadır (<https://www.wired.com/> ; <https://www.wsj.com/>). Veriye erişimdeki mahremiyet ihlallerinin önüne geçilmesi ve tıbbi görüntüleme alanında yapılacak çalışmaların da sektöre uğramaması için Avrupa'da biyo bankalar konusunda bir farkındalık gelişmeye başlamıştır (ESR, 2015; Neri ve Regge, 2017).

Son yıllarda özellikle derin öğrenme alanında bilgisayarlı görü konusu sayesinde farklı sektörlerin yapay zekâyâ olan ilgisi bir başka boyuta ulaşmıştır. Tıp bilimi kendi içerisinde birçok disiplini barındırmaktadır. Hasta psikolojisini düşünme, mahremiyeti ihlal etmeme, empati kurma gibi etik ve insani konular elbette ki otomatikleştirilemez. Böylesi durumlar insanidir. Yapay zekânın tıbbi görüntülemede kullanımı konusu sadece doktorlara ve diğer sağlık personellerine ve bu alandaki uzmanlara yardımcı olabilecek uygulamalar geliştirilmesinden ibarettir. Bu durum hem sağlık çalışanlarının hem de hastaların menfaatine olacaktır. Tıbbi görüntülerin yorumlanması ve bu doğrultuda konulacak bir teşhisin en kısa sürede ve büyük bir doğruluk ile konulması, hasta durumunun takibi vb. konularda yapay zekâ sayesinde hem zamandan tasarruf edilecek, hem maddi kaynaklar en az düzeyde kullanılacak hem de yapılacak çalışmalara olan güven ve dolayısı ile yatırım artacaktır (Ranschaert, Morozov ve Algra, 2019).

3.1. Bilgisayar Destekli Tanı Sistemleri

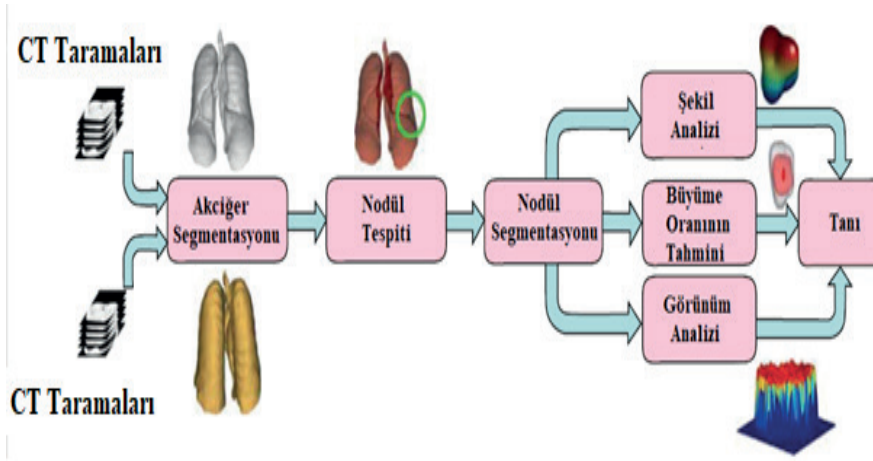
Medikal görüntülerin manuel olarak analiz edilmesi, değerlendirilmesi ve sınıflandırılması gibi işlemler hem maliyet hem de zaman açısından oldukça büyük bir yük teşkil etmektedir. Özellikle böylesine hassas görüntüler değerlendirileceği zaman analiz eden kişi ya da kişiler tarafından önerilen yanlış bir yorum ya da değerlendirme, hasta açısından hayati derecede bir hataya sebebiyet verebilmektedir. Ayrıca, görüntülemeyle ilgili tıbbi bakım maliyetleri de çok hızlı bir şekilde artmaktadır (Cheng, Cai, Chen, Hu ve Lou, 2003). Bu sebeplerden dolayı yeni teşhis yöntemlerine ihtiyaç duyulmaktadır. Günümüzde bilgisayar destekli tanı (CAD), tanısal radyoloji ve tıbbi görüntülemede çok sayıda önemli araştırma konularından biridir (Murino, Puppo, Sona, Cristani ve Sansone, 2015; Al-shamasneh ve Obaidallah, 2017).

CAD yaklaşımı, doktorların muayene süresini ve maliyetini en aza indirirken gereksiz biyopsi prosedürlerinden kaçınmanın haricinde hastalıkları daha yüksek bir verimlilik ile teşhis etmelerinde büyük fayda sağlamaktadır. CAD, günümüzde BT, MRI, X-Işını ya da mamogram görüntülerinin uygulanması ile birincil kanser teşhisi için çok daha uygun bir yöntemdir (Doi, 2007; Al-shamasneh ve Obaidellah, 2017). CAD, medikal giriş görüntüleri ile uzman (radyolog) arasında bir ara katman görevi görmektedir. Burada dikkat edilmesi gereken nokta CAD çıktısının bir sonuç olarak kabul edilemeyeceğidir. Fakat elde edilen çıktı, ilgili alanda yapılan ek testler için referans olarak kullanılabilir (Al-shamasneh ve Obaidellah, 2017).

CAD, kanserin erken ve daha kesin tespitinde doktorlara yardımcı olmaktadır. CAD sistemi yapay zekâ, dijital görüntü işleme, tıbbi bilgi işleme, görüntü analizi ve örüntü tanıma gibi alanlarda birden fazla bilim dalı ile ilişkili olarak geliştirilebilmektedir (Murino, Puppo, Sona, Cristani ve Sansone, 2015; El-Baz ve ark., 2013).

CAD sistemi çoğunlukla beyin, akciğer, meme ve başka kanser türleri için faydalıdır. CAD günümüz teknolojisi sayesinde oldukça ilerlemiş ve yüksek düzeyde doğruluğa ulaşmıştır. CAD sisteminin başlıca parametreleri de özgüllük (specifity), hassasiyet (sensitivity) ve mutlak tespit oranını (absolute detection rate) içermektedir (Al-shamasneh ve Obaidellah, 2017).

Şekil 12’de akciğer kanseri için BT görüntülerinin giriş olarak kullanıldığı bir CAD sistemi yapısı verilmektedir.



Şekil 12. Akciğer kanseri için CAD sistemi diyagramı (El-Baz ve ark., 2013)

Şekil 12’de görüldüğü gibi bir CAD sisteminin girişi, uygun bir yöntem kullanılarak elde edilen tıbbi görüntülerdir. Akciğer nodüllerinin arama alanını azaltmak için bir akciğer bölütlenmesi (segmentasyonu) adımı kullanılır. Nodül tespiti, akciğer nodüllerinin yerlerini belirlemek için kullanılır. Tespit edilen nodüller segmentlere ayrılır. Daha sonra hacim, şekil ve görünüm özellikleri gibi bir öznitelik seti çıkarılır ve tanı için kullanılır (El-Baz ve Suri, 2019). Öznitelik verilerinin tanı için kullanılmasında yapay zekâ tekniklerinden sıklıkla yararlanılmaktadır.

4. LİTERATÜR ÇALIŞMASI

Literatür içerisinde Tıbbi görüntüleme ve yapay zekâ teknolojilerini birlikte ele alan, yapay zekânın tıbbi görüntüleme kullanımını inceleyen bir makale çalışması bulunmamaktadır. Var olan çalışmaların hemen hepsi, bir tıbbi görüntüleme teknolojisinden elde edilen tıbbi bir görüntünün analizini yapan çalışmalardır. Bu çalışmanın en özgün yanı, yapay zekâ ve onun alt disiplinlerini açıklayıp ve aralarındaki farkı ortaya koyması, farklı enerji türlerini kullanan tıbbi görüntüleme teknolojilerini bir arada vermesi ve en önemlisi tıbbi görüntüleme yapay zekâ kullanımının mevcut durumu hakkında açık ve kapsamlı bilgiler vererek tüm bunları kavramsal olarak ortaya koymasındadır.

Tıbbi görüntüleme yapay zekânın kullanımına dair birçok çalışma bulunmakta olup, bu çalışmalar hastalıklara ait tanı konmasında ve uzmanlara yardımcı bilgiler vermeleri açısından önemlidir. Bunun yanında ileride yapılabilecek diğer çalışmalara da ışık tutmaktadırlar. Güncel çalışmalara ilişkin problem tanımları, yöntemler ve elde edilen sonuçlar Tablo 1’de verilmiştir.

Tablo 1

Tıbbi görüntüleme yapay zekâ kullanımı ile ilgili yapılmış çalışmalar

Yazar(lar)	Görüntüleme Tekniği	Araştırma Konusu	Yöntem*	Sonuç
Talo (2019)	Radyografi	Akciğer radyografi görüntüleri ile pnömoni hastalığının teşhisi	Transfer Öğrenme ile ResNET-152 Evrişimsel Sinir Ağı	Doğruluk (Accuracy) = 0.974
Chow ve ark. (2020)	Floroskopi	Cerrahi hassasiyeti artırmak için tek ve çift düzlemli floroskopi görüntülerinin kendi kendini kalibre etmesi	Yumuşatılmış k-En Yakın Komşu Regresyonu	Floroskopi görüntülerine ait 3 boyutlu haritalama hatası %94.2 ve %95.7 oranında azaltılmıştır.
Tiryaki (2020)	Mamografi	Kitleli ve kalsifikasyonlu mamografi görüntülerinin sınıflandırılması	Yapay Sinir Ağları	Doğruluk Oranı = %85.74
Budak (2019)	BT	BT görüntülerinde karaciğer bölgesinin bölütlenmesi	SegNet	Bölütleme ortalama başarı = 67.41
Sarhan (2020)	MRI	Beyin tümörlerinin sınıflandırılması için yeni bir Bilgisayar Destekli Tespit tekniği	Dalgacık tabanlı Evrişimli Sinir Ağı (WCNN)	Doğruluk = %99.3 Duyarlılık = %97.4 Özgüllük = %95.5
Vakanski ve ark. (2020)	Ultrason	Meme kanserinin teşhisinde ultrason görüntülerinin görsel belirginliğinin artırılması	Dikkat Blokları eklenmiş U-Net Modeli (U-NET-SA-C)	Zar Benzerlik Katsayısı (DSC) = %90.5 Doğruluk = %95.7
Shia ve Chen (2020)	Doppler Ultrason	İyi huylu ve kötü huylu tümör sınıflandırılması	Derin Öğrenme (Deep Learning)	Duyarlılık = %77.08 Özgüllük = %91.07
Cai ve ark. (2018)	NTDG	Üst üste binmiş olan NTDG görüntülerinin tek çekim haritalaması	ResNet tabanlı uçtan uca derin Evrişimli Sinir Ağı	Önerilen modelde geleneksel yöntemlere göre daha başarılı haritalama sonucu elde edilmiştir.
Shi ve ark. (2020)	SPECT	Derin Öğrenme yöntemleri ile zayıflama haritalarının tahmin edilerek SPECT görüntülerinin doğruluğunun artırılması	Derin Evrişimsel Sinir Ağı	Model uygulanmadan önce; SPECT görüntüleri arasındaki Normalleştirilmiş Ortalama Mutlak Hata (NMAE) = %3.60 iken, Model uygulandıktan sonra; SPECT görüntüleri arasındaki Normalleştirilmiş Ortalama Mutlak Hata (NMAE) = %0.26 olarak bulunmuştur.
Dong ve ark. (2020)	PET	PET görüntüleme zayıflama düzeltmesi için Derin Öğrenme tabanlı model geliştirilmesi	3 Boyutlu yama tabanlı Döngü Bağımlı Çekişmeli Üretici Ağlar (CycleGAN)	Orijinal PET görüntüleri ile modelin uygulandığı PET görüntüleri kıyaslandığında ölçülen yoğunluk değişiklikleri %3'ten daha düşük bir oranda farklılık gösterdiği görülmüştür.

*: Tablo 1'de verilen yöntemler, 4.1. Yöntemler bölümünde kısa ve genel bir şekilde açıklanmaktadır.

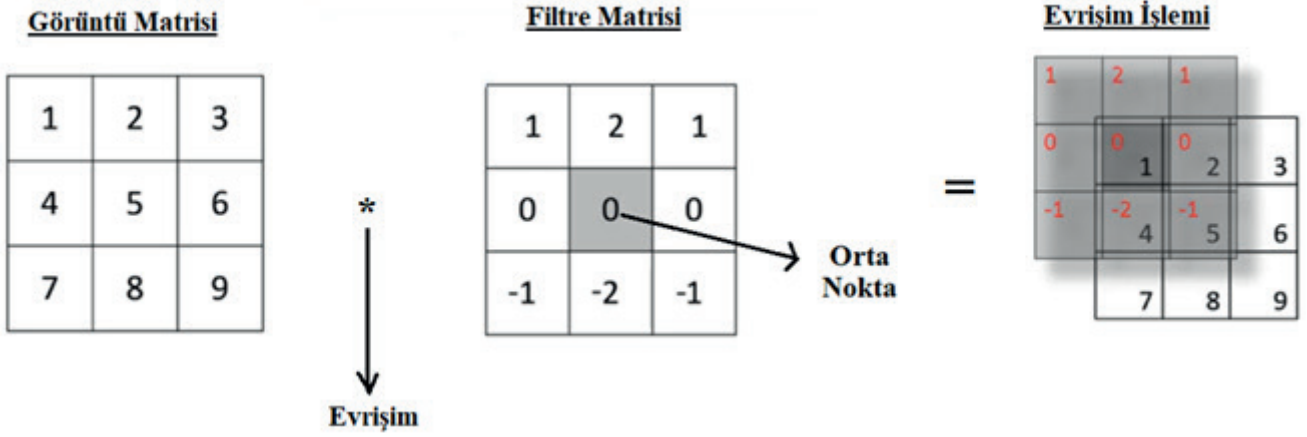
4.1. Yöntemler

Çalışmanın bu bölümünde Tablo 1'de verilen yöntemlerin klasik mimarileri açıklanmaktadır. Bu sebeple, ilk olarak derin öğrenme alanında bilgisayarlı görünün temeli olan Evrişimli Sinir Ağları anlatılmaktadır. Bilgisayarlı görü alanında Evrişimli Sinir Ağları'ndan sonra geliştirilen tüm ağlar ondan esinlendiği için bu ağların en temel katmanları açıklanmaktadır.

4.1.1. Evrişimli Sinir Ağları (Convolutional Neural Network)

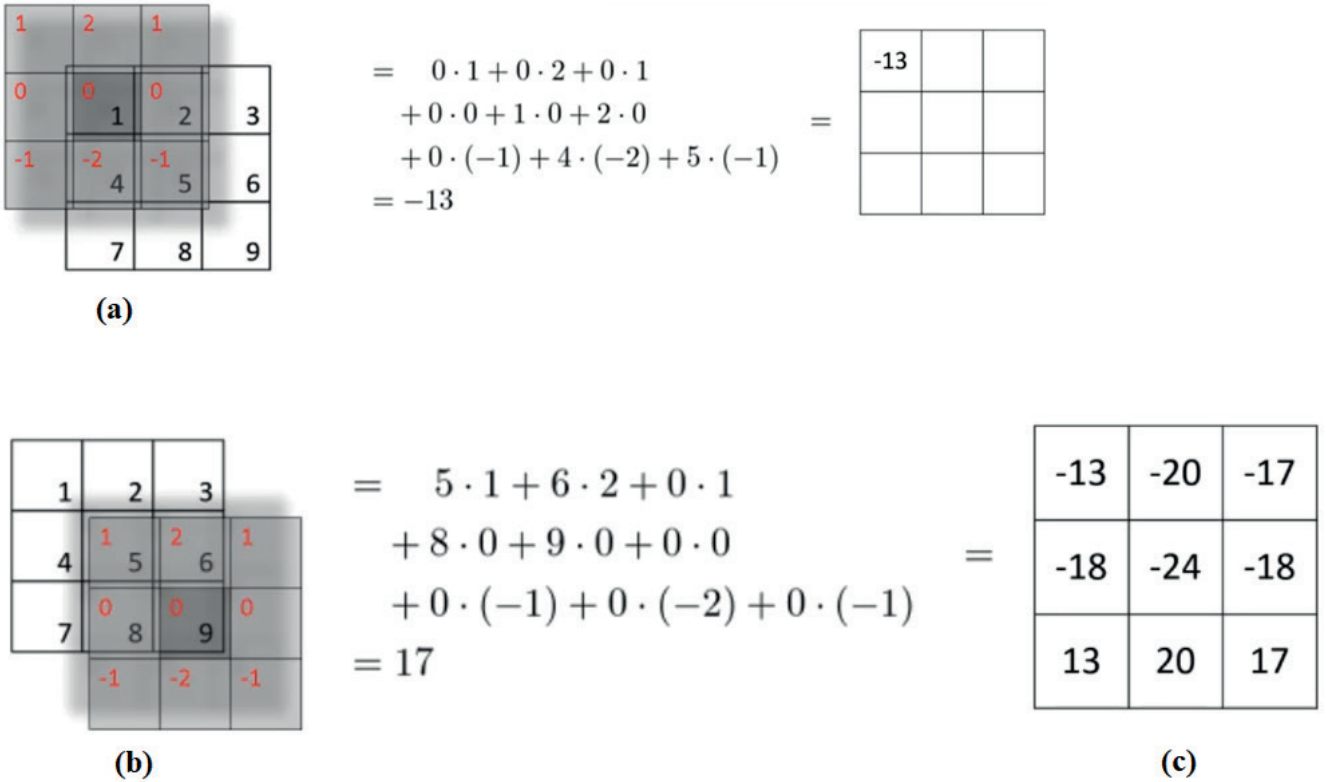
Derin öğrenmenin bir alt dalı olan Evrişimli Sinir Ağları (LeCun, Bottou, Bengio ve Haffner, 1998), özellikle çek-senet ve sokak tabelaları üzerindeki rakamların tanınması amacıyla oluşturulmuş bir ağıdır. Genel itibari ile bir Evrişimli Sinir Ağı (ESA) 4 katmandan oluşur. Bunlar: Evrişim katmanı, ortaklama katmanı ve düzleştirme katmanı ve tam bağlantı katmanıdır.

Evrişim Katmanı (Convolutional Layer), bilgisayarlar için görüntüler, satır ve sütunlardan oluşan matrisler ile ifade edilir. Görüntünün en küçük birimi ise piksel olarak ifade edilir (Gonzalez ve Rafael, 2018). ESA'da bir görüntünün sınıflandırma, bölütleme veya başka bir işlem için kullanılması için filtre matrisi denilen öznitelik çıkarım matrislerinden faydalanılır. Filtre matrisi, resim (görüntü matrisi) üzerinde satır satır (stride) kaydırılarak görüntü matrisinden öznitelikler çıkarır. Bu işlem aslında basit bir çarpma ve toplama işlemi olarak düşünülebilir. Şekil 13'te evrişim işleminin nasıl gerçekleştiği gösterilmektedir.



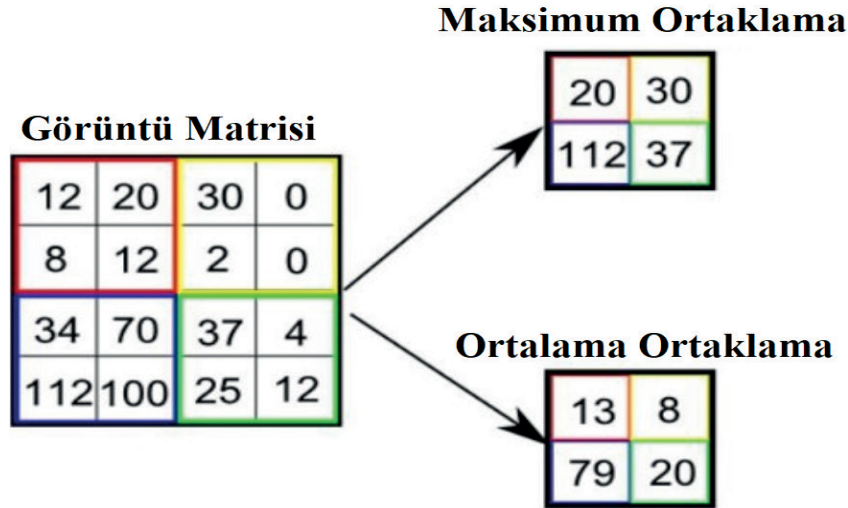
Şekil 13. Filtre matrisinin orta noktasının görüntü matrisinin ilk pikseli üzerine yerleştirilerek evrşim işleminin başlatılması

Şekil 13'te gösterilen evrşim işleminin sonucunda elde edilen yeni görüntü matrisi Şekil 14'te gösterilmektedir.



Şekil 14. Görüntü matrisinin; (a) ilk değerinden başlanarak, (b) son değerine ulaşıncaya kadar filtre matrisi ile evrşim işlemine girmesi sonucunda oluşan (c) yeni görüntü matrisi.

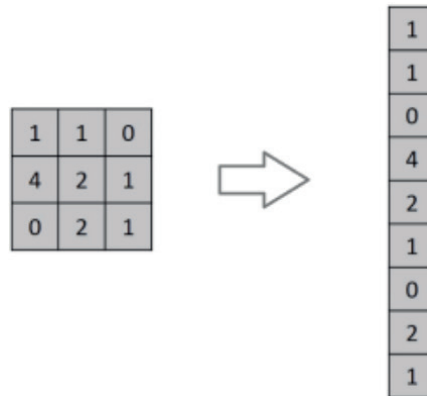
Ortaklama Katmanı (Pooling Layer), klasik ESA'da alt örnekleme (subsampling) olarak ifade edilmektedir (LeCun, Bottou, Bengio ve Haffner, 1998). Ortaklama katmanı, görüntünün boyutunu azaltmak için kullanılır. 2 tür ortaklama vardır: ortalama ortaklama (average pooling) ve maksimum ortaklama (maximum pooling). Ortalama ortaklama, giriş görüntüsünde n adet pikseli toplayıp n sayısına böler. Maksimum ortaklama, görüntü matrisi üzerindeki en yüksek piksel değerini baz alır. Şekil 15'te ortalama ortaklama ve maksimum ortaklama gösterilmektedir.



Şekil 15. Görüntü matrisinin maksimum ortaklama/ortalama ortaklama ile boyutunun azaltılması (<https://towardsdatascience.com/>)

Şekil 15'te gösterilen örnekte 4x4'lük 16 pikselden oluşan 2 boyutlu bir görüntü matrisi, ortaklama işlemi sonucunda 2x2'lik 4 piksellik bir görüntü haline getirilmiştir.

Düzleştirme Katmanı (Flattening Layer), ESA'daki bu katmanın görevi, görüntüyü yapay sinir ağına (YSA) iletmek için vektör (tek sütunluk dizi) haline getirmektir. Şekil 16'da tam bağlantı katmanında vektör haline getirilen bir görüntü matrisi gösterilmektedir.



Şekil 16. Düzleştirme katmanı sayesinde 3x3'lük bir görüntünün 1x9'lük bir vektöre dönüştürülmesi

Tam Bağlantı Katmanı (Fully Connected Layer), ESA'nın son katmanıdır ve düzleştirme katmanında vektör haline getirilen verileri YSA ile öğrenme işleminin gerçekleşmesini sağlar.

4.1.2. Artık Ağlar (Residual Networks)

Derin öğrenme, büyük verilerle uğraşmaktadır. Bilgisayarlı görü alanında yapılan çalışmalarda genellikle önceden eğitilmiş ağlar (pre-trained) kullanılmaktadır. Bunun sebebi, bir problemin çözümünde yeniden bir ağ kurmak yerine, benzer bir problem için oluşturulmuş ve kullanılmış bir ağın transfer edilerek kullanılması hem zaman açısından hem de maliyet açısından oldukça kazançlı bir fikirdir. Transfer öğrenme, bu tür sorunlar için çok iyi bir seçenektir.

Artık ağlar (ResNet), 2015 yılında düzenlenen "The ImageNet Large Scale Visual Recognition Challenge (ILSVRC)" yarışmasında %3.57 hata oranı ve %95.51'lik başarı oranı ile birinciliği kazanan bir Derin Nöral Ağ (Deep Neural Network) mimarisidir (He, Zhang, Ren ve Sun, 2016).

Özellikle ESA'ların gelişmesi ile ağlar, genişlemeye değil derinleşmeye başladı. ResNet, artık değerlerin (residual value) yani öğrenme sırasında öğrenme işleminin tam başarılı olmadığını değerlerin mevcudiyeti durumunda, bir önceki katmandan gelen çıktı değerinin bu değere eklenerek öğrenme işlemine katkıda bulunmasını mümkün kılan bir mimaridir.

4.1.3. K-En Yakın Komşu

Denetimli öğrenmede sıklıkla kullanılan bir makine öğrenmesi tekniğidir. Diğer denetimli öğrenme yöntemlerinin aksine eğitim kısmı yoktur. Bu algoritmada eğitim ve test neredeyse birbirine benzerdir. Sınıfları zaten belli olan örnek bir veri setinden faydalanarak kullanılır. Veri setine eklenmek istenen yeni verinin, diğer verilere olan uzaklığı, kullanılan uzaklık türüne göre hesaplanır. K sayıdaki en yakın komşulara bakılır ve elde edilen sonuç hangisine yakınsa o veri grubuna göre etiketleme yapılır. Tüm olasılıklar hesaplandığı için donanım açısından oldukça maliyetlidir. Yüksek bellek gerektirir (Cover ve Hart, 1967).

4.1.4. Yapay Sinir Ağları (YSA)

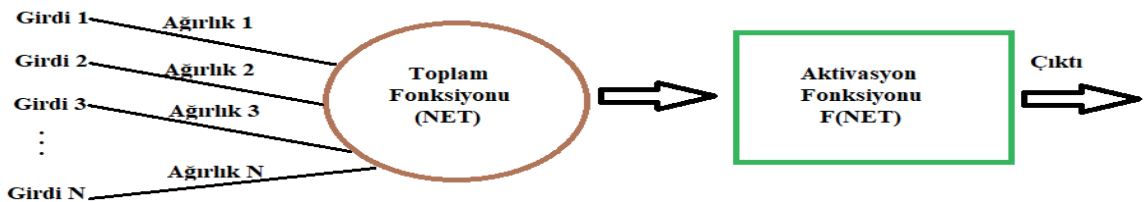
Yapay Sinir Ağları (YSA), insanlara ait gerçek beyin fonksiyonlarının ürünü olan örneklerden yola çıkarak öğrenebilen ve bu öğrenmenin sonucunda kendi karşılaştığı sorunlar karşısında nasıl çözümler üretileceğini belirleyebilen sistemlerdir (Öztemel, 2006). YSA'lar insan beyninin çalışma prensibinden yola çıkılarak geliştirilmiştir. Yapay sinir hücreleri bir araya gelerek yapay sinir ağını oluştururlar. Bu hücrelerin bir araya gelmeleri tesadüfi değildir. Belli katmanlara ayrılıp, katmanların içerisinde paralel bir şekilde dizilerek sinir ağını meydana getirirler. Bir YSA'da genel olarak 3 tip katman bulunmaktadır:

Girdi Katmanı (Input Layer): Çevreden alınan bilgileri bir sonraki katmana (ara katmana) iletmekle görevlidirler. Bu katmanda herhangi bir bilgi işleme söz konusu değildir.

Ara Katman (Hidden Layer): Girdi katmanından alınan bilgiler bu katmanda işlenmektedir. Bir YSA'da birden fazla sayıda ara katman bulunabilir. Bilgi işleme sonrası, bilgiler bu katmandan çıktı katmanına gönderilir.

Çıktı Katmanı (Output Layer): Bu katmanda yer alan hücreler, ara katmandan gönderilen bilgileri işleyerek girdi katmanının çevreden aldığı bilgilere örnek teşkil etmesi gereken çıktıları üretir ve çevreye (dış dünyaya, kullanıcıya) gönderir (Öztemel, 2006).

YSA'larda girdi katmanından alınan bilgiler, ağırlıklar (weights) ile çarpılarak bir sonraki katmana (ara katman) aktarılır ve bir toplama fonksiyonundan geçirilirler. Bu durum, girdi katmanının tüm nöronlarından gelen bilgilerin ilgili ağırlıklar ile çarpılarak toplanmasıdır. Daha sonra çıkış katmanına varıncaya dek her ara katmanda bu işlem tekrar edilir ve buna "ileri besleme (feed forward)" denilir. Çıktı katmanına gelindiğinde ise bu değerler bir aktivasyon fonksiyonundan geçirilirler ve nihai çıktı elde edilir. Elde edilen son çıktı (tahmin edilen çıktı) ile olması gereken çıktı arasındaki fark "hata" olarak adlandırılır ve bu hata, istenen düzeye indirgenmek için geriye dönük olarak tüm ağa yayılır. Hatanın tüm ağa yayılması ise "geriye yayılım (back propagation)" olarak isimlendirilir. Burada amaç, ağırlıkların güncellenmesidir. İleri besleme ve geriye yayılım ile öğrenme işlemi gerçekleşmiş olur. Şekil 17'de YSA hücrelerinin yapısı gösterilmektedir.

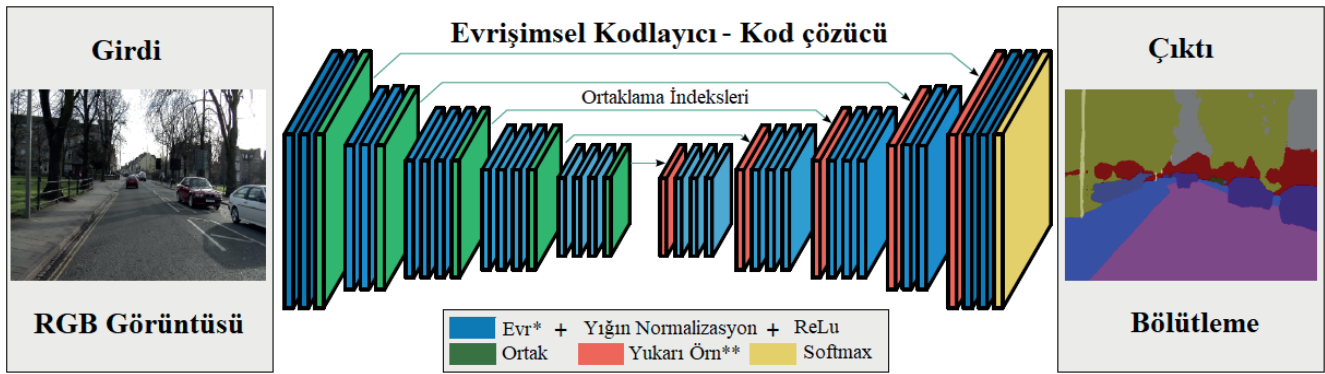


Şekil 17. YSA'nın yapısı

4.1.5. SegNet

SegNet, anlamsal bir bölütleme (segmentasyon) modelidir. Kodlayıcı (encoder) ve kod çözücü (decoder) katmanlardan oluşan

bir yapıya sahiptir. SegNet üzerindeki kodlayıcılar evrişim ve yığın normalizasyon (batch normalization) işlemleri uygular. Bu işlemlerden sonra her çerçeveden çıkardığı indeks değerini saklar ve bu işlemlerin sonucunda elde ettiği sonuç üzerinde maksimum ortaklama işlemini gerçekleştirir. SegNet üzerindeki kodlayıcılar mimari açıdan VGG16 ağı ile aynıdır. Kod çözümler ise düşük çözünürlüğe sahip kodlayıcı öznitelik haritalarını piksel temelli sınıflandırma yapmak için eksiksiz bir çözünürlüklü öznitelik haritalarına eşler. SegNet'in en önemli özelliği düşük çözünürlüğe sahip olan giriş görüntülerinin özniteliklerini yüksek örnekleme (upsampling) ile artırmaktır (Badrinarayanan, Kendall ve Cipolla, 2017). Şekil 18'de SegNet mimarisi ve yapılan işlemler sonucunda araçların ve çevrenin anlamsal bölütlemeye çıktısı verilmektedir.



Şekil 18. SegNet mimarisi (Badrinarayanan, Kendall ve Cipolla, 2017).

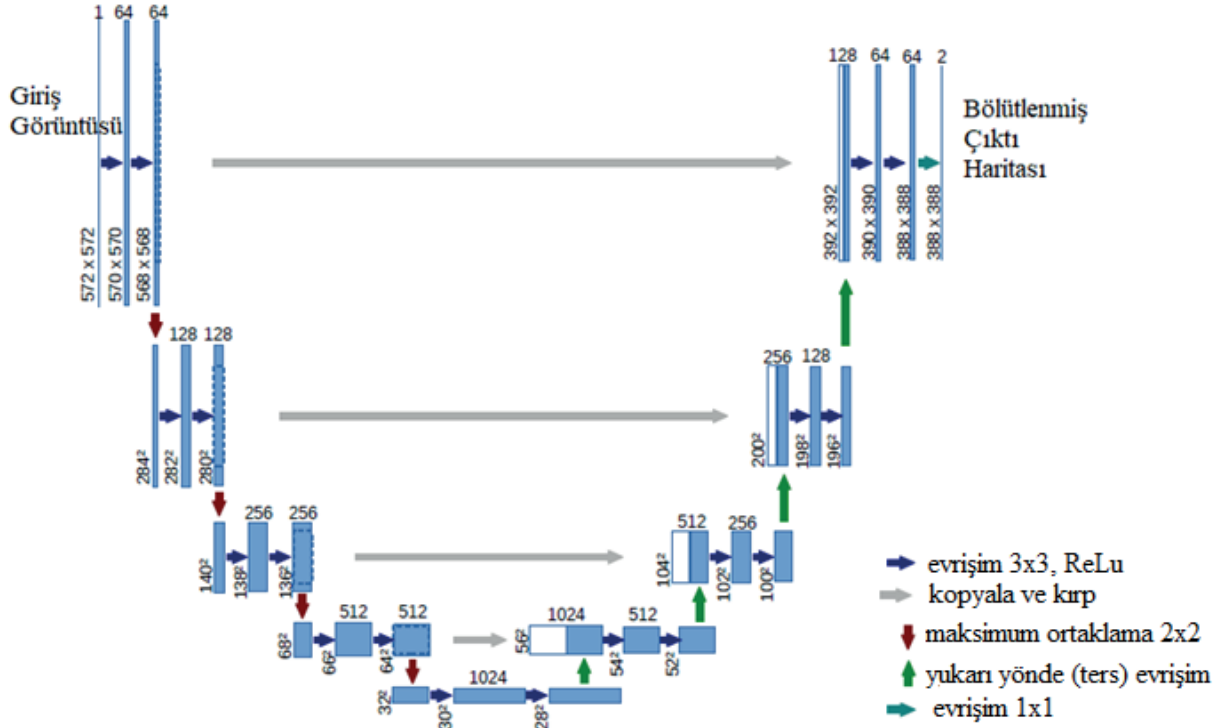
Not: *=Evrişim, **=Yukarı örnekleme

4.1.6. Dalgacık Dönüşümü (Wavelet Transform)

Dalgacık dönüşümü, sinyallerin zaman-frekans gösteriminin sağlanması için kullanılan matematiksel bir dönüşüm türüdür (Alp, Akıncı ve Albora, 2008). Dalgacık dönüşümü esnasında ana dalgacığın muhtemel bütün ölçekleri ve sinyalin bütün zaman aralığı kullanıldığında hesaplanması gereken katsayılar oldukça büyük veri yığınlarına neden olur. Bu sebeple, belirli bazı ölçek ve konum aralıklarında dönüşüm gerçekleştirilir. Bu dönüşüm işleminin çok hızlı bir şekilde gerçekleşmesini sağlamak için ölçek ve konum verileri 2'nin üsleri olacak şekilde seçilir (Öner, Yeşilyurt ve Yılmaz, 2017; Alp, Akıncı ve Albora, 2008; Coşkun, 2019).

4.1.7. U-net

Derin öğrenme, genellikle çok sayıda veri ve bunların etiketlenmesi gibi süreçleri gerektirir. Özellikle elde edilmesi zor görüntülerde (tıbbi görüntü, uydu görüntüsü gibi) hem bu görüntülerin elde edilmesi hem de piksel bazlı konum verilerine ihtiyaç duyulmaktadır. Bu durum bazen sorunlara yol açabilir. U-net mimarisi, ESA katmanlarından oluşan ve YSA katmanı barındırmayan yapısı ile piksel bazlı görüntü bölütleme konusunda oldukça başarılı sonuçlar vermektedir. Şekil 19'da U-net model mimarisi verilmektedir.



Şekil 19. U-net mimarisi (Ronneberger vd., 2015)

U-net modeli, Şekil 19'dan da görüleceği üzere adını model mimarisinden alır. U-net, aşağı yönde bir büzülme yolu (contracting path) ve yukarı yönde bir açılma yolundan (expansive path) oluşur. Büzülme yolunda, klasik ESA katmanlarının işlemleri meydana gelir. Bunlar sırasıyla iki defa 3x3'lük bir evrişim işlemi, geri dönüşümlü doğrusal birim fonksiyonu (ReLU) ve 2 adımlık (stride) bir 2x2'lik maksimum ortaklama işlemleridir. Açılma yolunda ise öznelik kanallarını yarıya indiren 2x2'lik evrişim işlemi, büzülme yolundaki eş katmanlara karşılık gelen kırpılmış özellik eşleşmesi ile birleştirme işlemi ve 3x3'lük evrişim işleminin iki defa tekrarlanması işlemleri gerçekleşir (Ronneberger vd., 2015).

4.1.7. Çekişmeli Üretici Ağlar (Generative Adversarial Networks)

Derin öğrenme alanında, görüntü sınıflandırması, bölütleme, kümeleme vb. işlemler yapılmasına olanak sağlayan birçok model, mimari geliştirilmiştir. Çekişmeli Üretici Ağlar (ÇÜA), ilk defa görüntü üretebilen bir model ortaya koymuştur (Goodfellow ve ark., 2014). Klasik ESA'lardan farklı olarak ÇÜA, kendi içerisinde bir üretici (generative) ve bir ayırıcı (discriminator) olarak adlandırılan 2 farklı derin ağ barındırmaktadır. ÇÜA'nın işleyişi oldukça basittir. Üretici, veri setindeki görüntülerden yola çıkarak gerçek görüntülere benzeyen sentetik (sahte) görüntüler üretir ve ayırıcıyı kandırmaya çalışır. İki ağ arasında da bir nevi paralel bir rekabet (çekişme) vardır ve buna göre eğitilirler. Üretici ağ, belirli bir aşamadan sonra ayırıcı ağın ayırt edemeyeceği derecede yani gerçek görüntülere çok benzeyen görüntüler üretene kadar bu işlem devam eder. ÇÜA'lar son yıllarda üzerine en fazla çalışma yapılan derin öğrenme konularının başında gelmektedir (Goodfellow ve ark., 2014). Görüntü üretiminden video üretime, çözünürlük artırımından doku sentezine varıncaya dek oldukça geniş bir yelpazeye ve başarılı sonuçlara sahiptir (Çelik ve Talu, 2019; Wu, Xu, Hall, 2017).

4.2. Yöntemlerin Değerlendirilmesi

Çalışmanın önceki alt bölümünde, Tablo 1'deki literatür çalışmalarında kullanılan yöntemlerin genel itibarı ile yapısı anlatılmıştır. Bilgisayar donanımlarının gücünün gelişmesi ile beraber derin öğrenme alanında yapılan çalışmaların sayısı artmış ve içeriği oldukça zenginleşmiştir. Gerek makine öğrenmesinde gerek derin öğrenmede bir problem alanında kesin ve en iyi sonucu veren sabit bir yöntem ve parametre bulunmamaktadır. Elde edilen başarılı sonuçlara rağmen her geçen gün yapılan çalışmalar da bu değerlendirmeyi desteklemektedir. Ancak derin öğrenme alanında belirli görüntü veri setleri

üzerinde sınıflandırma, bölütleme gibi amaçlarla geliştirilen modellerin parametre sayısı o modellerin tercih edilmesinde etkili olabilir. Çünkü bu parametreler doğrudan donanım gücü ile ilişkilidir. Ele alınan sorunun içeriğine, yapılmak istenen işleme ve çalışmanın amacına göre bir model ve parametre seçimi yapılabilir. Bu çalışmada verilen yöntemler birbirlerinden farklı sorunlara, amaçlara ve metriklere yönelik olduğu için bu yöntemlerin kıyaslanması eksik ve hatalı olacaktır. Burada asıl odaklanılması gereken nokta, genel itibari ile yapay zekânın tıbbi görüntüleme alanında ne denli geniş bir yelpazede kullanıldığıdır.

5. TARTIŞMA ve SONUÇ

Günümüzde yapay zekânın kullanım alanları artmaktadır. Tıbbi görüntülemede yapay zekânın kullanımı sayesinde uzmanların farklı tıbbi görüntüleme teknolojilerinden elde edilen görüntüleri daha iyi yorumlaması, hastalık tanısı konulması gibi işlemler zaman ve maliyet açısından daha ekonomik kullanılmakta ve bu işlemlerin otomatikleştirilmesi sağlanmaktadır. Kullanılan enerji türüne göre tıbbi görüntüleme teknolojileri ve bunlardan elde edilen görüntüler ile bu görüntülerin yorumlanması da farklılık göstermektedir. Bu durum, uzmanların işini zorlaştırmaktadır. Bu nedenle tıbbi görüntüleme alanında yapay zekânın kullanımı ile ilgili çalışmalar gün geçtikçe artmakta ve geliştirilmektedir.

Bu çalışma, tıbbi görüntülemede yapay zekânın kullanımı ile ilgili yapılacak çalışmalara dair bir teorik bir temel inşa etmek ve genel bir bakış açısı sunmak için amacıyla hazırlanmıştır. Literatürde özellikle nükleer tıp görüntüleme başta olmak üzere tıbbi görüntüleme teknolojilerini kapsayan, bu teknolojileri açıklayan çalışmaların yetersiz olmasından dolayı söz konusu eksiklik bu çalışmada bir bütünlük içerisinde giderilmiştir. Ayrıca, tıbbi görüntülemede yapay zekâ kullanımının mevcut durumu, geleceği, bu alanda karşılaşılan sorunlar ve bunların çözümleri gibi hususların açık bir şekilde ele alınması, çalışmanın literatüre olan katkısını göstermektedir. Tıbbi görüntülemede yapay zekâ kullanımını tartışma ve sonuç bölümünde ele alınmaması, bu durumun bir başlık altında özel olarak incelenmesi ve hemen arkasından tıbbi görüntülemede yapay zekâ kullanımının popüler ve işlevsel bir örneği olan bilgisayar destekli tanı sistemlerinin verilmesi, çalışmanın anlam bütünlüğünü sağlamaya yönelik özel bir tutumdur. Tıbbi görüntülemede yapay zekâ kullanımına ilişkin olarak Bilgisayar Destekli Tanı Sistemlerinin ve diğer gelişmelerin, sağlık çalışanlarının ve bilhassa doktorların rakibi olmayıp aksine onların yorum-teşhis gücünü artırmaya yönelik yardımcı uygulamalar olması hususuna değinilmiştir. Çalışmada sadece tıbbi görüntüleme teknolojileri açıklanmamış bu alanda yapay zekânın kullanımı ile ilgili her bir tıbbi görüntüleme tekniği için yapılmış güncel ve başarılı çalışmalar da sonuçları ile birlikte verilmiştir. Dolayısıyla literatürdeki tıbbi görüntüleme teknolojilerinin kapsayıcı, açıklayıcı ve yapay zekâ ile ilişkisinin net bir şekilde ortaya konulması ile ilgili boşluk, bu çalışma ile doldurulmaya çalışılmıştır.

İleriki çalışmalarda tıbbi görüntü formatları, bu görüntülerin ön işleme ile nasıl kullanıma hazır hale getirildiği ve tıbbi görüntü veri setleri gibi konularda da çalışmaların gerçekleştirilmesi faydalı olacaktır. Özellikle gerçek tıbbi görüntülere erişim, bu alanda çalışan araştırmacılar için ciddi bir zorluktur. Bu sebeple, tıbbi görüntüleri açık erişime sunan hastaneler, vakıflar, kliniklerden elde edilen görüntüleri bir araya toplayan kişi ya da kurumların oluşturdukları açık veri setlerinin tanıtılması, tıbbi görüntülerin yapay zekâ ile işlenmesi ve çalışılması, ilgili konularda çalışan araştırmacılara büyük kolaylık sağlayacaktır.

Hakem Değerlendirmesi: Dış bağımsız.

Çıkar Çatışması: Yazarlar çıkar çatışması bildirmemiştir.

Finansal Destek: Yazarlar bu çalışma için finansal destek almadığını beyan etmiştir.

Peer-review: Externally peer-reviewed.

Conflict of Interest: The authors have no conflict of interest to declare.

Grant Support: The authors declared that this study has received no financial support.

Kaynaklar/References

- 3 Boyutlu Mamogram Görüntüsü. WAKE RADIOLOGY. (2020, 8 Haziran). Erişim Adresi: <https://www.wakerad.com/whats-new/study-3-d-scans-accurate-standard-mammograms/>
- ABD’li ve Çinli şirketler Yapay Zekaya hükmetmek için yarışıyor. (2020, 17 Ekim). Erişim Adresi: <https://www.wsj.com/articles/why-u-s-companies-may-lose-the-ai-race-1516280677>
- Alp, H., Akıncı, T. Ç., & Albora, M. (2008). Jeofizik Uygulamalarda Fourier ve Dalgacık Dönüşümlerinin Karşılaştırılması. *Pamukkale Üniversitesi Mühendislik Bilimleri Dergisi*, 14(1), 67-76.
- Al-shamasneh, A. R. M., & Obaidallah, U. H. B. (2017). Artificial intelligence techniques for cancer detection and classification: review study. *European Scientific Journal*, 13(3), 342-370.
- Arslan, T. X Işınları ve Kullanım Alanları. Gazi Üniversitesi Tezi, 2010.
- Badrinarayanan, V., Kendall, A., & Cipolla, R. (2017). Segnet: A deep convolutional encoder-decoder architecture for image segmentation. *IEEE transactions on pattern analysis and machine intelligence*, 39(12), 2481-2495.
- Bailey, D. L., Maisey, M. N., Townsend, D. W., & Valk, P. E. (2005). *Positron emission tomography* (Vol. 2). London: Springer.
- Bankman, I. N. (2009). *Handbook of medical image processing and analysis*. Boston: Academic Press.
- Bray, F., Ferlay, J., Soerjomataram, I., Siegel, R. L., Torre, A., & Jemal, A. (2018). Global cancer statistics 2018: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. *CA: a cancer journal for clinicians*, vol. 6, no. 8, pp. 394-424.
- Buck, A. K., Nekolla, S., Ziegler, S., Beer, A., Krause, B. J., Herrmann, K., ... & Drzezga, A. (2008). Spect/ct. *Journal of Nuclear Medicine*, 49(8), 1305-1319.
- Budak, Ü. (2019). SegNet Mimarisi ile Bilgisayarlı Tomografi Görüntülerinden Karaciğer Bölgesinin Bölütlenmesi. *Fırat Üniversitesi Mühendislik Bilimleri Dergisi*, 31(1), 215-222.
- Bushberg, J. T., Seibert, J. A., Leidholdt Jr, E. M., Boone, J. M., & Goldschmidt Jr, E. J. (2011). Introduction to Medical Imaging. The essential physics of medical imaging. 3 rd edition: Lippincott Williams & Wilkins.
- Buxton, R. B. (2009). *Introduction to Functional Magnetic Resonance Imaging Principles and Techniques*. Cambridge: Cambridge University Press.
- Cai, C., Wang, C., Zeng, Y., Cai, S., Liang, D., Wu, Y., Chen, Z., Ding, X. and Zhong, J. (2018). Single-shot T2 mapping using overlapping-echo detachment planar imaging and a deep convolutional neural network. *Magnetic resonance in medicine*, 80(5), 2202-2214.
- Cheng, H. D., Cai, X., Chen, X., Hu, L., & Lou, X. (2003). Computer-aided detection and classification of microcalcifications in mammograms: a survey. *Pattern recognition*, 36(12), 2967-2991.
- Chow, J. C., Boyd, S. K., Lichti, D. D., & Ronsky, J. L. (2020). Robust Self-Supervised Learning of Deterministic Errors in Single-Plane (Monoplanar) and Dual-Plane (Biplanar) X-ray Fluoroscopy. *IEEE Transactions on Medical Imaging*, 39(6), 2051-2060.
- Coşkun, Y. (2019). *Ayrık dalgacık dönüşümü tabanlı paralel görüntü sıkıştırma sistemi tasarımı* (Master’s thesis, Maltepe Üniversitesi, Fen Bilimleri Enstitüsü).
- Cover, T., & Hart, P. (1967). Nearest neighbor pattern classification. *IEEE transactions on information theory*, 13(1), 21-27.
- Çelik, G., & Talu, M. F. (2019). Çekişmeli üretken ağ modellerinin görüntü üretme performanslarının incelenmesi. *Balıkesir Üniversitesi Fen Bilimleri Enstitüsü Dergisi*, 22(1), 181-192.
- Dandıl, E., Serin, Z. (2020). Derin Sinir Ağları Kullanarak Histopatolojik Görüntülerde Meme Kanseri Tespiti . *Avrupa Bilim ve Teknoloji Dergisi* , Ejosat Özel Sayı 2020 (HORA): 451-463.
- Deng, J., Berg, A., Satheesh, S., Su, H., Khosla, A., & Fei-Fei, L. (2012). Imagenet large scale visual recognition competition 2012 (ILSVRC2012).
- Doi, K. (2007). Computer-aided diagnosis in medical imaging: historical review, current status and future potential. *Computerized medical imaging and graphics*, 31(4-5), 198-211.
- Dong, X., Lei, Y., Wang, T., Higgins, K., Liu, T., Curran, W. J., ... & Yang, X. (2020). Deep learning-based attenuation correction in the absence of structural information for whole-body positron emission tomography imaging. *Physics in Medicine & Biology*, 65(5), 055011.
- Düzleştirme Katmanı ile görüntü matrislerinin vektör hâline getirilmesi. (2021, 14 Ocak). Erişim Adresi: <https://medium.com/@tuncerergin/convolutional-neural-network-convnet-yada-cnn-nedir-nasil-calisir-97a0f5d34cad>
- Ehrlich, Ruth Ann., Daly, Joan A. (2008). Patient Care in Radiography: With an introduction to medical imaging (Seventh Edition). MO: Elsevier.
- El-Baz, A. S., & Suri, J. S. (2019). *Lung imaging and CADx*. Boca Raton, FL: CRC Press/Taylor & Francis.
- El-Baz, A., Beache, G. M., Gimel’farb, G., Suzuki, K., Okada, K., Elnakib, A., ... & Abdollahi, B. (2013). Computer-aided diagnosis systems for lung cancer: challenges and methodologies. *International journal of biomedical imaging*.
- Ertel, W., Black, N., & Mast, F. (2017). *Introduction to artificial intelligence*. Cham, Switzerland: Springer.
- European Society of Radiology (ESR). (2015). ESR position paper on imaging biobanks. *Insights into imaging*, 6(4), 403-410.
- Floroskopi Prosedürü ve Görüntüsü, Angela Betsaida B. Laguipo. (2020, 8 Haziran). Erişim Adresi: <https://www.news-medical.net/health/Fluoroscopy-Procedure.aspx>
- Fotin, S. V., Yin, Y., Haldankar, H., Hoffmeister, J. W., & Periaswamy, S. (2016, March). Detection of soft tissue densities from digital breast tomosynthesis: comparison of conventional and deep learning approaches. In *Medical Imaging 2016: Computer-Aided Diagnosis* (Vol. 9785, p. 97850X). International Society for Optics and Photonics.
- Ghoneim, A., Muhammad, G., & Hossain, M. S. (2020). Cervical cancer classification using convolutional neural networks and extreme learning machines. *Future Generation Computer Systems*. 102: 643-649.

- Gonzalez, Rafael (2018). *Digital image processing*. New York, NY: Pearson.
- Goodfellow, I., Pouget-Abadie, J., Mirza, M., Xu, B., Warde-Farley, D., Ozair, S., ... & Bengio, Y. (2014). Generative adversarial nets. *Advances in neural information processing systems*, 27, 2672-2680.
- Greenspan, H., Van Ginneken, B., & Summers, R. M. (2016). Guest editorial deep learning in medical imaging: Overview and future promise of an exciting new technique. *IEEE Transactions on Medical Imaging*, 35(5), 1153-1159.
- Gulli, A., Kapoor, A., & Pal, S. (2019). Deep learning with TensorFlow 2 and Keras: Regression, ConvNets, GANs, RNNs, NLP, and more with TensorFlow 2 and the Keras API. Birmingham: Packt Publishing.
- Haidekker, M. A. (2013). *Medical Imaging Technology*. New York, NY: Springer.
- He, K., Zhang, X., Ren, S., & Sun, J. (2016). Deep residual learning for image recognition. In *Proceedings of the IEEE conference on computer vision and pattern recognition* (pp. 770-778).
- Hou, W., Zhang, D., Wei, Y., Guo, J., & Zhang, X. (2020). Review on Computer Aided Weld Defect Detection from Radiography Images. *Applied Sciences*, 10(5), 1878.
- Huettel, S. A., Song, A. W., & McCarthy, G. (2004). *Functional magnetic resonance imaging* (Vol. 1). Sunderland, MA: Sinauer Associates.
- Hutton, L., & Henderson, T. (2017). Beyond the EULA: Improving consent for data mining. In *Transparent Data Mining for Big and Small Data* (pp. 147-167). Cham: Springer International Publishing.
- ILSVRC2012. Large Scale Visual Recognition Challenge 2012. (2020, 22 Aralık). Erişim Adresi: <https://www.kdnuggets.com/2018/12/deep-learning-major-advances-review.html>
- Imagenet'te Görüntü Sınıflandırması. (2020, 6 Eylül). Erişim Adresi: <https://paperswithcode.com/sota/image-classification-on-imagenet>
- Kiani, A., Uyumazturk, B., Rajpurkar, P., Wang, A., Gao, R., Jones, E., ... & Martin, B. A. (2020). Impact of a deep learning assistant on the histopathologic classification of liver cancer. *NPJ digital medicine*. 3(1): 1-8.
- Krizhevsky, A., Sutskever, I., & Hinton, G. E. (2012). Imagenet classification with deep convolutional neural networks. In *Advances in neural information processing systems*. 1097-1105.
- Kubat, M. (2018). *Introduction To Machine Learning*. Place of publication not identified: SPRINGER INTERNATIONAL PU.
- Lameka, K., Farwell, M. D., & Ichise, M. (2016). Positron emission tomography. In *Handbook of clinical neurology* (Vol. 135, pp. 209-227). Amsterdam: Elsevier.
- Larson, C., Lionhart, P., Roh, A., & Colglazier, R. (2018). *Introduction to fluoroscopy: For residents & professionals alike*. Place of publication not identified: Prometheus Lionhart.
- LeCun, Y., Bottou, L., Bengio, Y., & Haffner, P. (1998). Gradient-based learning applied to document recognition. *Proceedings of the IEEE*, 86(11), 2278-2324.
- Leighton, T. G. (2007). What is ultrasound? *Progress in biophysics and molecular biology*, 93(1-3), 3-83.
- Litjens, G., Kooi, T., Bejnordi, B. E., Setio, A. A. A., Ciompi, F., Ghafoorian, M., ... & Sánchez, C. I. (2017). A survey on deep learning in medical image analysis. *Medical image analysis*, 42, 60-88.
- Maulik, D., & Zalud, I. (Eds.). (2005). Doppler ultrasound in obstetrics and gynecology (pp. 363-374). Berlin: Springer.
- Mettler, F. A., & Guiberteau, M. J. (2012). *Essentials of nuclear medicine imaging*. Philadelphia, PA: Elsevier/Saunders.
- Morra, L., Delsanto, S., & Correale, L. (2019). *Artificial Intelligence in Medical Imaging: From Theory to Clinical Practice*. Boca Raton: CRC Press.
- Murino, V., Puppo, E., Sona, D., Cristani, M., & Sansone, C. (Eds.). (2015). *New Trends in Image Analysis and Processing--ICAP 2015 Workshops: ICAP 2015 International Workshops, BioFor, CTMR, RHEUMA, ISCA, MADiMa, SBMI, and QoEM, Genoa, Italy, September 7-8, 2015, Proceedings* (Vol. 9281). Heidelberg: Springer.
- Neri, E., & Regge, D. (2017). Imaging biobanks in oncology: European perspective. *Future Oncology*, 13(5), 433-441.
- Ortaklama katmanı türleri. (2021, 14 Ocak). Erişim Adresi: <https://towardsdatascience.com/a-comprehensive-guide-to-convolutional-neural-networks-the-eli5-way-3bd2b1164a53>
- Öner, İ. V., Yeşilyurt, M. K., & Yılmaz, E. Ç. (2017). WAVELET ANALİZ TEKNİĞİ VE UYGULAMA ALANLARI. *Ordu Üniversitesi Bilim ve Teknoloji Dergisi*, 7(1), 42-56.
- Öztad, E. (2020). Meme Kanseri Tespitinde Sınıflandırma ve Sinir Ağları Yöntemlerinin Karşılaştırılması. İleri Mühendislik Çalışmaları ve Teknolojileri Dergisi. 1(1): 49-54.
- Öztemel, E. (2006). Yapay Sinir Ağları. 2. Baskı. Papatya Yayıncılık: İstanbul.
- Pisano, E. D., & Yaffe, M. J. Digital mammography. *Radiology*, vol. 234, no. 2, pp. 353-362, 2005.
- Ranschaert, E. R., Morozov, S., & Algra, P. R. (2019). *Artificial Intelligence in Medical Imaging: Opportunities, Applications and Risks*. Cham: Springer International Publishing.
- Ronneberger, O., Fischer, P., & Brox, T. (2015). U-net: Convolutional networks for biomedical image segmentation. In *International Conference on Medical image computing and computer-assisted intervention* (pp. 234-241). Springer, Cham.
- Sağlık hizmetleri verileri ve gevşek kurallar Çin'in Yapay Zekâ konusunda başarılı olmasına nasıl yardımcı oluyor? (2020, 17 Ekim). Erişim Adresi: <https://www.wired.com/story/health-care-data-laxrules-help-china-prosper-ai/>
- Sahiner, B., Pezeshk, A., Hadjiiski, L. M., Wang, X., Drukker, K., Cha, K. H., ... & Giger, M. L. (2019). Deep learning in medical imaging and radiation therapy. *Medical physics*, 46(1), e1-e36.
- Sarhan, A. M. (2020). Brain Tumor Classification in Magnetic Resonance Images Using Deep Learning and Wavelet Transform. *Journal of Biomedical Science and Engineering*, 13(06), 102.

- Sato, E., Nakayama, K., Nakamura, K., Ishikawa, M., Katagiri, H., & Kyo, S. (2015). A case with life-threatening uterine bleeding due to postmenopausal uterine arteriovenous malformation. *BMC Women's Health*, 15(1), 1-5.
- Shi, L., Onofrey, J. A., Liu, H., Liu, Y. H., & Liu, C. (2020). Deep learning-based attenuation map generation for myocardial perfusion SPECT. *European Journal of Nuclear Medicine and Molecular Imaging*, 1-13.
- Shia, W., & Chen, D. (2020). Abstract P1-02-10: Using deep residual networks for malignant and benign classification of two-dimensional Doppler breast ultrasound imaging.
- Shung, K. K. (2015). *Diagnostic ultrasound: Imaging and blood flow measurements*. Boca Raton, BR: CRC press.
- Sluimer, I., Schilham, A., Prokop, M., & Van Ginneken, B. (2006). Computer analysis of computed tomography scans of the lung: a survey. *IEEE transactions on medical imaging*, 25(4), 385-405.
- Smith, N. & Webb, A. (2010). Nuclear medicine: Planar scintigraphy, SPECT and PET/CT. In *Introduction to Medical Imaging: Physics, Engineering and Clinical Applications* (Cambridge Texts in Biomedical Engineering, pp. 89-144). Cambridge: Cambridge University Press. doi:10.1017/CBO9780511760976.003
- SPECT görüntüleri ile Alzheimer Hastalığı Teşhisi. (2020, 10 Haziran). Erişim Adresi: <https://www.pinterest.co.uk/pin/350717889706920247/>
- Szabo, T.L. Doppler Models. (2004). *Diagnostic ultrasound imaging: inside out*. 2 nd edition: Academic Press.
- Talo, M. (2019, April). Pneumonia Detection from Radiography Images using Convolutional Neural Networks. In *2019 27th Signal Processing and Communications Applications Conference (SIU)* (pp. 1-4). IEEE.
- The Economist. "Çalışmamaktan Sinir Ağına". (2020, 3 Eylül). Erişim Adresi: <https://www.economist.com/special-report/2016/06/23/from-not-working-to-neural-networking>
- The, L. (2018). Artificial intelligence in health care: within touching distance. *Lancet (London, England)*, 390(10114), 2739.
- Tiryaki, V. (2020). Mamografi görüntülerindeki anormalliklerin yerel ikili örüntü ve varyantları kullanılarak sınıflandırılması. *Bitlis Eren Üniversitesi Fen Bilimleri Dergisi*, 9 (1), 297-305 . DOI: 10.17798/bitlisfen.557411.
- Vakanski, A., Xian, M., & Freer, P. E. (2020). Attention-Enriched Deep Learning Model for Breast Tumor Segmentation in Ultrasound Images. *Ultrasound in Medicine & Biology*, 46(10), 2819-2833.
- Wang, H., & Raj, B. (2017). On the origin of deep learning. arXiv preprint arXiv:1702.07800.
- Webb, S., & Flower, M. A. (2016). *Webb's physics of medical imaging*. Place of publication not identified: CRC Press.
- Wilhelm Conrad Röntgen'in eşine ait ilk radyografi görüntüsü. (2020, 8 Ekim). Erişim Adresi: <https://www.winally.com/2018/09/yeni-nesil-goruntuleme/>
- Wu, X., Xu, K., & Hall, P. (2017). A survey of image synthesis and editing with generative adversarial networks. *Tsinghua Science and Technology*, 22(6), 660-674.
- Yapay Zekâ, Tıbbi Teşhise karşı. Teşhis otomatikleştirildiğinde ne olur? (2020, 15 Eylül). Erişim Adresi: <https://www.newyorker.com/magazine/2017/04/03/ai-versus-md>
- Zekâ teriminin tanımı. Türk Dil Kurumu. (10 Eylül, 2020). Erişim Adresi: <https://www.sozluk.gov.tr> (Erişim zamanı: 10.10.2020).

Sağlık Hizmetleri ve Yönetiminde Yapay Zekâ

Artificial Intelligence in Health Services and Management

Betül Akalın¹ , Ülkü Veranyurt² 



ÖZ

Günümüzde artan kronik hastalıklar, Covid-19 pandemisi gibi salgınlar, ortalama insan ömrünün uzamasına bağlı olarak artan yaşlı nüfusun evde bakım hizmetlerine olan ihtiyacının artışı ve sağlık okuryazarlığının gelişmesine paralel olarak bireylerin sağlık hizmetlerinden beklentilerindeki değişiklikler; sağlık hizmetleri ve yönetiminde de değişimi beraberinde getirmektedir. Sağlık hizmetleri ve yönetimi, sağlık sektörü ve diğer sektörlerin işbirliği ile geniş kitlelere hitap etmektedir. Sağlık profesyonelleri ile birlikte diğer meslek gruplarının bir arada koordineli bir şekilde çalışabilmesi gerekmektedir. Bunun yanında hastalıkların tanı, tedavi, rehabilitasyonunda ve toplum sağlığının geliştirilmesinde sağlık hizmetlerinin yönetiminde sağlık bilgi teknolojilerinin kullanımına ihtiyaç vardır. Tüm bunlar dikkate alındığında, artan iş yükü yanında yetersiz sayıdaki sağlık insan gücü sebebi ile sağlık hizmetleri ve yönetiminde yapay zekâ uygulamalarının kullanılması kaçınılmazdır. Sağlık alanında ve teknolojide güncel gelişmeler doğrultusunda hasta odaklı dijital bir sağlık ekosistemi yaratılmaya başlanmıştır. Umut vaat eden çözümleriyle yapay zekâ uygulamalarının kullanımında gerekli yasal düzenlemelerin yapılması önerilmektedir.

Anahtar kelimeler: Sağlık Hizmetleri, Sağlık Hizmetleri Yönetimi, Yapay Zekâ

ABSTRACT

Increasing chronic diseases and epidemics, such as the Covid-19 pandemic, shows a greater need in home care services of the elderly population due to the prolongation of the average human life span and changes in the expectations of individuals from health services in parallel with the development of health literacy. It also brings about change in health services and management. The cooperation of health services and management, the health sector and other sectors would benefit a wide audience. Health professionals and other occupational groups should be able to work in coordination. In addition, there is a need to use health information technologies in the diagnosis, treatment, rehabilitation of diseases and in the management of health services for the development of public health. Considering all of these, it is inevitable to use artificial intelligence applications in healthcare services and management due to the increasing workload and insufficient number of health workers. A patient-oriented digital health ecosystem is being created in line with current developments in healthcare and technology. It is recommended to make the necessary legal regulations in the use of artificial intelligence applications with promising solutions.

Keywords: Healthcare, Healthcare Management, Artificial Intelligence

¹ (Dr. Öğr. Üyesi), Sağlık Bilimleri Üniversitesi, Hamidiye Sağlık Bilimleri Fakültesi, Sağlık Yönetimi Bölümü, İstanbul, Türkiye

² Sağlık Bilimleri Üniversitesi, Hamidiye Sağlık Bilimleri Fakültesi, Sağlık Yönetimi Bölümü, İstanbul, Türkiye

ORCID: B.A. 0000-0003-0402-2461;
Ü.V. 0000-0003-4838-3373

Corresponding author:

Betül AKALIN
Sağlık Bilimleri Üniversitesi, Hamidiye Sağlık Bilimleri Fakültesi, Sağlık Yönetimi Bölümü, İstanbul, Türkiye
E-mail address: betul.akalin@sbu.edu.tr

Submitted: 30.12.2020
Revision Requested: 05.02.2021
Last Revision Received: 07.02.2021
Accepted: 18.03.2021
Published Online: 20.04.2021

Citation: Akalın, B. ve Veranyurt, U. (2021). Sağlık hizmetleri ve yönetiminde yapay zekâ. *Acta Infologica*, 5(1), 231-240. <https://doi.org/10.26650/acin.850857>

1. Giriş

Sağlık hizmetleri ve yönetimi, sağlık sektörü ile diğer sektörlerin işbirliği sayesinde geniş kitlelere hitap etmektedir. Sağlık profesyonelleri ile birlikte diğer meslek gruplarının koordineli bir şekilde çalışmasını gerektirmektedir. Bunun yanında hastalıkların tanı, tedavi, rehabilitasyonunda, toplumun sağlık düzeyinin geliştirilmesinde ve sağlık hizmetlerinin yönetiminde çeşitli teknolojik yöntemlere ve araçlara ihtiyaç vardır.¹ Günümüzde artan kronik hastalıklar, Covid-19 pandemisi gibi salgınlar, ortalama insan ömrünün uzamasına bağlı olarak artan yaşlı nüfusun evde bakım hizmetlerine ihtiyacının artışı ve sağlık okuryazarlığının gelişmesine paralel olarak bireylerin sağlık hizmetlerinden beklentilerindeki değişiklikler; sağlık hizmetleri ve yönetiminde de değişimi beraberinde getirmektedir. Tüm bunlar dikkate alındığında artan iş yükü yanında yetersiz sayıdaki insan gücü sebebi ile sağlık hizmetleri ve yönetiminde yapay zekâ uygulamalarının kullanılması kaçınılmazdır.²

Yapay zekâ, insan beyninin fonksiyonlarını temel alarak, insana özgü düşünme, yorum yapabilme, çözümlenme, karar verme gibi yetilerin bilgisayar, robot, program gibi sistemlere aktarılması olarak tanımlanabilir.³

Sağlık yönetiminde yapay zekâ uygulama alanları oldukça geniş kapsamlıdır. Dünyada pek çok ülkede yapay zekâ uygulamaları, ulusal ve bölgesel sağlık kurum/kuruluşlarının genel idari kararların tespit edilmesinde, alınmasında ve yürütülmesinde kullanılmaktadır.

Türkiye’de Sağlık Bakanlığı; Microsoft, Oracle gibi teknoloji firmalarının ürünlerini kullanmaktadır. Bakanlığın sahip olduğu yapay zekâ uygulamaları; MHRS (Merkezi Hekim Randevu Sistemi) kullanım oranları, Aile hekimliği performans raporları, hastaneye yatış, ameliyat, tanı vb. raporlar, eNabız değerlendirme raporlarıdır.⁴

Sağlık hizmetleri ve yönetiminde yapay zekâ uygulamaları gün geçtikçe popülerlik kazanmaktadır. Sağlık hizmetlerinde yapay zekâ uygulamalarının kullanılması için dünya genelinde pek çok bilimsel çalışma yapılmaktadır. Günümüzde geleneksel sağlık hizmetlerinde yapay zekâ uygulamaları kullanılması yönünde adımlar atılmaktadır.

Sağlık hizmetlerinde ve yönetiminde yapay zekâ uygulamaları, süreçlerde maliyetlerin azaltılması, kalitenin artırılması, insan kaynaklı hataların azaltılması ve performansın artırılması amacıyla kullanılmaktadır. Covid-19 pandemisinin başlamasıyla birlikte dünya sağlık sistemi öngörülemez bir talep ile karşılaşmıştır. Bu noktada hastanelerin kapasitelerini öngörmek güçleşmektedir. Yapılan bir çalışmada makine öğrenmesi tabanlı bir Covid-19 kapasite planlama ve analiz sistemi geliştirilmiştir (Capacity Planning and Analysis System (CPAS)). Bu sistem İngiltere’de belirli hastanelerde kurulmuş olup NHS (National Health Service) tarafından onaylı şekilde kullanılmaktadır.⁵ Sağlık yöneticilerinin ve sağlık profesyonellerinin iş süreçlerinde yapay zekâ uygulamalarını kullanabilmeleri için gerekli yasal düzenlemelerin yapılması önerilmektedir.

Bu çalışmada, ülkemizden ve dünyadan sağlık hizmetleri ve yönetiminde kullanılan/kullanılabilecek yapay zekâ uygulamalarına örnekler verilerek, sağlıkta yapay zekânın kullanılabileceğini belirtmek amaçlanmıştır.

2. SAĞLIK HİZMETLERİ SUNUMUNDA VE YÖNETİMİNDE YAPAY ZEKÂNIN ÖNEMİ

Sağlık hizmetlerinin sunumunda ve yönetiminde meydana gelen hatalar ve olumsuzlukların giderilmesinde, sağlık hizmetlerinde kalitenin geliştirilmesinde ve verimliliğin artırılmasında yapay zekâ uygulamalarının kullanılmasının gerekliliği ortaya çıkmıştır. Bu gereklilikler aşağıdaki başlıklarda verilmiştir⁶

2.1. Dünyada 65 Yaş ve Üstü Nüfustaki Artış

Dünyada yaşanan insan sayısının giderek artışı, artan yaşlı nüfusa yönelik uzun süreli tedavi ve bakım gerektiren sağlık hizmeti süreçlerinin de yönetimini gerektirmektedir. OECD verilerine göre 2017 yılında Finlandiya’da 65 yaş ve üstü nüfusun toplam nüfusa oranı 2000 yılına göre %6,1 artmıştır. 2017 yılında Türkiye’de 65 yaş ve üstü nüfusun toplam nüfusa oranı 2000 yılına göre %2,9 artmıştır.⁷ OECD verilerine göre hem ülkemizde hem de dünyada nüfusun yaşlandığı anlaşılmaktadır. Türkiye’de TÜİK 2013-2015 yılları arası Hayat Tabloları istatistiğine göre doğuştan beklenen yaşam süresi 78 yıl iken, TÜİK 2017-2019 yılları arası Hayat Tabloları istatistiğine göre doğuştan beklenen yaşam süresi 78,6 yıla yükseldiği belirtilmektedir.⁸ Önümüzdeki 30 yıl içerisinde dünyada 65 yaş üstü nüfusun 1.6 milyar olacağı tahmin edilmektedir.⁹

2017 OECD verilerine göre Türkiye’de 100 bin kişiye 187 doktor düşmektedir.¹⁰ 2017 OECD verilerine göre 1,4 milyarlık nüfusa sahip olan Hindistan’da 100 bin kişiye 78 doktor düşmektedir.⁷ Artan yaşlı nüfusa karşılık sağlık personeli istihdamı oldukça zordur. Bir doktorun ya da sağlık profesyonelinin istihdam edilebilmesi için hem insan kaynağı, eğitim için zaman ve gerekli çalışma ortamının oluşturulması gerekmektedir. Sınırlı kaynaklarla sağlık hizmeti talebinin karşılanabilmesi için akılcı teknolojiler kullanılması artık bir gerekliliktir.¹¹

1.2 Sağlık Tehdit Eden Hastalıklar

DSÖ’nün 2019 yılında açıkladığı raporda küresel anlamda insan sağlığını tehdit eden 10 büyük tehlikeye dikkat çekilmiştir.¹² Bu raporda sanayinin ve teknolojinin gelişmesiyle artan hava kirliliği ve salgın hastalıklara neden olan virüsler, yetersiz sağlık hizmeti sebebiyle hastalık yaparak topluma yayılmasına neden olmaktadır. Bunun yanında sosyal, ekonomik ve kültürel nedenlerden dolayı temiz içme suyuna, besine ve sağlıklı yaşam ortamlarına sahip olunmamasından kaynaklı sağlık sorunları artmaktadır. Günümüzde yaşanan Covid-19 pandemisi de bugün tüm dünyada mücadele edilen önemli bir halk sağlığı sorunudur. Aynı raporda kalp hastalıkları, kanser, diyabet gibi bulaşıcı olmayan kronik hastalıkların dünyadaki ölümlerin %70’inden fazlasına sebep olduğu belirtilmektedir.¹² Kronik hastalıkların uzun süreli tedavi süreçlerini gerektirmesi sebebi ile hem maddi kaynaklar daha fazla harcanmakta hem de daha fazla sayıda iş gücü kaynağına ihtiyaç duyulmaktadır.

1.3. Düşük Kalite Yaşam Standartı

Sağlığı tehdit eden bulaşıcı olan ya da olmayan hastalıklara insanların maruz kalmasının en önemli sebeplerinden biri de düşük kalite yaşam standartıdır. Özellikle, tüm dünyada 30-69 yaş arası ölümlerin nedeni tütün ve alkol ürünlerinin tüketiminin artışıdır. Ayrıca, sağlıksız beslenme, fiziksel hareketsizlik, hava kirliliği de erken ölümlere neden olabilmektedir.¹² Sosyo-ekonomik yönden yetersiz olan ülkelerde alım gücü düşük olduğundan ya da savaşlar, çatışmalar sebebi ile besin maddelerine, içme suyuna ve hijyen ürünlerine ulaşım oldukça sınırlıdır. Sonuç olarak, artan sağlık hizmeti talebine karşılık verilmesi ülkelerin sağlık sistemlerini zorlamaktadır.

1.4. Maliyetlerin Artması ve Rekabet

Dünya genelinde 65 yaş üstü bireylerin yaşam sürelerinin artması beraberinde uzun süreli tedavi süreçleri ve yüksek maliyete neden olmaktadır. Her ülkenin, mevcut sağlık politikalarının uygulanması küresel dünyada oldukça güç hale gelmiştir. Bunun başlıca sebeplerinden biri aniden çıkan salgın hastalıklar çok sayıda bireyin hastalanmasına ve sağlık hizmeti talebi yaratmasına neden olmaktadır. Sınırlı sayıda istihdam ve kabiliyeti olan sağlık kurumları talebi karşılamakta yetersiz kalabilmektedir. Özellikle aşı, ilaç gibi tedavi ve korumaya yönelik çalışmalar oldukça maliyetli ve uzun yıllar gerektirmektedir.^{13, 14, 15}

Sağlık hizmeti sunun devlet kurumu ya da özel kurumlar belli bir hizmet kalitesinin üzerinde sağlık hizmeti sunmak zorundadır. Sağlık işletmesinin varlığını devam ettirebilmesi için pazardaki rakiplerinden farklılıkları, üstünlükleri olmalıdır. Sağlık kuruluşunun kapasitesi düşünüldüğünde insan üstü çalışmalar doktor ve sağlık profesyonellerinden beklenilmektedir. Böyle bir durumda artan hasta sayısına karşılık kısıtlı insan gücü yanında yapay zekâ teknolojilerinden yararlanılması kaçınılmazdır.¹

2.5. Bilişim Teknolojilerindeki Gelişmeler

İnternetin gelişmesi ile birlikte pek çok sektörde olduğu gibi sağlık sektöründe de dijital bir dönüşüm gerçekleşmektedir. Bu süreçte yapay zekâ uygulamaları bir çok alanda olduğu gibi sağlık hizmetlerinde ve yönetiminde de meydana gelen problemlerin çözülmesinde, hastalıkların teşhis ve tedavisinde, hastalıkların tahmin edilmesi ve daha bir çok konuda kullanılmaya başlamıştır. Sağlıkta bilişim teknolojilerinin en önemli uygulama alanlarından biri de yapay zekâ uygulamalarıdır.⁴

2. Sağlık Hizmetleri ve Yönetiminde Yapay Zekânın Uygulama Alanları

3.1 Yapay Zekânın Yönetim Amaçlı Uygulama Alanları

Genel sağlık yönetimi: MHRS kullanım oranları, eNabız verileri vb. gibi veriler kullanılarak hastaneye başvurular tahmin edilebilir, acil servislerde bekleme süreleri azaltılabilir, ambulans ulaşım süreleri standardize edilebilir, bebek/çocukların aşılama takibi takip edilebilir.

Dokümantasyon yönetimi: Elektronik Sağlık Kayıt Sistemleri ile sağlık profesyonellerin bilgilerinin düzenlenmesi, saklanması ve yeniden kullanılması sağlanabilir. NLP (Doğal Dil İşleme) araçlarıyla doktor ve sağlık profesyonellerin sesli kayıtları ile raporlar yapay zekâ tabanlı uygulamalarla çok kısa süre yazdırılabilir.

Maliyet ve kalite yönetimi: Sağlık kuruluşları diğer işletmelerin aksine, sağlık hizmeti sunarken hastayı müşteri olarak görmemelidir. Sağlık kuruluşunun varlığını idame ettirebilmesi için katlanması gereken maliyeti azaltıp, sunduğu hizmetin kalitesini arttırması gerekmektedir. Centerstone Araştırma Enstitüsü yaptığı bir araştırmada yapay zekâ kullanarak tanı koymanın, geleneksel tanı koymaya göre daha az maliyetli olduğunu göstermiştir. Bu çalışmada, rastgele seçilmiş 500 hasta bireyin fiziksel performansları, hastalıkları ile ilgili sonuçları yapay zekânın algoritmasına göre karar verme modelleriyle kıyaslanmıştır. Sonuç olarak, birim başına maliyetler arasında anlamlı bir farklılık olduğu ortaya çıkmıştır. Yapay zekâ algoritmaları 189 dolar maliyet çıkarırken, geleneksel yöntemlerin maliyeti 497 doları bulduğu görülmüştür.¹⁶ Sağlık işletmesinin maruz kalabileceği risk ve maliyetleri önceden tespit edebilen yapay zekâ uygulamaları kullanılmaktadır.

Sağlık kurumlarının kapasitesinin akılcı kullanımı: Boş yatak kapasitesinin anlık olarak kontrol edilmesinde kullanılabilirdiği gibi taburcu edilen hastaların tekrar yatış yapma olasılığı da hesaplanabilmektedir. Özellikle, son yıllarda acil servislerde hasta sayısı ve bu servislere talebin artması yoğunluğun önceden hesaplanabilmesi için bir gereksinim doğurmuştur. Yapılan bir çalışmada, Long Short-Term Memory (LSTM) derin öğrenme modeli kullanılarak zamana bağlı acil servis yoğunluğunu hesaplayan, ilerleyen gün ve aylar için günlük hasta sayılarını önceden hesaplayan bir uygulama yapılmıştır.¹⁷ Covid-19 gibi salgınlarda sınırlı sayıda yatağın efektif bir şekilde kullanılması açısından oldukça faydalı kullanım alanı olabilir.

Sağlık hizmeti ve yönetimindeki olumsuzlukların iyileştirilmesi: Hasta kayıtlarının arşivlenmesinde yaşanan sıkıntılar ve siber saldırılar gibi sebeplerden meydana gelen olumsuzlukların giderilmesinde kullanılabilir. Her geçen saniye dünyanın bir ucundan sağlık ile ilgili veriler insanların giydikleri cihazlar, kullandıkları uygulamalar, elektronik sağlık kayıtları olarak üretilmektedir. Sağlık verilerinin oluşturduğu büyük verileri analiz edebilmek, arşivlemek, gerektiğinde kullanabilmek maliyetli ve insan gücü gerektirmektedir. ABD’de bulunan Montefiore Health System, Intel ile işbirliği yaparak çeşitli hasta popülasyonuna daha etkili bir hizmet sunabilmek üzere çok büyük miktardaki hasta verilerindeki ortak modelleri görmek için yapay zekâ çözümlerini ve analizlerini devreye almıştır.¹⁸ Bu şekilde, insan kaynaklı hataların azaltılması, ilaç dozu hesaplamalarındaki yanlışlıklardan kaynaklanan ilaç israfı ve hastaya uygulanan yanlış tıbbi tedavi sonucu oluşan zararların engellenmesini mümkün olabilmektedir.

Uzaktan önleyici ve tamamlayıcı sağlık hizmeti sunumu: Sağlık hizmeti sunumu yanında bireyin sağlığını koruması, hastalanmasının önlenmesi oldukça önemlidir. Sağlığın geliştirilmesinde ve takibinde hastaneye gitmeden uzaktan takip edebilen sistemler sayesinde sınırlı kaynaklarda %60 oranında tasarruf sağlanabilir. Dijital sağlık ortamında karar destek sistemleri ve makine öğrenmesi önemli rol oynamaktadır. Yapılan bir çalışmada, çocuklarda obezitenin engellenmesi için makine öğrenmesi tabanlı bir karar destek sistemi önerilmiştir. Yapılan incelemede makine öğrenmesi tabanlı karar destek sisteminin yapay sinir ağları ile kullanılması durumunda obezite riski olan çocukları öngörebildiği tespit edilmiştir.¹

3.2 Yapay Zekânın Klinik Amaçlı Uygulama Alanları

Yapay zekânın klinik amaçlı kullanılması ile ilgili dünya genelinde pek çok çalışma mevcuttur. Çok geniş kapsamlı olmasına rağmen aşağıdaki başlıklar ile kategorize edebiliriz.⁴

Halk sağlığı yönetimi: yapay zekâ uygulamaları, bireylerin, hastalıklar hakkında bilinçlendirilmeleri, çok sayıda hedef kitleye ulaşarak tarama testleri yapılması ve değerlendirilmesi gibi konularda kullanılabilir. Covid-19 pandemisi nedeni ile kişilerin kendilerini riskli görülen bölgelerden koruması, hastaneye gitmeden mevcut sağlık durumunun değerlendirilmesi ile günümüzde yapay zekâ uygulamaları kullanılmaktadır.

Covid-19 pandemisi sürecinde FİTAS (Filyasyon ve izolasyon takip sistemi) ile filyasyon yapılan alanda veriler anlık olarak sisteme kaydedilerek, hızlıca analiz edilmesinde kullanılmaktadır. Ülkemizde ilk Covid-19 vakasından sonra tespit edilen vakalar için filyasyon çalışmaları yapılarak, Covid-19 pandemisi yayılım haritası çıkarılmaktadır. FİTAS ile elde edilen veriler doğrultusunda alınan tedbirle birlikte vaka artış hızının düşürülmesi amaçlanmaktadır.²⁰

Erken tanı ve acil müdahale: Özellikle kanser hastalıklarının erken teşhisi ve tedavisi son derece önemlidir. Hastalık gözle görülür belirtiler vermesini beklemeden hastalığın başlangıç evresinde teşhis edilmesi hayat kurtarıcıdır. Çeşitli makine öğrenmesi algoritmaları hastalıkların önceden tahmin edilmesinde kullanılmaktadır. Parkinson, erken tanının çok önemli olduğu bir hastalıktır. Yapılan bir çalışmada EEG (elektroensefalogram) sinyalleri, fotoğraf simülasyonları, PDC (partial directed coherence) verilerini makine öğrenmesi teknikleri kullanılarak hasta kişiler, ilaç kullanması gerekenler ve sağlıklı bireyler olarak sınıflandırılmıştır. Çalışmada testlerde %99 oranında doğruluk gözlemlenmiştir.²¹

Dünyada yaygın olarak bilinen yapay zekâ sağlık hizmetleri destek sistemlerinden biri de IBM Watson'dur. Makine öğrenimi ve doğal dil işleme yetenekleri ile bu sistem, doktorların hastaların elektronik sağlık kayıtlarının incelemelerine ve aramayla ilgili tıbbi araştırma yayınlarını ve kılavuzlarını incelemelerine yardımcı olmak için tasarlanmıştır. Çift kör bir çalışmada, bir tümör kurulu tarafından alınan kararlarla Watson Onkoloji Sistemi (Watson for Oncology) tarafından alınan kararları karşılaştırmıştır. Sonuçlar, sistem tarafından yapılan önerilerin %90'ının tümör kurulu tarafından yapılan önerilere uygun olduğunu göstermiştir.²²

Radyoloji görüntülerinin değerlendirilmesi: Radyoloji görüntülerinin incelenmesi, raporlanması, raporların hastanın doktoru tarafından değerlendirilmesi, uzun bir süreçtir. Yapay zekâ uygulamaları kullanılarak, radyoloji görüntülerinde çeşitli nodüllere/lezyonlara rastlananlar hastaların doktorlarına görüntülerin gönderilmesi hem zamandan hem de iş yükünden tasarruf sağlayabilmektedir. Zamanında teşhis konulması için hastaya kaliteli bir sağlık hizmeti de böylece sunulabilmektedir. Yapay zekânın radyoloji alanında kullanımı ile ilgili yapılan bir çalışmada kemik yaşını değerlendirmek için Evrişimsel Sinir Ağları (CNN) kullanılmıştır. Çalışma sonucunda CNN ile pediatrik el radyografilerin yaş değerlendirmeleri, radyologların değerlendirmelerine benzer doğrulukta çıkmıştır.²³

Test sonuçlarının takibi: Bir çok hastalığın tekrardan ortaya çıkması ya da başka bir hastalığa sebep olması gibi nedenlerden ötürü hastalığın takibi son derece önemlidir. Hastaya yapay zekâ tabanlı mobil cihazlar giydirilerek, dünyanın neresinde olursa olsun hasta, doktoru tarafından takip edilebilmektedir. Hastanın yaptırdığı test sonuçları hastanın izni ile sisteme kayıt edildiği için doktor yer/zaman fark etmeksizin hastasının verilerine ulaşabilmektedir.

Tedavi: Olası tedavi protokollerinin oluşturulmasında ve hangisi kullanıldığında hastaya ne gibi yarar/zarar sağlayabileceğini yapay zekâ tabanlı sistemlerle tahmin etme yaparak doktora destek olunmaktadır. Hindistan'da bir Kanser Merkezi'nde onkologlar tarafından meme kanseri hastaları için uygulanan tedaviler; geriye dönük gözlemsel bir çalışmada "Watson for Oncology" adındaki yapay zekâ sistemi kullanılarak karşılaştırılmıştır. Sistem, genel olarak kanser evresi ve yaşına göre %93'lük (%80-%97) uyum göstermiştir.²⁴ Yapılan başka bir çalışmada; birden fazla hastaneden doğumsal kataraktlarla ilgili veriler elde edilmiştir. Yapay zekâ tabanlı kullanılan sistem ile çerçevenin tanımlama ağlarında % 98,25 doğruluk ve tedavi önerileri için % 92,86 doğruluk gösterdiği belirlenmiştir.²⁵ Yapay zekâ, oftalmolojide mükemmel bakım sağlama amacına hizmet edebilecek çeşitli kullanımlar göstermiştir.²⁶

Yapay zekâ destekli robotik ameliyat: Günümüzde bazı hastanelerde kullanılmaya başlamıştır. Yer/zaman farkı olmaksızın anlık dünyanın herhangi bir ucundan doktor destekli yapay zekâ destekli ameliyatlar yapılmaktadır. Robotlar yardımcı cerrah olarak ameliyatta da kullanılmaktadırlar. Da Vinci Cerrahi Sistemi (The da Vinci Surgical System), en çok tercih edilen robotik cerrahi sistemlerden biridir.²⁷ Bunun yanında otonom ameliyat yapabilen yapay zekâ destekli robotlar için çalışmalar devam etmektedir.

Patoloji sonuçlarının değerlendirilmesi: Çok nadir görülen hücre görüntülerinin varlığı, boyamadan kaynaklı hatalar, patoloğun yorgun olmasından kaynaklı hataların azaltılması amacıyla kullanılmaktadır. Sepsis, ölüm oranı yüksek bir hastalık olup hastalarda ilgili patojenin tespiti zor bir süreçtir. Yapılan bir çalışmada klinik-metabolik sepsis vakalarını içeren bir veri tabanı kullanılarak farklı makine öğrenmesi teknikleri ile bir sınıflandırıcı model oluşturulmuştur. Veri tabanında 100 hastaya ait veri kullanılarak 29 klinik ve metabolik özellikler paneline bakılmıştır. Elde edilen en iyi AUC değeri (area under curve) 0,94'tür. Çalışma sonucuna göre seçilmiş panel ile, sepsisli hastaları ayırt etmek için güçlü biyobelirteçler olabileceği ortaya konulmuştur.²⁸

İlaç tedavisi takibi: Hasta bireylerin ilaçlarını doğru dozda, doğru zamanda, doğru yöntemle alması önemlidir. Yapay zekâ uygulamaları, Alzheimer gibi kronik hastalıklarda ilaçların takip edilmesi için kullanılmaktadır. Yüz tanıma teknolojisi ile yapay zekâ tabanlı sistemler kullanılarak, hastanın yüzü tanınmakta, hastanın almak istediği ilacın doğru olup olmadığı tespit edilebilmektedir.

Gelişen teknoloji Parkinson hastalığının tedavisinin uzaktan takibi için yeni yöntemler sunmaktadır. Yapılan bir çalışmada hastaların motor sinir davranışları geliştirilen bir cihazda tuşlara basış şekil ve sürelerine bakılarak 6 ay süre ile gözlemlenmiştir. Elde edilen veriler üzerinde geliştirilen yeni bir derin öğrenme algoritması ile hasta durumları sınıflandırılmıştır. Bu çalışma kullanılan gözetimsiz motor sinir verilerinin, ilaçlara verilen tepkinin takibi ve gözetiminde kullanılmasına bir örnektir.²⁹

İlaç geliştirme: İlaç geliştirme pahalı, zaman alan ve yoğun emek isteyen süreçleri kapsamaktadır. Yapay zekâ teknolojileri kullanılarak moleküler düzeyde ilaç geliştirme çalışmalarına katkı sağlanmaktadır. Örneğin, pek çok ilaç veya besin takviyesi gibi maddeler insan vücudunda hücre bileşenlerine oksidatif zararlar verebilmektedir. Bu tip maddelerin verebileceği zararların ön görülmesinde hidrojen atom transferi kullanılmakta ve sonuçların değerlendirilmesi için kuantum bilgisayarlar gerekmektedir. Bu etkinin simüle edilip, tahmin etmesi için makine öğrenmesi kullanılarak bir model önerilmiştir. Böylece, bu hesaplamaların yapılabilmesi için donanım ve zaman maliyeti ciddi ölçüde optimize edilmiştir.³⁰

Kişiyeye özel tedavi: Günümüzde bazı hastalıkları tedavi etmede geleneksel tedavi yöntemleri yetersiz kalmaktadır. Genetik yapı, bağışıklık sistemi, yaşam tarzı gibi farklılıklar sebebiyle kişiyeye özel tedavi yöntemleri geliştirilmesi gerekmektedir. Kişiyeye özgü tedavi yöntemleri geliştirilmesinde, yapay zekâ uygulamalarından yararlanılmaktadır. Kore’de yapılan bir çalışmada Technology Acceptance Model (TAM) kullanılarak hastaların davranışları nelerden zevk aldıkları, birlikte yaşadığı kişi sayısı ve karakterleri gibi parametreleri göz önünde bulundurarak kişiyeye özel evde bakım hizmeti sağlayan bir robot önerilmektedir. Bu çalışmada 403 hasta ve hastanın ailesine ait bilgiler gözlemlenmiştir. Robotlar ile birlikte geçirdikleri süre değerlendirilip, sağlıklı yaşam konusundaki katma değeri incelenmiştir. Özellikle robotun kişiyeye özel davranması ve kullanım kolaylığı kullanıcıların en beğendiği faktörler olmuştur.³¹

Tedavi sonrası klinik karar destek: Klinik karar destek sistemleri doktorlara tedavi sonrası komplikasyonlar ile uygulanan tedavinin devam ettirilip/ettirilemeyeceği konusunda destek olmaktadır.

Hastaya evde bakım: Bilindiği üzere ortalama yaşamdan beklenen yaşam süresi artmıştır. Buna bağlı kronik hastalıkları, Alzheimer ve bunama gibi rahatsızlığı olan birey sayısı gün geçtikçe artmaktadır. Bunun yanında evde yalnız başına yaşayan yaşlı birey sayısı azımsanmayacak sayıdadır. Örneğin evde bakım hizmetleri ile ilgili Finlandiya’da 2018 yılında 68,400 hasta için sağlık kontrolleri evden sanal olarak gerçekleştirilmiştir. 250 hasta için belirli testler 60 adet medikal robot kullanılarak yapılmıştır.³² Yapay zekâ tabanlı robotlar ile hastaların evlerinde temel sağlık verilerinin, ilaçlarının takibi, gerekli durumlarda acil sağlık ekiplerinin gelmesini sağlamaktadır.

4.Sağlık Hizmetleri Ve Yönetiminde Yapay Zekâ Uygulamalarının Avantajları/Dezavantajları

1.1 Sağlık Hizmetleri ve Yönetiminde Yapay Zekâ Uygulamalarının Avantajları

Sağlık hizmetleri ve yönetiminde yapay zekâ uygulamaları ile ilgili çalışmalarda, yapılan işlerin yeniden tanımlandığı, kalite ve verimliliğin arttığı, maliyetlerin azaldığı, insan/sistem kaynaklı hataların azaltıldığı belirtilmiştir. Yapay zekâ uygulamalarının sağladığı bazı avantajlar aşağıdaki gibi sıralanabilir: ^{33, 34}

- Doktorlardan gelen girdiler koordineli olarak geliştirilirse, girdiler ve çıktılar uygun şekilde tanımlanabilirse, yapay zekâ sayesinde hastalık karakterizasyonu, hastalık tespiti ile raporlamada bir standardizasyon sağlanabilir.
- Sürekli çalışmalar yapıldığı için problemlerin çözüm önerileri geliştirilmeye çalışılmaktadır. Birden fazla algoritma birleştirilerek ortaya daha iyi bir algoritma çıkarılarak (tümeleştirmeye yöntemi) yeni çözüm önerileri sunulabilmektedir.
- Nitel ve nicel veri gruplarında başarı sağlar. Örneğin, tahlil sonuçlarının değerlendirilmesi, kanser tanısının konmasının tahmin edilmesinde kullanılabilir.

- Farklı klinik alanlardan (radyoloji, biyokimya, mikrobiyoloji, patoloji vb. gibi) veriler kullanılarak çeşitli bilgiler elde edilebilir. Bu elde edilen bilgilerin yapay zekâ teknolojileri kullanılarak anlamlı bilgilere dönüştürülerek çeşitli hastalıkların teşhisinde kullanılabilir.
- Zaman tasarrufu sağlar.
- Emek isteyen, maliyetli ve zaman alan işleri otomatikleştirerek işleri daha yapılabilir hale getirir.
- Verimliliği artırarak, görüntülerin daha hızlı ve güvenilir raporlanmasını sağlar.
- Arşivlemede fiziksel alan sıkıntısı yaratmadan depolamayı sağlar.
- Manuel ve beceri isteyen birçok işi otomatize ederek standardize edilmesini sağlar.
- İş yükünü azaltır.
- Göz taraması yapay zekâ teknolojileri kullanılarak yapıldığında, doktora başvuru yapacak hastalar önceden filtrelenir. Bu şekilde, hasta olmayanların sağlık kuruluşlarına başvurmaları önlenirken, hastalık tespit edilen kişilerin daha hızlı doktora yönlendirilmesine yardımcı olur. Dolaylı olarak, sağlık maliyetlerinin hem hesaplanmasında hem de maliyetlerin azaltılmasında önemli rol oynar.

4.2 Sağlık Hizmetleri ve Yönetiminde Yapay Zekâ Uygulamalarının Dezavantajları

Yapay zekâ uygulamalarının dezavantajlarından bazıları aşağıdaki gibidir: ^{33,34}

- Algoritmayı geliştiren mühendislerin yeterli tıbbi bilgiye sahip olmaması, radyoloğun iş akışının anlaşılmasına bağlı oluşabilecek hatalar. Bununla birlikte, algoritmayı kullanacak radyoloğun konuya hâkim olmaması yanlış kullanım ve yorumlamaya bağlı hatalara sebep olabilir. Bu yüzden multidisipliner bir ekip çalışması gereklidir. Radyolog ve algoritmayı geliştiren mühendis iş birliği içinde olmalıdır. Görüntülerin eğitimi sırasında mutlaka radyolog görüntüleri incelemelidir.
- Görüntü analizi için doğru ve yeterli sayıda veri seti gereklidir. Veri setlerinin oluşması için belli bir zaman gerekmektedir ayrıca, veri setlerinin doğruluklarının kontrol edilmesi süreçleri de oldukça zahmetlidir.
- Kullanılan veri seti çok önemlidir. İyi eğitim gerçekleşmezse güvenilir olmayan sonuçlar verebilir.
- Veri seti yeterli sayıda örnek içermiyorsa yanıltıcı sonuçlar verebilir.
- Sadece görüntülerin kullanılması yeterli sonuç vermeyebilir. Tanı için doğru özellikler hekim ile birlikte karar verilmesi gerekir.
- Klinik ilaç tedavisi uygulamalarında henüz çok somut örnekleri yoktur.
- Sorumluluk ve etik açısından risk teşkil etmektedir.
- Çok sayıda görüntünün arşivlenebilmesi için dijital görüntü haline getirilmesi gerektiğinden, büyük bir iş yüküne de sebep olmaktadır.
- Teknolojinin ilerlemesiyle birlikte yapay zekâ tabanlı daha güçlü modeller geliştirilmelidir.
- Sistemin kurulumu için gerekli cihazların ve malzemelerin yetersizliği ve mali yetersizlikler için gerekli önlemler alınmalıdır.
- Ahlaki ve etik boyutu ile ilgili birçok soru cevapsız kaldığı için belirsizlik mevcuttur.

5. Tartışma

Yapay zekâ uygulamaları uzay, askeri, pazarlama, turizm vb. gibi pek çok sektörde kullanım alanı bulmaktadır. Günümüzde sağlık hizmetleri ve yönetiminde de yapay zekâ uygulamaları kullanılarak erken tanı, tedavi, rehabilitasyon, sağlığın geliştirilmesi vb. gibi süreçlerde uygulama alanları mevcuttur. Sınırlı kaynaklarla, geniş bir kitleye kaliteli ve düşük maliyetli, etkili, zamanında ve coğrafi bölge ayrımı yapmaksızın sağlık hizmeti sunmak oldukça zordur. Artan sağlık hizmeti ihtiyacı karşısında sınırlı sağlık insan gücü nedeniyle yapay zekâ uygulamalarının kullanılması kaçınılmazdır. Bu doğrultuda bireylerin yaşam kalitelerinin artırılması, sağlık maliyetlerinin azaltılması, yerinde ve zamanında ve yeterli sağlık hizmeti sunumu açısından süreçlerin optimize edilmesi ve süreçlere yapay zekâ teknolojilerinin entegre edilmesi artık bir gerekliliktir.

Küresel boyutta sağlık sektörü diğer sektörlerde olduğu gibi dünyadaki teknolojik gelişmeleri takip etmektedir. Bu sektörün paydaşları olan sigorta şirketleri, ilaç ve medikal cihaz üreten firmalar, hizmet sunumunda yer alan kurum/ kuruluşlar sağlık sektörünün büyümesi ve geliştirilmesi için büyük önem arz etmektedir.³⁵ Ülkelerin uyguladığı sağlık politikalarının temelini maliyeti düşük, kaliteli ve erişilebilir sağlık hizmeti başlıkları oluşturmaktadır. Sağlık sektöründe pazardan pay almak ve sektördeki varlığının devam ettirmek isteyen paydaşlar artan nüfus, kronik hastalık ve salgınlar, yaşam standartlarındaki değişimler, sağlık okuryazarlığının artışı ve kriz ortamlarında sürdürülebilir sağlık hizmeti için yeni yollar aramaktadırlar. Geleneksel yöntemlerin sebep olduğu dezavantajları avantaja çevirecek çözüm önerilerinin başında yapay zekâ uygulamaları gelmektedir.

Günümüzde sağlık alanında ve teknolojiye meydana gelen gelişmeler doğrultusunda hasta odaklı dijital bir sağlık ekosistemi yaratılmaya başlanmıştır. Umut vaat eden çözümleriyle yapay zekâ uygulamalarının kullanımında gerekli yasal düzenlemelerin yapılması önerilmektedir.

Hakem Değerlendirmesi: Dış bağımsız.

Çıkar Çatışması: Yazarlar çıkar çatışması bildirmemiştir.

Finansal Destek: Yazarlar bu çalışma için finansal destek almadığını beyan etmiştir.

Yazar Katkıları: Çalışma Konsepti/Tasarım- B.A., Ü.V.; Veri Toplama- B.A., Ü.V.; Veri Analizi/Yorumlama- B.A., Ü.V.; Yazı Taslağı- B.A., Ü.V.; İçeriğin Eleştirel İncelemesi- B.A., Ü.V.; Son Onay ve Sorumluluk- B.A., Ü.V.

Peer-review: Externally peer-reviewed.

Conflict of Interest: The authors have no conflict of interest to declare.

Grant Support: The authors declared that this study has received no financial support.

Author Contributions: Conception/Design of Study- B.A., Ü.V.; Data Acquisition- B.A., Ü.V.; Data Analysis/Interpretation- B.A., Ü.V.; Drafting Manuscript- B.A., Ü.V.; Critical Revision of Manuscript- B.A., Ü.V.; Final Approval and Accountability- B.A., Ü.V.

Kaynaklar/References

- Veranyurt U, Deveci AF, Esen MF, Veranyurt O. (2020). Disease Classification By Machine Learning Techniques: Random Forest, K-Nearest Neighbor and Adaboost Algorithms Applications. *Usaysad Derg.* 2020; 6(2):275-286.
- Hayran O. (2012). Sağlık Yönetimi Yazıları. Ankara: Sage Yayıncılık.
- Elmas Ç. 2018. Yapay Zekâ Uygulamaları. 4. Baskı. Ankara: Seçkin Yayıncılık, 25-45.
- Akalın B. 2020. Sağlık Hizmetleri ve Yönetiminde Yapay Zekâ. İstanbul: Hiper Yayıncılık, 125-132.
- Qian, Z., Alaa, A. M., & van der Schaar, M. (2020). CPAS: the UK's National Machine Learning-based Hospital Capacity Planning System for COVID-19.
- Jiang F, Jiang Y, Zhi H, et al. (2017). Artificial intelligence in healthcare: past, present and future. *Stroke and Vascular Neurology*;2: e000101. doi:10.1136/svn-2017-000101.
- OECD. Organisation for Economic Co-operation and Development. (2019). Health at a Glance 2019: OECD Indicators. Paris: OECD Publishing; 2019. doi:10.1787/4dd50c09-en.
- TÜİK. Türkiye İstatistik Kurumu. <https://www.tuik.gov.tr/> (Erişim Tarihi: 01.11.2020).
- Öcal H., Doğru İ. A. ve Barışçı N. (2019). Akıllı ve geleneksel giyilebilir sağlık cihazlarında nesnelerin interneti. *Politeknik Dergisi*, 22(3): 695-714.
- T.C. Sağlık Bakanlığı Sağlık İstatistikleri Yıllığı 2018. <https://dosyasb.saglik.gov.tr/Eklenti/36134,siy2018trpdf.pdf?0> (Erişim Tarihi: 01.11.2020).
- Çoban, H. (2009). Sağlık Ekonomisi ve Türkiye'de Sağlık Hizmetlerinin Yeniden Yapılandırılması. Dokuz Eylül Üniversitesi, Doktora Tezi, İzmir.
- World Health Organization (WHO). "Ten threats to global health in 2019", <https://www.who.int/emergencies/ten-threats-to-global-health-in-2019> (Erişim Tarihi: 01.11.2020).

- Fleming, N. (2018). How Artificial intelligence is changing drug discovery. *Nature*, 557(7707), ss 55-57. Doi: 10.1038/d41586-018-05267-x.
- Hecht, D. (2011). Applications of Machine Learning and Computational Intelligence to Drug Discovery and Development. *Drug Development Research*, 72: 53–65.
- Gavin, B., Hayden, J., Adamis, D., & McNicholas, F. (2020). Caring for the psychological well-being of healthcare professionals in the Covid-19 pandemic crisis. *Ir Med J*, 113(4), 51.
- Mesko, B. (2017). Yapay Zekâyla Tıbbi Karar Almak. B. Mesko içinde, *Tıbbın Geleceğine Yolculuk* (s. 174-183). İstanbul: Optimist Yayın Grubu.
- Kadri, F., Baraoui, M., & Nouaouri, I. (2019, September). An LSTM-based Deep Learning Approach with Application to Predicting Hospital Emergency Department Admissions. In 2019 International Conference on Industrial Engineering and Systems Management (IESM) IEEE. (1-6).
- İntel. (2018). Sağlık Dönüşümü. <https://www.intel.com.tr/content/www/tr/healthcare-it/healthcare-overview.html> (Erişim Tarihi: 01.11.2020).
- Triantafyllidis, A., Polychronidou, E., Alexiadis, A., Rocha, C. L., Oliveira, D. N., da Silva, A. S., Freire, A. L., Macedo, C., Sousa, I. F., Werbet, E., Lillo, E. A., Luengo, H. G., Ellacuria, M. T., Votis, K., & Tzovaras, D. (2020). Computerized decision support and machine learning applications for the prevention and treatment of childhood obesity: A systematic review of the literature. *Artificial Intelligence In Medicine*, 104.
- T.C. Sağlık Bakanlığı Sağlık Bilgi Sistemleri Genel Müdürlüğü, FİTAS (Filyasyon ve İzolasyon Takip Sistemi). <https://sbsgm.saglik.gov.tr/TR,73584/fitas.html> (Erişim Tarihi: 01.11.2020).
- Yan, Y. (2020). MCI progression classification for early diagnosis of Alzheimer's disease using machine learning and deep learning methods. *Res. Biomed.* 36:311–331.
- Curioni-Fontecedro, A. (2017). A new era of oncology through artificial intelligence. *ESMO Open*, 2(2): e000198.
- Larson, D.B., Chen, M.C., Lungren, M.P., Halabi, S.S., Stence, N.V., Langlotz, C.P. 2018. Performance of a deep-learning neural network model in assessing skeletal maturity on pediatric hand radiographs. *Radiology* 287, 313–322.
- Somashekhar, S.P., Sepulveda, M.J., Puglielli, S., Norden, A.D., Shortliffe, E.H., Rohit Kumar, C., Rauthan, A., Arun Kumar, N., Patil, P., Rhee, K., Ramya, Y. 2018. Watson for Oncology and breast cancer treatment recommendations: agreement with an expert multidisciplinary tumor board. *Ann. Oncol.* 29, 418–423.
- Long E, Lin H, Liu Z, et al. 2017. An artificial intelligence platform for the multihospital collaborative management of congenital cataracts. *Nat Biomed Eng*;1. Article number 24.
- Ting, D. S. W., Pasquale, L. R., Peng, L., Campbell, J. P., Lee, A. Y., Raman, R., ... & Wong, T. Y. (2019). Artificial intelligence and deep learning in ophthalmology. *British Journal of Ophthalmology*, 103(2), 167-175.
- Guo J, Li B (2018) The application of medical artificial intelligence technology in rural areas of developing countries, *Health Equity* 2:1, 174–181.
- Zheng, L., Lin, F., Zhu, C., Liu, G., Wu, X., Wu, Z., Zheng, J., Xia, H., Cai, Y., & Liang, H. (2020). Machine Learning Algorithms Identify Pathogen-Specific Biomarkers of Clinical and Metabolomic Characteristics in Septic Patients with Bacterial Infections. *BioMed Research International*, 1–11.
- Pedrosa, T. Í., Vasconcelos, F. F., Medeiros, L., & Silva, L. D. (2018). Machine Learning Application to Quantify the Tremor Level for Parkinson's Disease Patients. *Procedia Computer Science*, 138, 215–220.
- Muraro, C., Polato, M., Bortoli, M., Aiolfi, F., & Orian, L. (2020). Radical scavenging activity of natural antioxidants and drugs: Development of a combined machine learning and quantum chemistry protocol. *Journal of Chemical Physics*, 153(11), 1.
- Thadatritharntip, W., & Vongurai, R. (2020). Artificial Intelligence Healthcare: An Empirical Study on Users' Attitude and Intention to Use toward a Personal Home Healthcare Robot to Improve Health and Wellness Conditions in Bangkok, Thailand. *UTCC International Journal of Business & Economics*, 12(1), 3–25.
- Itkonen, P. (2019). Artificial intelligence in Home Care Settings in South Karelia Social and Healthcare District in Finland. 2019 IEEE World Congress on Services (SERVICES), 2642–939X, 238–239.
- Şensu S, Erdoğan N, Gürbüz, YS. 2020. Patolojide Dijital Çağ ve Yapay Zekâ: Temel Bilgiler. *Türkiye Klinikleri J Med Sci*; 40(1):104-12.
- Oral, C. 2011. Sayısal Mamografi Görüntülerine Yapay Zekâ Yöntemlerinin Uygulanması. Doktora Tezi, Ondokuz Mayıs Üniversitesi, Samsun, 35-56.
- Thinktech STM. İleri Sağlık Teknolojileri I - Akıllı Sağlık Uygulamaları Ve Veri Analizi İle Sağlık Sorunlarını Tanımlamak (Araştırma Raporu 2019). <http://thinktech.stm.com.tr> (Erişim Tarihi: 01.11.2020).

YAZARLARA BİLGİ

TANIM

Acta INFOLOGICA (ACIN), İstanbul Üniversitesi Enformatik Bölümü'nün yayınıdır. Açık-erişimli, bilimsel ve hakemli bir dergi olarak yılda iki defa Haziran ve Aralık aylarında yayınlanır. Derginin başlangıç tarihi 2017'dir.

AMAÇ VE KAPSAM

ACIN, veri-enformasyon-bilgi kavramlarını, bilgi-iletişim teknolojileri ve uygulamalarını temel alarak gerek enformatik alanında gerekse disiplinler arası gerçekleştirilen çalışmalar için bilimsel bir yayın ortamı sunmayı, yayınlanan çalışmalar ile bu alanda çalışan, alana ilgi duyan araştırmacılar ve ilgililerin gelişimine katkı sağlamayı amaçlamaktadır.

ACIN'ın çalışma alanları aşağıda listelenmiştir. Bu alanlarda ya da bu alanlarla ilgili olduğu düşünülen alanlardaki tüm çalışmalar derginin kapsamındadır.

Akıllı Sistemler
Bilgi Güvenliği ve Hukuk
Bilgi Yönetimi
Bilgisayar Ağları
Bilgisayar Mimarisi
Bilişim Sistemleri
Biyoenformatik
Coğrafi Bilgi Sistemleri
E-Uygulamalar
İnternet Teknolojileri
Karar Destek Sistemleri ve İş Zekası
Mikro Denetleyici ve Uygulamaları
Mobil Sistemler
Modelleme ve Optimizasyon
Sosyal ve Dijital Medya
Veri Madenciliği
Veri Tabanı Sistemleri
Yapay Zeka ve Makine Öğrenmesi
Yazılım Mühendisliği
Yönetim Bilişim Sistemleri

EDİTORYAL POLİTİKALAR VE HAKEM SÜRECİ

Yayın Politikası

Dergiye yayınlanmak üzere gönderilen makalelerin içeriği derginin amaç ve kapsamı ile uyumlu olmalıdır. Dergi, orijinal araştırma niteliğindeki yazıları yayınlamaya öncelik vermektedir.

Genel İlkeler

Daha önce yayınlanmamış ya da yayınlanmak üzere başka bir dergide halen değerlendirmede olmayan ve her bir yazar tarafından onaylanan makaleler değerlendirilmek üzere kabul edilir.

Ön değerlendirmeyi geçen yazılar iThenticate intihal tarama programından geçirilir. İntihal incelemesinden sonra, uygun makaleler Editör tarafından orijinaliteleri, metodolojileri, makalede ele alınan konunun önemi ve derginin kapsamına uygunluğu açısından değerlendirilir.

Bilimsel toplantılarda sunulan özet bildirimler, makalede belirtilmesi koşulu ile kaynak olarak kabul edilir. Editör, gönderilen makale biçimsel esaslara uygun ise, gelen yazıyı yurtiçinden ve /veya yurtdışından en az iki hakemin değerlendirmesine sunar, hakemler gerek gördüğü takdirde yazıda istenen değişiklikler yazarlar tarafından yapıldıktan sonra yayınlanmasına onay verir.

YAZARLARA BİLGİ

Makale yayınlanmak üzere Dergiye gönderildikten sonra yazarlardan hiçbirinin ismi, tüm yazarların yazılı izni olmadan yazar listesinden silinemez ve yeni bir isim yazar olarak eklenemez ve yazar sırası değiştirilemez.

Yayına kabul edilmeyen makale, resim ve fotoğraflar yazarlara geri gönderilmez.

AÇIK ERİŞİM İLKESİ

ACIN'in tüm içeriği okura ya da okurun dahil olduğu kuruma ücretsiz olarak sunulur. Okurlar, ticari amaç haricinde, yayıncı ya da yazardan izin almadan dergi makalelerinin tam metnini okuyabilir, indirebilir, kopyalayabilir, arayabilir ve link sağlayabilir.

ACIN makaleleri açık erişimlidir ve Creative Commons Atıf-GayrıTicari 4.0 Uluslararası (CC BY-NC 4.0) (<https://creativecommons.org/licenses/by-nc/4.0/deed.tr>) olarak lisanslıdır.

İşleme Ücreti

Derginin tüm giderleri İstanbul Üniversitesi tarafından karşılanmaktadır. Dergide makale yayını ve makale süreçlerinin yürütülmesi ücrete tabi değildir. Dergiye gönderilen ya da yayın için kabul edilen makaleler için işleme ücreti ya da gönderim ücreti alınmaz.

Hakem Süreci

Daha önce yayınlanmamış ya da yayınlanmak üzere başka bir dergide halen değerlendirmede olmayan ve her bir yazar tarafından onaylanan makaleler değerlendirilmek üzere kabul edilir. Gönderilen ve ön kontrolü geçen makaleler iThenticate yazılımı kullanılarak intihal için taranır. İntihal kontrolünden sonra, uygun olan makaleler baş editör tarafından orijinallik, metodoloji, işlenen konunun önemi ve dergi kapsamı ile uyumluluğu açısından değerlendirilir. Baş editör, makaleleri, yazarların etnik kökeninden, cinsiyetinden, cinsel yöneliminden, uyruğundan, dini inancından ve siyasi felsefesinden bağımsız olarak değerlendirir. Yayına gönderilen makalelerin adil bir şekilde çift taraflı kör hakem değerlendirmesinden geçmelerini sağlar.

Seçilen makaleler en az iki ulusal/uluslararası hakeme değerlendirmeye gönderilir; yayın kararı, hakemlerin talepleri doğrultusunda yazarların gerçekleştirdiği düzenlemelerin ve hakem sürecinin sonrasında baş editör tarafından verilir.

Hakemlerin değerlendirmeleri objektif olmalıdır. Hakem süreci sırasında hakemlerin aşağıdaki hususları dikkate alarak değerlendirmelerini yapmaları beklenir.

- Makale yeni ve önemli bir bilgi içeriyor mu?
- Öz, makalenin içeriğini net ve düzgün bir şekilde tanımlıyor mu?
- Yöntem bütünlüklü ve anlaşılır şekilde tanımlanmış mı?
- Yapılan yorum ve varılan sonuçlar bulgularla kanıtlanıyor mu?
- Alandaki diğer çalışmalara yeterli referans verilmiş mi?
- Dil kalitesi yeterli mi?

Hakemler, gönderilen makalelere ilişkin tüm bilginin, makale yayınlanana kadar gizli kalmasını sağlamalı ve yazar tarafında herhangi bir telif hakkı ihlali ve intihal fark ederlerse editöre raporlamalıdır. Hakem, makale konusu hakkında kendini vasıflı hissetmiyor ya da zamanında geri dönüş sağlaması mümkün görünmüyorsa, editöre bu durumu bildirmeli ve hakem sürecine kendisini dahil etmemesini istemelidir.

Değerlendirme sürecinde editör hakemlere gözden geçirme için gönderilen makalelerin, yazarların özel mülkü olduğunu ve bunun imtiyazlı bir iletişim olduğunu açıkça belirtir. Hakemler ve yayın kurulu üyeleri başka kişilerle makaleleri tartışamazlar. Hakemlerin kimliğinin gizli kalmasına özen gösterilmelidir.

TELİF HAKKINDA

Yazarlar Acta INFOLOGICA (ACIN) dergisinde yayınlanan çalışmalarının telif hakkına sahiptirler ve çalışmaları Creative Commons Atıf-GayrıTicari 4.0 Uluslararası (CC BY-NC 4.0) olarak lisanslıdır. Creative Commons Atıf-GayrıTicari 4.0 Uluslararası (CC BY-NC 4.0) lisansı, eserin ticari kullanım dışında her boyut ve formatta paylaşılmasına, kopyalanmasına, çoğaltılmasına ve orijinal esere uygun şekilde atıfta bulunmak kaydıyla yeniden düzenleme, dönüştürme ve eserin üzerine inşa etme dâhil adapte edilmesine izin verir.

YAZARLARA BİLGİ

YAYIN ETİĞİ VE İLKELER

Acta INFOLOGICA (ACIN), yayın etiğinde en yüksek standartlara bağlıdır ve Committee on Publication Ethics (COPE), Directory of Open Access Journals (DOAJ), Open Access Scholarly Publishers Association (OASPA) ve World Association of Medical Editors (WAME) tarafından yayınlanan etik yayıncılık ilkelerini benimser; Principles of Transparency and Best Practice in Scholarly Publishing başlığı altında ifade edilen ilkeler için: <https://publicationethics.org/resources/guidelines-new/principles-transparency-and-best-practice-scholarly-publishing>

Gönderilen tüm makaleler orijinal, yayınlanmamış ve başka bir dergide değerlendirme sürecinde olmamalıdır. Her bir makale editörlerden biri ve en az iki hakem tarafından çift kör değerlendirilmeden geçirilir. İntihal, duplikasyon, sahte yazarlık/inkar edilen yazarlık, araştırma/veri fabrikasyonu, makale dilimleme, dilimleyerek yayın, telif hakları ihlali ve çıkar çatışmasının gizlenmesi, etik dışı davranışlar olarak kabul edilir.

Kabul edilen etik standartlara uygun olmayan tüm makaleler yayından çıkarılır. Buna yayından sonra tespit edilen olası kuraldışı, uygunsuzluklar içeren makaleler de dahildir.

ARAŞTIRMA ETİĞİ

Acta INFOLOGICA (ACIN) araştırma etiğinde en yüksek standartları gözetir ve aşağıda tanımlanan uluslararası araştırma etiği ilkelerini benimser. Makalelerin etik kurallara uygunluğu yazarların sorumluluğundadır.

- Araştırmanın tasarlanması, tasarımın gözden geçirilmesi ve araştırmanın yürütülmesinde, bütünlük, kalite ve şeffaflık ilkeleri sağlanmalıdır.
- Araştırma ekibi ve katılımcılar, araştırmanın amacı, yöntemleri ve öngörülen olası kullanımları; araştırmaya katılımın gerektirdikleri ve varsa riskleri hakkında tam olarak bilgilendirilmelidir.
- Araştırma katılımcılarının sağladığı bilgilerin gizliliği ve yanıt verenlerin gizliliği sağlanmalıdır. Araştırma katılımcıların özerkliğini ve saygınlığını koruyacak şekilde tasarlanmalıdır.
- Araştırma katılımcıları gönüllü olarak araştırmada yer almalı, herhangi bir zorlama altında olmamalıdır.
- Katılımcıların zarar görmesinden kaçınılmalıdır. Araştırma, katılımcıları riske sokmayacak şekilde planlanmalıdır.
- Araştırma bağımsızlığıyla ilgili açık ve net olunmalı; çıkar çatışması varsa belirtilmelidir.
- Deneysel çalışmalarda, araştırmaya katılmaya karar veren katılımcıların yazılı bilgilendirilmiş onayı alınmalıdır. Çocukların ve vesayet altındakilerin veya tasdiklenmiş akıl hastalığı bulunanların yasal vasisinin onayı alınmalıdır.
- Çalışma herhangi bir kurum ya da kuruluşta gerçekleştirilecekse bu kurum ya da kuruluştan çalışma yapılacağına dair onay alınmalıdır.
- İnsan ögesi bulunan çalışmalarda, “yöntem” bölümünde katılımcılardan “bilgilendirilmiş onam” alındığının ve çalışmanın yapıldığı kurumdaki etik kurul onayı alındığı belirtilmesi gerekir.

YAZARLARIN SORUMLULUĞU

Makalelerin bilimsel ve etik kurallara uygunluğu yazarların sorumluluğundadır. Yazar makalenin orijinal olduğu, daha önce başka bir yerde yayınlanmadığı ve başka bir yerde, başka bir dilde yayınlanmak üzere değerlendirmede olmadığı konusunda teminat sağlamalıdır. Uygulamadaki telif kanunları ve anlaşmaları gözetilmelidir. Telifle bağlı materyaller (örneğin tablolar, şekiller veya büyük alıntılar) gerekli izin ve teşekkürle kullanılmalıdır. Başka yazarların, katkıda bulunanların çalışmaları ya da yararlanılan kaynaklar uygun biçimde kullanılmalı ve referanslarda belirtilmelidir.

Gönderilen makalede tüm yazarların akademik ve bilimsel olarak doğrudan katkısı olmalıdır, bu bağlamda “yazar” yayınlanan bir araştırmanın kavramsallaştırılmasına ve tasarımına, verilerin elde edilmesine, analizine ya da yorumlanmasına belirgin katkı yapan, yazının yazılması ya da bunun içerik açısından eleştirel biçimde gözden geçirilmesinde görev yapan birisi olarak görülür. Yazar olabilmenin diğer koşulları ise, makaledeki çalışmayı planlamak veya icra etmek ve / veya revize etmektir. Fon sağlanması, veri toplanması ya da araştırma grubunun genel süpervizyonu tek başına yazarlık hakkı kazandırmaz. Yazar olarak gösterilen tüm bireyler sayılan tüm ölçütleri karşılamalıdır ve yukarıdaki ölçütleri karşılayan her birey yazar olarak gösterilebilir. Yazarların isim sıralaması ortak verilen bir karar olmalıdır. Tüm yazarlar yazar sıralamasını Telif Hakkı Devir Formunda imzalı olarak belirtmek zorundadırlar.

Yazarlık için yeterli ölçütleri karşılamayan ancak çalışmaya katkısı olan tüm bireyler “teşekkür / bilgileri” kısmında sıralanmalıdır. Bunlara

YAZARLARA BİLGİ

örnek olarak ise sadece teknik destek sağlayan, yazıma yardımcı olan ya da sadece genel bir destek sağlayan, finansal ve materyal desteği sunan kişiler verilebilir.

Bütün yazarlar, araştırmanın sonuçlarını ya da bilimsel değerlendirmeyi etkileyebilme potansiyeli olan finansal ilişkiler, çıkar çatışması ve çıkar rekabetini beyan etmelidirler. Bir yazar kendi yayınlanmış yazısında belirgin bir hata ya da yanlışlık tespit ederse, bu yanlışlıklara ilişkin düzeltme ya da geri çekme için editör ile hemen temasa geçme ve işbirliği yapma sorumluluğunu taşır.

EDİTÖR VE HAKEM SORUMLULUKLARI

Baş editör, makaleleri, yazarların etnik kökeninden, cinsiyetinden, cinsel yöneliminden, uyruğundan, dini inancından ve siyasi felsefesinden bağımsız olarak değerlendirir. Yayına gönderilen makalelerin adil bir şekilde çift taraflı kör hakem değerlendirmesinden geçmelerini sağlar. Gönderilen makalelere ilişkin tüm bilginin, makale yayınlanana kadar gizli kalacağını garanti eder. Baş editör içerik ve yayının toplam kalitesinden sorumludur. Gereğinde hata sayfası yayınlamalı ya da düzeltme yapmalıdır.

Baş editör; yazarlar, editörler ve hakemler arasında çıkar çatışmasına izin vermez. Hakem atama konusunda tam yetkiye sahiptir ve Dergide yayınlanacak makalelerle ilgili nihai kararı vermekle yükümlüdür.

Hakemlerin araştırmayla ilgili, yazarlarla ve/veya araştırmanın finansal destekçileriyle çıkar çatışmaları olmamalıdır. Değerlendirmelerinin sonucunda tarafsız bir yargıya varmalıdırlar. Gönderilmiş yazılara ilişkin tüm bilginin gizli tutulmasını sağlamalı ve yazar tarafında herhangi bir telif hakkı ihlali ve intihal fark ederlerse editöre raporlamalıdırlar. Hakem, makale konusu hakkında kendini vasıflı hissetmiyor ya da zamanında geri dönüş sağlaması mümkün görünmüyorsa, editöre bu durumu bildirmeli ve hakem sürecine kendisini dahil etmemesini istemelidir.

Değerlendirme sürecinde editör hakemlere gözden geçirme için gönderilen makalelerin, yazarların özel mülkü olduğunu ve bunun imtiyazlı bir iletişim olduğunu açıkça belirtir. Hakemler ve yayın kurulu üyeleri başka kişilerle makaleleri tartışamazlar. Hakemlerin kimliğinin gizli kalmasına özen gösterilmelidir. Bazı durumlarda editörün kararıyla, ilgili hakemlerin makaleye ait yorumları aynı makaleyi yorumlayan diğer hakemlere gönderilerek hakemlerin bu süreçte aydınlatılması sağlanabilir.

YAZILARIN HAZIRLANMASI

DİL

Derginin yayın dili Türkçe ve Amerikan İngilizcesi'dir.

Yazıların Hazırlanması ve Yazım Kuralları

Aksi belirtilmedikçe gönderilen yazılarla ilgili tüm yazışmalar ilk yazarla yapılacaktır. Makale gönderimi online olarak ve <http://acin.istanbul.edu.tr> adresinden erişilen <http://dergipark.gov.tr/login> üzerinden yapılmalıdır. Gönderilen yazılar, makale türünü belirten ve makaleyle ilgili detayları içeren (bkz: Son Kontrol Listesi) kapak sayfası; editöre mektup, yazının elektronik formunu içeren Microsoft Word 2003 ve üzerindeki versiyonları ile yazılmış elektronik dosya ve tüm yazarların imzaladığı Telif Hakkı Devir Formu eklenerek gönderilmelidir.

- Microsoft Word 6.0 ya da üstü bir versiyon kullanıyorsanız ACIN Makale Şablonunu kullanabilirsiniz. Aksi halde, bu doküman bir yönerge olarak kullanılabilir.
- Makale başlıkları büyük harf ve küçük harflerden oluşmalı, bütün harfler büyük olmamalıdır. Başlığa formül yazmaktan kaçınılmalıdır. Başlıkta "(Davetli)" ya da benzeri ifadeler yer almamalıdır.
- Öz 150-250 kelime arasında olmalıdır, bir paragraf olarak yazılmalı ve matematiksel denklem ya da tablo içermemelidir. Öz, okuyucunun kolaylıkla bulabilmesi için, üç ya da dört anahtar kelime ya da ifade içermelidir. Öz iyi okunabilir ve de dilbilgisi açısından doğru olmalıdır.
- Özün altında çalışmanın içeriğini temsil eden üç anahtar kelime olmalıdır. Anahtar kelimelerin, "TR Dizin Anahtar Terimler Listesi", "Medical Subject Headings", "CAB Theasarus", "JISCT", "ERIC" vd. tarafından tanımlanmış olmasına önem verilmelidir.
- Çalışmaların başlıca şu unsurları içermesi gerekmektedir: Türkçe başlık, öz ve anahtar kelimeler; İngilizce başlık öz ve anahtar kelimeler; ana metin bölümleri, kaynaklar, tablolar ve şekiller.
- Yayınlanmak üzere gönderilen makale ile birlikte yazar bilgilerini içeren kapak sayfası gönderilmelidir. Kapak sayfasında, makalenin

YAZARLARA BİLGİ

başlığı, yazar veya yazarların bağlı oldukları kurum ve unvanları, kendilerine ulaşılabilecek adresler, cep, iş ve faks numaraları, ORCID ve e-posta adresleri yer almalıdır (bkz. Son Kontrol Listesi).

7. Referanslar APA 6 stiline uygun olarak hazırlanmalıdır.

KAYNAKLAR

Referans Stili ve Formatı

Acta INFOLOGICA (ACIN), metin içi alıntılama ve kaynak gösterme için APA (American Psychological Association) kaynak sitilinin 6. edisyonunu benimser. APA 6. Edisyon hakkında bilgi için:

- American Psychological Association. (2010). Publication manual of the American Psychological Association (6th ed.). Washington, DC: APA.

- <http://www.apastyle.org/>

Kaynakların doğruluğundan yazar(lar) sorumludur. Tüm kaynaklar metinde belirtilmelidir. Kaynaklar aşağıdaki örneklerdeki gibi gösterilmelidir.

Metin İçinde Kaynak Gösterme

Kaynaklar metinde parantez içinde yazarların soyadı ve yayın tarihi yazılarak belirtilmelidir. Birden fazla kaynak gösterilecekse kaynaklar arasında (;) işareti kullanılmalıdır. Kaynaklar alfabetik olarak sıralanmalıdır.

Örnekler:

Birden fazla kaynak;

(Esin ve ark., 2002; Karasar 1995)

Tek yazarlı kaynak;

(Akyolcu, 2007)

İki yazarlı kaynak;

(Saymer ve Demirci, 2007, s. 72)

Üç, dört ve beş yazarlı kaynak;

Metin içinde ilk kullanımda: (Ailen, Ciambrene ve Welch, 2000, s. 12–13) Metin içinde tekrarlayan kullanımlarda: (Ailen ve ark., 2000)

Altı ve daha çok yazarlı kaynak;

(Çavdar ve ark., 2003)

Kaynaklar Bölümünde Kaynak Gösterme

Kullanılan tüm kaynaklar metnin sonunda ayrı bir bölüm halinde yazar soyadlarına göre alfabetik olarak numaralandırılmadan verilmelidir.

Kaynak yazımı ile ilgili örnekler aşağıda verilmiştir.

Kitap

a) *Türkçe Kitap*

Karasar, N. (1995). Araştırmalarda rapor hazırlama (8.bs). Ankara: 3A Eğitim Danışmanlık Ltd.

b) *Türkçeye Çevrilmiş Kitap*

Mucchielli, A. (1991). Zihniyetler (A. Kotil, Çev.). İstanbul: İletişim Yayınları.

c) *Editörlü Kitap*

Ören, T., Üney, T. ve Çölkesen, R. (Ed.). (2006). Türkiye bilişim ansiklopedisi. İstanbul: Papatya Yayıncılık.

d) *Çok Yazarlı Türkçe Kitap*

Tonta, Y., Bitirim, Y. ve Sever, H. (2002). Türkçe arama motorlarında performans değerlendirme. Ankara: Total Bilişim.

e) *İngilizce Kitap*

Kamien R., & Kamien A. (2014). Music: An appreciation. New York, NY: McGraw-Hill Education.

f) *İngilizce Kitap İçerisinde Bölüm*

Bassett, C. (2012). Cultural studies and new media. In G. Hall & C. Birchall (Eds.), New cultural studies: Adventures in theory (pp. 220–237). Edinburgh, UK: Edinburgh University Press.

g) *Türkçe Kitap İçerisinde Bölüm*

Erkmen, T. (2012). Örgüt kültürü: Fonksiyonları, öğeleri, işletme yönetimi ve liderlikteki önemi. M. Zencirkıran (Ed.), Örgüt sosyolojisi kitabı içinde (s. 233–263). Bursa: Dora Basım Yayın.

h) Yayıncının ve Yazarın Kurum Olduğu Yayın

Türk Standartları Enstitüsü. (1974). Adlandırma ilkeleri. Ankara: Yazar.

Makale

a) Türkçe Makale

Mutlu, B. ve Savaşer, S. (2007). Çocuğu ameliyat sonrası yoğun bakımda olan ebeveynlerde stres nedenleri ve azaltma girişimleri. İstanbul Üniversitesi Florence Nightingale Hemşirelik Dergisi, 15(60), 179–182.

b) İngilizce Makale

de Cillia, R., Reisingl, M., & Wodak, R. (1999). The discursive construction of national identity. *Discourse and Society*, 10(2), 149–173. <http://dx.doi.org/10.1177/0957926599010002002>

c) Yediden Fazla Yazarlı Makale

Lal, H., Cunningham, A. L., Godeaux, O., Chlibek, R., Diez-Domingo, J., Hwang, S.-J. ... Heineman, T. C. (2015). Efficacy of an adjuvanted herpes zoster subunit vaccine in older adults. *New England Journal of Medicine*, 372, 2087–2096. <http://dx.doi.org/10.1056/NEJMoa1501184>

d) DOI'si Olmayan Online Edinilmiş Makale

Al, U. ve Doğan, G. (2012). Hacettepe Üniversitesi Bilgi ve Belge Yönetimi Bölümü tezlerinin atıf analizi. *Türk Kütüphaneciliği*, 26, 349–369. Erişim adresi: <http://www.tk.org.tr/>

e) DOI'si Olan Makale

Turner, S. J. (2010). Website statistics 2.0: Using Google Analytics to measure library website effectiveness. *Technical Services Quarterly*, 27, 261–278. <http://dx.doi.org/10.1080/07317131003765910>

f) Advance Online Olarak Yayımlanmış Makale

Smith, J. A. (2010). Citing advance online publication: A review. *Journal of Psychology*. Advance online publication. <http://dx.doi.org/10.1037/a45d7867>

g) Popüler Dergi Makalesi

Semerçioğlu, C. (2015, Haziran). Sıradanlığın rayihası. *Sabit Fikir*, 52, 38–39.

Tez, Sunum, Bildiri

a) Türkçe Tezler

Sarı, E. (2008). Kültür kimlik ve politika: Mardin’de kültürlerarasılık. (Doktora Tezi). Ankara Üniversitesi Sosyal Bilimler Enstitüsü, Ankara.

b) Ticari Veritabanında Yer Alan Yüksek Lisans Ya da Doktora Tezi

Van Brunt, D. (1997). Networked consumer health information systems (Doctoral dissertation). Available from ProQuest Dissertations and Theses. (UMI No. 9943436)

c) Kurumsal Veritabanında Yer Alan İngilizce Yüksek Lisans/Doktora Tezi

Yaylalı-Yıldız, B. (2014). University campuses as places of potential publicness: Exploring the political, social and cultural practices in Ege University (Doctoral dissertation). Retrieved from: <http://library.iyte.edu.tr/hizli-erisim/iyte-tez-portali>

d) Web’de Yer Alan İngilizce Yüksek Lisans/Doktora Tezi

Tonta, Y. A. (1992). An analysis of search failures in online library catalogs (Doctoral dissertation, University of California, Berkeley). Retrieved from <http://yunus.hacettepe.edu.tr/~tonta/yayinlar/phd/ickapak.html>

e) Dissertations Abstracts International’da Yer Alan Yüksek Lisans/Doktora Tezi

Appelbaum, L. G. (2005). Three studies of human information processing: Texture amplification, motion representation, and figure-ground segregation. *Dissertation Abstracts International: Section B. Sciences and Engineering*, 65(10), 5428.

f) Sempozyum Katkısı

Krinsky-McHale, S. J., Zigman, W. B., & Silverman, W. (2012, August). Are neuropsychiatric symptoms markers of prodromal Alzheimer’s disease in adults with Down syndrome? In W. B. Zigman (Chair), Predictors of mild cognitive impairment, dementia, and mortality in adults with Down syndrome. Symposium conducted at American Psychological Association meeting, Orlando, FL.

g) Online Olarak Erişilen Konferans Bildiri Özeti

Çınar, M., Doğan, D. ve Seferoğlu, S. S. (2015, Şubat). Eğitimde dijital araçlar: Google sınıf uygulaması üzerine bir değerlendirme [Öz]. Akademik Bilişim Konferansında sunulan bildiri, Anadolu Üniversitesi, Eskişehir. Erişim adresi: <http://ab2015.anadolu.edu.tr/index.php?menu=5&submenu=27>

h) Düzenli Olarak Online Yayımlanan Bildiriler

Herculano-Houzel, S., Collins, C. E., Wong, P., Kaas, J. H., & Lent, R. (2008). The basic nonuniformity of the cerebral cortex. *Proceedings of the National Academy of Sciences*, 105, 12593–12598. <http://dx.doi.org/10.1073/pnas.0805417105>

i) Kitap Şeklinde Yayımlanan Bildiriler

Schneider, R. (2013). Research data literacy. S. Kurbanoğlu ve ark. (Ed.), *Communications in Computer and Information Science: Vol. 397. Worldwide Communalities and Challenges in Information Literacy Research and Practice içinde* (s. 134–140). Cham, İsviçre: Springer. <http://dx.doi.org/10.1007/978-3-319-03919-0>

YAZARLARA BİLGİ

j) Kongre Bildirisi

Çepni, S., Bacanak A. ve Özsevgeç T. (2001, Haziran). Fen bilgisi öğretmen adaylarının fen branşlarına karşı tutumları ile fen branşlarındaki başarılarının ilişkisi. X. Ulusal Eğitim Bilimleri Kongresi'nde sunulan bildiri, Abant İzzet Baysal Üniversitesi, Bolu

Diğer Kaynaklar

a) Gazete Yazısı

Toker, Ç. (2015, 26 Haziran). 'Unutma' notları. Cumhuriyet, s. 13.

b) Online Gazete Yazısı

Tamer, M. (2015, 26 Haziran). E-ticaret hamle yapmak için tüketiciyi bekliyor. Milliyet. Erişim adresi: <http://www.milliyet>

c) Web Page/Blog Post

Bordwell, D. (2013, June 18). David Koepp: Making the world movie-sized [Web log post]. Retrieved from <http://www.davidbordwell.net/blog/page/27/>

d) Online Ansiklopedi/Sözlük

Bilgi mimarisi. (2014, 20 Aralık). Vikipedi içinde. Erişim adresi: http://tr.wikipedia.org/wiki/Bilgi_mimarisi

Marcoux, A. (2008). Business ethics. In E. N. Zalta (Ed.), The Stanford encyclopedia of philosophy. Retrieved from <http://plato.stanford.edu/entries/ethics-business/>

e) Podcast

Radyo ODTÜ (Yapımcı). (2015, 13 Nisan). Modern sabahlar [Podcast]. Erişim adresi: <http://www.radyoodtu.com.tr/>

f) Bir Televizyon Dizisinden Tek Bir Bölüm

Shore, D. (Senarist), Jackson, M. (Senarist) ve Bookstaver, S. (Yönetmen). (2012). Runaways [Televizyon dizisi bölümü]. D. Shore (Baş yapımcı), House M.D. içinde. New York, NY: Fox Broadcasting.

g) Müzik Kaydı

Say, F. (2009). Galata Kulesi. İstanbul senfonisi [CD] içinde. İstanbul: Ak Müzik.

SON KONTROL LİSTESİ

Aşağıdaki listede eksik olmadığından emin olun:

- Editöre mektup
 - Makalenin türü
 - Başka bir dergiye gönderilmemiş olduğu bilgisi
 - Sponsor veya ticari bir firma ile ilişkisi (varsa belirtiniz)
 - Kaynakların APA6'ya göre belirtildiği
 - İngilizce yönünden kontrolünün yapıldığı
 - Yazarlara Bilgide detaylı olarak anlatılan dergi politikalarının gözden geçirildiği
- Telif Hakkı Devir Formu
- Daha önce basılmış materyal (yazı-resim-tablo) kullanılmış ise izin belgesi
- Kapak sayfası
 - Makalenin kategorisi
 - Makale dilinde ve İngilizce başlık

INFORMATION FOR AUTHORS

DESCRIPTION

Acta INFOLOGICA (ACIN) is the publication of Informatics Department of the Istanbul University. It is an open access, scholarly, peer-reviewed journal published biannually in June and December. The journal was founded in 2017.

AIM AND SCOPE

ACIN aims to contribute to the scientific community interested in the field of informatics and aims to provide a platform for researchers exploring issues based on the concepts of data-information-knowledge, information and communication technologies and applications. The journal welcomes multidisciplinary studies regarding the field as well.

The areas of study covered in the scope of ACIN are in below;

Intelligent Systems

Information Security and Law

Knowledge Management

Computer Networks

Computer Architecture

Information Systems

Bioinformatics

Geographic Information Systems

E-Applications

Internet Technologies

Decision Support Systems and Business Intelligence

Microcontroller and Applications

Mobile Systems

Modeling and Optimization

Project Management

Social and Digital Media

Data Mining

Database Systems

Artificial Intelligence and Machine Learning

Software Engineering

EDITORIAL POLICIES AND PEER REVIEW PROCESS

Publication Policy

The subjects covered in the manuscripts submitted to the Journal for publication must be in accordance with the aim and scope of the journal. The journal gives priority to original research papers submitted for publication.

General Principles

Only those manuscripts approved by its every individual author and that were not published before in or sent to another journal, are accepted for evaluation.

Submitted manuscripts that pass preliminary control are scanned for plagiarism using iThenticate software. After plagiarism check, the eligible ones are evaluated by editor-in-chief for their originality, methodology, the importance of the subject covered and compliance with the journal scope.

Short presentations that took place in scientific meetings can be referred if indicated in the article. The editor hands over the papers matching the formal rules to at least two national/international referees for evaluation and gives green light for publication upon modification by the authors in accordance with the referees' claims. Changing the name of an author (omission, addition or order) in papers submitted to the Journal requires written permission of all declared authors. Refused manuscripts and graphics are not returned to the author.

INFORMATION FOR AUTHORS

OPEN ACCESS STATEMENT

Acta INFOLOGICA (ACIN) is an open access journal which means that all content is freely available without charge to the user or his/her institution. Except for commercial purposes, users are allowed to read, download, copy, print, search, or link to the full texts of the articles in this journal without asking prior permission from the publisher or the author.

The articles in ACIN are open access articles licensed under the terms of the Creative Commons Attribution-NonCommercial 4.0 International (CC BY-NC 4.0) license (<https://creativecommons.org/licenses/by-nc/4.0/deed.en>)

Article Processing Charge

All expenses of the journal are covered by the Istanbul University. Processing and publication are free of charge with the journal. There is no article processing charges or submission fees for any submitted or accepted articles.

Peer Review Process

Only those manuscripts approved by its every individual author and that were not published before in or sent to another journal, are accepted for evaluation.

Submitted manuscripts that pass preliminary control are scanned for plagiarism using iThenticate software. After plagiarism check, the eligible ones are evaluated by Editor-in-Chief for their originality, methodology, the importance of the subject covered and compliance with the journal scope. Editor-in-Chief evaluates manuscripts for their scientific content without regard to ethnic origin, gender, sexual orientation, citizenship, religious belief or political philosophy of the authors and ensures a fair double-blind peer review of the selected manuscripts.

The selected manuscripts are sent to at least two national/international referees for evaluation and publication decision is given by Editor-in-Chief upon modification by the authors in accordance with the referees' claims.

Editor-in-Chief does not allow any conflicts of interest between the authors, editors and reviewers and is responsible for final decision for publication of the manuscripts in the Journal.

Reviewers' judgments must be objective. Reviewers' comments on the following aspects are expected while conducting the review.

- Does the manuscript contain new and significant information?
- Does the abstract clearly and accurately describe the content of the manuscript?
- Is the problem significant and concisely stated?
- Are the methods described comprehensively?
- Are the interpretations and conclusions justified by the results?
- Is adequate references made to other Works in the field?
- Is the language acceptable?

Reviewers must ensure that all the information related to submitted manuscripts is kept as confidential and must report to the editor if they are aware of copyright infringement and plagiarism on the author's side.

A reviewer who feels unqualified to review the topic of a manuscript or knows that its prompt review will be impossible should notify the editor and excuse himself from the review process.

The editor informs the reviewers that the manuscripts are confidential information and that this is a privileged interaction. The reviewers and editorial board cannot discuss the manuscripts with other persons. The anonymity of the referees is important.

COPYRIGHT NOTICE

Authors publishing with Acta INFOLOGICA (ACIN) retain the copyright to their work, licensing it under the Creative Commons Attribution-NonCommercial 4.0 International (CC BY-NC 4.0) license that gives permission to copy and redistribute the material in any medium or format other than commercial purposes as well as remix, transform and build upon the material by providing appropriate credit to the original work.

INFORMATION FOR AUTHORS

PUBLICATION ETHICS AND PUBLICATION MALPRACTICE STATEMENT

Acta INFOLOGICA (ACIN) is committed to upholding the highest standards of publication ethics and pays regard to Principles of Transparency and Best Practice in Scholarly Publishing published by the Committee on Publication Ethics (COPE), the Directory of Open Access Journals (DOAJ), to access the Open Access Scholarly Publishers Association (OASPA), and the World Association of Medical Editors (WAME) on <https://publicationethics.org/resources/guidelines-new/principles-transparency-and-best-practice-scholarly-publishing>. All parties involved in the publishing process (Editors, Reviewers, Authors and Publishers) are expected to agree on the following ethical principles.

All submissions must be original, unpublished (including as full text in conference proceedings), and not under the review of any other publication synchronously. Each manuscript is reviewed by one of the editors and at least two referees under double-blind peer review process. Plagiarism, duplication, fraud authorship/denied authorship, research/data fabrication, salami slicing/salami publication, breaching of copyrights, prevailing conflict of interest are unethical behaviors.

All manuscripts not in accordance with the accepted ethical standards will be removed from the publication. This also contains any possible malpractice discovered after the publication. In accordance with the code of conduct we will report any cases of suspected plagiarism or duplicate publishing.

RESEARCH ETHICS

Acta INFOLOGICA (ACIN) adheres to the highest standards in research ethics and follows the principles of international research ethics as defined below. The authors are responsible for the compliance of the manuscripts with the ethical rules.

- Principles of integrity, quality and transparency should be sustained in designing the research, reviewing the design and conducting the research.
- The research team and participants should be fully informed about the aim, methods, possible uses and requirements of the research and risks of participation in research.
- The confidentiality of the information provided by the research participants and the confidentiality of the respondents should be ensured. The research should be designed to protect the autonomy and dignity of the participants.
- Research participants should participate in the research voluntarily, not under any coercion.
- Any possible harm to participants must be avoided. The research should be planned in such a way that the participants are not at risk.
- The independence of research must be clear; and any conflict of interest or must be disclosed.
- In experimental studies with human subjects, written informed consent of the participants who decide to participate in the research must be obtained. In the case of children and those under wardship or with confirmed insanity, legal custodian's assent must be obtained.
- If the study is to be carried out in any institution or organization, approval must be obtained from this institution or organization.
- In studies with human subject, it must be noted in the method's section of the manuscript that the informed consent of the participants and ethics committee approval from the institution where the study has been conducted have been obtained.

AUTHOR RESPONSIBILITIES

It is authors' responsibility to ensure that the article is in accordance with scientific and ethical standards and rules. And authors must ensure that submitted work is original. They must certify that the manuscript has not previously been published elsewhere or is not currently being considered for publication elsewhere, in any language. Applicable copyright laws and conventions must be followed. Copyright material (e.g. tables, figures or extensive quotations) must be reproduced only with appropriate permission and acknowledgement. Any work or words of other authors, contributors, or sources must be appropriately credited and referenced.

All the authors of a submitted manuscript must have direct scientific and academic contribution to the manuscript. The author(s) of the original research articles is defined as a person who is significantly involved in "conceptualization and design of the study", "collecting the data", "analyzing the data", "writing the manuscript", "reviewing the manuscript with a critical perspective" and "planning/conducting the study of the manuscript and/or revising it". Fund raising, data collection or supervision of the research group are not sufficient roles to be accepted as an author. The author(s) must meet all these criteria described above. The order of names in the author list of an article must be a co-decision and it must be indicated in the Copyright Transfer Form. The individuals who do not meet the authorship criteria but contributed to the study must take place in the acknowledgement section. Individuals providing technical support, assisting writing, providing a general support, providing material or financial support are examples to be indicated in acknowledgement section.

INFORMATION FOR AUTHORS

All authors must disclose all issues concerning financial relationship, conflict of interest, and competing interest that may potentially influence the results of the research or scientific judgment.

When an author discovers a significant error or inaccuracy in his/her own published paper, it is the author's obligation to promptly cooperate with the Editor to provide retractions or corrections of mistakes.

RESPONSIBILITY FOR THE EDITOR AND REVIEWERS

Editor-in-Chief evaluates manuscripts for their scientific content without regard to ethnic origin, gender, sexual orientation, citizenship, religious belief or political philosophy of the authors. He/She provides a fair double-blind peer review of the submitted articles for publication and ensures that all the information related to submitted manuscripts is kept as confidential before publishing.

Editor-in-Chief is responsible for the contents and overall quality of the publication. He/She must publish errata pages or make corrections when needed.

Editor-in-Chief does not allow any conflicts of interest between the authors, editors and reviewers. Only he has the full authority to assign a reviewer and is responsible for final decision for publication of the manuscripts in the Journal.

Reviewers must have no conflict of interest with respect to the research, the authors and/or the research funders. Their judgments must be objective.

Reviewers must ensure that all the information related to submitted manuscripts is kept as confidential and must report to the editor if they are aware of copyright infringement and plagiarism on the author's side.

A reviewer who feels unqualified to review the topic of a manuscript or knows that its prompt review will be impossible should notify the editor and excuse himself from the review process.

The editor informs the reviewers that the manuscripts are confidential information and that this is a privileged interaction. The reviewers and editorial board cannot discuss the manuscripts with other persons. The anonymity of the referees must be ensured. In particular situations, the editor may share the review of one reviewer with other reviewers to clarify a particular point.

MANUSCRIPT ORGANIZATION

LANGUAGE

The language of the journal is both Turkish and American English.

Manuscript Organization and Submission

All correspondence will be sent to the first-named author unless otherwise specified. Manuscript is to be submitted online via <http://dergipark.gov.tr/login> that can be accessed at <http://acin.istanbul.edu.tr> and it must be accompanied by a title page specifying the article category (i.e. research article, review etc.) and including information about the manuscript (see the Submission Checklist) and cover letter to the editor. Manuscripts should be prepared in Microsoft Word 2003 and upper versions. In addition, Copyright Transfer Form that has to be signed by all authors must be submitted.

1. Use ACIN article document as a template if you are using Microsoft Word 6.0 or upper versions. Otherwise, use this document as an instruction set.
 2. The first letters of words in the article title should be written in uppercase; the entire title should not be capitalized. Avoid writing formulas in the title. Do not write "(Invited)" or similar expressions in the title.
 3. The abstract must be between 150–250 words and written as one paragraph. It should not contain displayed mathematical equations or tabular material. The abstract should include three or four different keywords or phrases, as this will help readers to find it. It is important to avoid over-repetition of such phrases as this can result in a page being rejected by search engines. Ensure that your abstract reads well and is grammatically correct.
 4. Underneath the abstracts, 3 keywords that inform the reader about the content of the study should be specified. Keywords must be defined by taking into consideration authorities like "TR Dizin Anahtar Terimler Listesi", "Medical Subject Headings", "CAB
-

INFORMATION FOR AUTHORS

Theasarus”, “JISCT, “ERIC”, etc.

5. The manuscripts should contain mainly these components: title, abstract and keywords; sections, references, tables and figures.
6. A title page including author information must be submitted together with the manuscript. The title page is to include fully descriptive title of the manuscript and, affiliation, title, e-mail address, ORCID, postal address, phone, mobile phone and fax number of the author(s) (see The Submission Checklist).
7. References should be prepared as APA 6th edition.

REFERENCES

Reference Style and Format

Acta INFOLOGICA (ACIN) complies with APA (American Psychological Association) style 6th Edition for referencing and quoting. For more information:

- American Psychological Association. (2010). Publication manual of the American Psychological Association (6th ed.). Washington, DC: APA.
- <http://www.apastyle.org>

Accuracy of citation is the author’s responsibility. All references should be cited in text. Reference list must be in alphabetical order. Type references in the style shown below.

Citations in the Text

Citations must be indicated with the author surname and publication year within the parenthesis.

If more than one citation is made within the same paranthesis, separate them with (;).

Samples:

More than one citation;

(Esin, et al., 2002; Karasar, 1995)

Citation with one author;

(Akyolcu, 2007)

Citation with two authors;

(Saymer & Demirci, 2007)

Citation with three, four, five authors;

First citation in the text: (Ailen, Ciembrune, & Welch, 2000) Subsequent citations in the text: (Ailen, et al., 2000)

Citations with more than six authors;

(Çavdar, et al., 2003)

Citations in the Reference

All the citations done in the text should be listed in the References section in alphabetical order of author surname without numbering. Below given examples should be considered in citing the references.

Basic Reference Types

Book

a) Turkish Book

Karasar, N. (1995). *Araştırmalarda rapor hazırlama* (8th ed.) [Preparing research reports]. Ankara, Turkey: 3A Eğitim Danışmanlık Ltd.

b) Book Translated into Turkish

Mucchielli, A. (1991). *Zihniyetler* [Mindsets] (A. Kotil, Trans.). İstanbul, Turkey: İletişim Yayınları.

c) Edited Book

Ören, T., Üney, T., & Çölkesen, R. (Eds.). (2006). *Türkiye bilişim ansiklopedisi* [Turkish Encyclopedia of Informatics]. İstanbul, Turkey: Papatya Yayıncılık.

d) Turkish Book with Multiple Authors

Tonta, Y., Bitirim, Y., & Sever, H. (2002). *Türkçe arama motorlarında performans değerlendirme* [Performance evaluation in Turkish search engines]. Ankara, Turkey: Total Bilişim.

e) Book in English

Kamien R., & Kamien A. (2014). *Music: An appreciation*. New York, NY: McGraw-Hill Education.

f) Chapter in an Edited Book

Bassett, C. (2006). Cultural studies and new media. In G. Hall & C. Birchall (Eds.), *New cultural studies: Adventures in theory* (pp. 220–237). Edinburgh, UK: Edinburgh University Press.

INFORMATION FOR AUTHORS

g) Chapter in an Edited Book in Turkish

Erkmen, T. (2012). Örgüt kültürü: Fonksiyonları, öğeleri, işletme yönetimi ve liderlikteki önemi [Organization culture: Its functions, elements and importance in leadership and business management]. In M. Zencirkiran (Ed.), *Örgüt sosyolojisi* [Organization sociology] (pp. 233–263). Bursa, Turkey: Dora Basım Yayın.

h) Book with the same organization as author and publisher

American Psychological Association. (2009). *Publication manual of the American psychological association* (6th ed.). Washington, DC: Author.

Article

a) Turkish Article

Mutlu, B., & Savaşer, S. (2007). Çocuğu ameliyat sonrası yoğun bakımda olan ebeveynlerde stres nedenleri ve azaltma girişimleri [Source and intervention reduction of stress for parents whose children are in intensive care unit after surgery]. *Istanbul University Florence Nightingale Journal of Nursing*, 15(60), 179–182.

b) English Article

de Cillia, R., Reisigl, M., & Wodak, R. (1999). The discursive construction of national identity. *Discourse and Society*, 10(2), 149–173. <http://dx.doi.org/10.1177/0957926599010002002>

c) Journal Article with DOI and More Than Seven Authors

Lal, H., Cunningham, A. L., Godeaux, O., Chlibek, R., Diez-Domingo, J., Hwang, S.-J. ... Heineman, T. C. (2015). Efficacy of an adjuvanted herpes zoster subunit vaccine in older adults. *New England Journal of Medicine*, 372, 2087–2096. <http://dx.doi.org/10.1056/NEJMoa1501184>

d) Journal Article from Web, without DOI

Sidani, S. (2003). Enhancing the evaluation of nursing care effectiveness. *Canadian Journal of Nursing Research*, 35(3), 26–38. Retrieved from <http://cjr.mcgill.ca>

e) Journal Article with DOI

Turner, S. J. (2010). Website statistics 2.0: Using Google Analytics to measure library website effectiveness. *Technical Services Quarterly*, 27, 261–278. <http://dx.doi.org/10.1080/07317131003765910>

f) Advance Online Publication

Smith, J. A. (2010). Citing advance online publication: A review. *Journal of Psychology*. Advance online publication. <http://dx.doi.org/10.1037/a45d7867>

g) Article in a Magazine

Henry, W. A., III. (1990, April 9). Making the grade in today's schools. *Time*, 135, 28–31.

Doctoral Dissertation, Master's Thesis, Presentation, Proceeding

a) Dissertation/Thesis from a Commercial Database

Van Brunt, D. (1997). *Networked consumer health information systems* (Doctoral dissertation). Available from ProQuest Dissertations and Theses database. (UMI No. 9943436)

b) Dissertation/Thesis from an Institutional Database

Yaylalı-Yıldız, B. (2014). *University campuses as places of potential publicness: Exploring the political, social and cultural practices in Ege University* (Doctoral dissertation). Retrieved from <http://library.iyte.edu.tr/hizli-erisim/iyte-tez-portali>

c) Dissertation/Thesis from Web

Tonta, Y. A. (1992). *An analysis of search failures in online library catalogs* (Doctoral dissertation, University of California, Berkeley). Retrieved from <http://yunus.hacettepe.edu.tr/~tonta/yayinlar/phd/ickapak.html>

d) Dissertation/Thesis abstracted in Dissertations Abstracts International

Appelbaum, L. G. (2005). Three studies of human information processing: Texture amplification, motion representation, and figure-ground segregation. *Dissertation Abstracts International: Section B. Sciences and Engineering*, 65(10), 5428.

e) Symposium Contribution

Krinsky-McHale, S. J., Zigman, W. B., & Silverman, W. (2012, August). Are neuropsychiatric symptoms markers of prodromal Alzheimer's disease in adults with Down syndrome? In W. B. Zigman (Chair), *Predictors of mild cognitive impairment, dementia, and mortality in adults with Down syndrome*. Symposium conducted at the meeting of the American Psychological Association, Orlando, FL.

f) Conference Paper Abstract Retrieved Online

Liu, S. (2005, May). *Defending against business crises with the help of intelligent agent based early warning solutions*. Paper presented at the Seventh International Conference on Enterprise Information Systems, Miami, FL. Abstract retrieved from http://www.iccis.org/iccis2005/abstracts_2005.htm

g) Conference Paper - In Regularly Published Proceedings and Retrieved Online

Herculano-Houzel, S., Collins, C. E., Wong, P., Kaas, J. H., & Lent, R. (2008). The basic nonuniformity of the cerebral cortex. *Proceedings of the National Academy of Sciences*, 105, 12593–12598. <http://dx.doi.org/10.1073/pnas.0805417105>

h) Proceeding in Book Form

INFORMATION FOR AUTHORS

Parsons, O. A., Pryzwansky, W. B., Weinstein, D. J., & Wiens, A. N. (1995). Taxonomy for psychology. In J. N. Reich, H. Sands, & A. N. Wiens (Eds.), *Education and training beyond the doctoral degree: Proceedings of the American Psychological Association National Conference on Postdoctoral Education and Training in Psychology* (pp. 45–50). Washington, DC: American Psychological Association.

i) Paper Presentation

Nguyen, C. A. (2012, August). *Humor and deception in advertising: When laughter may not be the best medicine*. Paper presented at the meeting of the American Psychological Association, Orlando, FL.

Other Sources

a) Newspaper Article

Browne, R. (2010, March 21). This brainless patient is no dummy. *Sydney Morning Herald*, 45.

b) Newspaper Article with no Author

New drug appears to sharply cut risk of death from heart failure. (1993, July 15). *The Washington Post*, p. A12.

c) Web Page/Blog Post

Bordwell, D. (2013, June 18). David Koepp: Making the world movie-sized [Web log post]. Retrieved from <http://www.davidbordwell.net/blog/page/27/>

d) Online Encyclopedia/Dictionary

Ignition. (1989). In *Oxford English online dictionary* (2nd ed.). Retrieved from <http://dictionary.oed.com>

Marcoux, A. (2008). Business ethics. In E. N. Zalta (Ed.). *The Stanford encyclopedia of philosophy*. Retrieved from <http://plato.stanford.edu/entries/ethics-business/>

e) Podcast

Dunning, B. (Producer). (2011, January 12). *inFact: Conspiracy theories* [Video podcast]. Retrieved from <http://itunes.apple.com/>

f) Single Episode in a Television Series

Egan, D. (Writer), & Alexander, J. (Director). (2005). Failure to communicate. [Television series episode]. In D. Shore (Executive producer), *House*; New York, NY: Fox Broadcasting.

g) Music

Fuchs, G. (2004). Light the menorah. On *Eight nights of Hanukkah* [CD]. Brick, NJ: Kid Kosher.

SUBMISSION CHECKLIST

Ensure that the following items are present:

- Cover letter to the editor
 - The category of the manuscript
 - Confirming that “the paper is not under consideration for publication in another journal”.
 - Including disclosure of any commercial or financial involvement.
 - Confirming that last control for fluent English was done.
 - Confirming that journal policies detailed in Information for Authors have been reviewed.
 - Confirming that the references cited in the text and listed in the references section are in line with APA 6.
 - Copyright Transfer Form
 - Permission of previous published material if used in the present manuscript
 - Title page
 - The category of the manuscript
 - The title of the manuscript
 - All authors’ names and affiliations (institution, faculty/department, city, country),
 - e-mail addresses
 - Corresponding author’s email address, full postal address, telephone and fax number
 - ORCIDs of all authors.
 - Main Manuscript Document
 - The title of the manuscript
 - Abstract (150-250 words)
 - Key words: 3 words
 - Grant support (if exists)
 - Conflict of interest (if exists)
 - Acknowledgement (if exists)
 - References
 - All tables, illustrations (figures) (including title, explanation, captions)
-

TELİF HAKKI DEVİR FORMU / COPYRIGHT TRANSFER FORM



Acta INFOLOGICA

Istanbul University
İstanbul Üniversitesi

Copyright Agreement Form
Telif Hakkı Anlaşması Formu

Responsible/Corresponding Author <i>Sorumlu Yazar</i>	
Title of Manuscript <i>Makalenin Başlığı</i>	
Acceptance date <i>Kabul Tarihi</i>	
List of authors <i>Yazarların Listesi</i>	

Sıra No	Name - Surname Adı-Soyadı	E-mail E-Posta	Signature İmza	Date Tarih
1				
2				
3				
4				
5				

Manuscript Type (Research Article, Review, Short communication, etc.) <i>Makalenin türü (Araştırma makalesi, Derleme, Kısa bildiri, v.b.)</i>
Responsible/Corresponding Author: <i>Sorumlu Yazar:</i>

University/company/institution	<i>Çalıştığı kurum</i>	
Address	<i>Posta adresi</i>	
E-mail	<i>E-posta</i>	
Phone; mobile phone	<i>Telefon no; GSM no</i>	

The author(s) agrees that:
The manuscript submitted is his/her/their own original work, and has not been plagiarized from any prior work, all authors participated in the work in a substantive way, and are prepared to take public responsibility for the work, all authors have seen and approved the manuscript as submitted, the manuscript has not been published and is not being submitted or considered for publication elsewhere, the text, illustrations, and any other materials included in the manuscript do not infringe upon any existing copyright or other rights of anyone. İSTANBUL UNIVERSITY will publish the content under Creative Commons Attribution-NonCommercial 4.0 International (CC BY-NC 4.0) license that gives permission to copy and redistribute the material in any medium or format other than commercial purposes as well as remix, transform and build upon the material by providing appropriate credit to the original work. The Contributor(s) or, if applicable the Contributor's Employer, retain(s) all proprietary rights in addition to copyright, patent rights; to use, free of charge, all parts of this article for the author's future works in books, lectures, classroom teaching or oral presentations, the right to reproduce the article for their own purposes provided the copies are not offered for sale. All materials related to manuscripts, accepted or rejected, including photographs, original figures etc., will be kept by İSTANBUL UNIVERSITY for one year following the editor's decision. These materials will then be destroyed. I/We indemnify İSTANBUL UNIVERSITY and the Editors of the Journals, and hold them harmless from any loss, expense or damage occasioned by a claim or suit by a third party for copyright infringement, or any suit arising out of any breach of the foregoing warranties as a result of publication of my/our article. I/We also warrant that the article contains no libelous or unlawful statements, and does not contain material or instructions that might cause harm or injury. This Copyright Agreement Form must be signed/ratified by all authors. Separate copies of the form (completed in full) may be submitted by authors located at different institutions; however, all signatures must be original and authenticated.

Yazar(lar) aşağıdaki hususları kabul eder
Sunulan makalenin yazar(lar)ın orijinal çalışması olduğunu ve intihal yapmadıklarını, Tüm yazarların bu çalışmaya aslı olarak katılmış olduklarını ve bu çalışma için her türlü sorumluluğu aldıklarını, Tüm yazarların sunulan makalenin son halini gördüklerini ve onayladıklarını, Makalenin başka bir yerde basılmadığını veya basılmak için sunulmadığını, Makalede bulunan metnin, şekillerin ve dokümanların diğer şahıslara ait olan Telif Haklarını ihlal etmediğini kabul ve taahhüt ederler. İSTANBUL ÜNİVERSİTESİ'nin bu fikri eseri, Creative Commons Atf-GayriTicari 4.0 Uluslararası (CC BY-NC 4.0) lisansı ile yayınlamasına izin verirler. Creative Commons Atf-GayriTicari 4.0 Uluslararası (CC BY-NC 4.0) lisansı, eserin ticari kullanım dışında her boyut ve formatta paylaşılmasına, kopyalanmasına, çoğaltılmasına ve orijinal esere uygun şekilde atıfta bulunmak kaydıyla yeniden düzenleme, dönüştürme ve eserin üzerine inşa etme dâhil adapte edilmesine izin verir. Yazar(lar)ın veya varsa yazar(lar)ın işvereninin telif dâhil patent hakları, yazar(lar)ın gelecekte kitaplarında veya diğer çalışmalarında makalenin tümünü ücret ödemeksizin kullanma hakkı makaleyi satmamak koşuluyla kendi amaçları için çoğaltma hakkı gibi fikri mülkiyet hakları saklıdır. Yayımlanan veya yayıma kabul edilmeyen makalelerle ilgili dokümanlar (fotoğraf, orijinal şekil vb.) karar tarihinden başlamak üzere bir yıl süreyle İSTANBUL ÜNİVERSİTESİ'nce saklanır ve bu sürenin sonunda imha edilir. Ben/Biz, telif hakkı ihlali nedeniyle üçüncü şahıslarla vuku bulacak hak talebi veya açılacak davalarda İSTANBUL ÜNİVERSİTESİ ve Dergi Editörlerinin hiçbir sorumluluğunun olmadığını, tüm sorumluluğun yazarlara ait olduğunu taahhüt ederim/ederiz. Ayrıca Ben/Biz makalede hiçbir suç unsuru veya kanuna aykırı ifade bulunmadığını, araştırma yapılırken kanuna aykırı herhangi bir malzeme ve yöntem kullanılmadığını taahhüt ederim/ederiz. Bu Telif Hakkı Anlaşması Formu tüm yazarlar tarafından imzalanmalıdır/onaylanmalıdır. Form farklı kurumlarda bulunan yazarlar tarafından ayrı kopyalar halinde doldurularak sunulabilir. Ancak, tüm imzaların orijinal veya kanıtlanabilir şekilde onaylı olması gerekir.

Responsible/Corresponding Author; <i>Sorumlu Yazar;</i>	Signature / İmza	Date / Tarih
	/...../.....

