

Yıl:2022

Cilt:6

Sayı:2

Year:2022

Vol:6

No:2

**UYBİSBBD**

ULUSLARARASI YÖNETİM BİLİŞİM SİSTEMLERİ  
VE  
BİLGİSAYAR BİLİMLERİ DERGİSİ

ULUSLARARASI INTERNATIONAL JOURNAL OF  
YÖNETİM MANAGEMENT  
BİLİŞİM SİSTEMLERİ INFORMATION SYSTEMS  
VE AND  
BİLGİSAYAR BİLİMLERİ DERGİSİ COMPUTER SCIENCE

Cilt: 6 • Sayı: 2 • Aralık 2022  
Vol: 6 • No: 2 • December 2022

e-ISSN: 2618 - 5954

**ULUSLARARASI YÖNETİM BİLİŞİM SİSTEMLERİ  
VE  
BİLGİSAYAR BİLİMLERİ DERGİSİ**

**INTERNATIONAL JOURNAL OF MANAGEMENT INFORMATION SYSTEMS  
AND  
COMPUTER SCIENCE**

---

Cilt: 6 • Sayı: 2 • Aralık 2022  
Vol: 6 • No: 2 • December 2022

e-ISSN: 2618-5954

**E-mail :** [ybsbb.info@gmail.com](mailto:ybsbb.info@gmail.com)

**Web :** [dergipark.gov.tr/uybisbbd](http://dergipark.gov.tr/uybisbbd)

UYBİSBBD, uluslararası hakemli, uluslararası indeksli, açık erişimli bilimsel bir dergidir



---

**ULUSLARARASI YÖNETİM BİLİŞİM SİSTEMLERİ VE BİLGİSAYAR BİLİMLERİ  
DERGİSİ**  
**INTERNATIONAL JOURNAL OF MANAGEMENT INFORMATION SYSTEMS  
AND COMPUTER SCIENCE**

---

**Kurucu (Founder)**

Dr. Öğr. Üyesi Adem KORKMAZ

**Baş Editör (Editor-in-Chief)**

Dr. Öğr. Üyesi Selma BÜYÜKGÖZE

**Editörler (Editors)**

Prof. Dr. Aysun COŞKUN

Dr. Öğr. Üyesi Mustafa Mikail ÖZÇİLOĞLU

Dr. Öğr. Üyesi Tarık TALAN

Dr. Öğr. Üyesi Adem KORKMAZ

Dr. Öğr. Üyesi Selma BÜYÜKGÖZE

**Yayın Kurulu (Editorial Board)**

Prof. Dr. Florentin SMARANDACHE	(University of New Mexico, USA)
Prof. Dr. Aysun COŞKUN	(Gazi Üniversitesi)
Dr. Öğr. Üyesi Mustafa Mikail ÖZÇİLOĞLU	(Kilis 7 Aralık Üniversitesi)
Dr. Öğr. Üyesi Ayşe ÇİÇEK KORKMAZ	(Bandırma Onyediy Eylül Üniversitesi)
Dr. Öğr. Üyesi Tarık TALAN	(Gaziantep İslam Bilim ve Teknoloji Üniversitesi)
Dr. Öğr. Üyesi Adem KORKMAZ	(Bandırma Onyediy Eylül Üniversitesi)
Dr. Bogdan PATRUT	(Alexandru Ioan Cuza University of Iasi, Romania)
Dr. Iulian FURDU	(Vasile Alecsandri University of Bacau, Romania)
Dr. Sadiq HUSSAIN	(Dibrugarh University, India)
Dr. Svitlana ILNYTSKA	(National Aviation University, Ukraine)

**Danışma Kurulu (Advisory Board)**

Prof. Dr. Abdulkadir YILDIZ	(Kahramanmaraş Sütçü İmam Üniversitesi)
Prof. Dr. Aysun COŞKUN	(Gazi Üniversitesi)
Prof. Dr. Erdem UÇAR	(Trakya Üniversitesi)
Prof. Dr. Florentin SMARANDACHE	(University of New Mexico)
Prof. Dr. H. Mustafa PAKSOY	(Gaziantep Üniversitesi)
Prof. Dr. İsmail Rakıp KARAS	(Karabük Üniversitesi)
Prof. Dr. Sadettin PAKSOY	(Gaziantep Üniversitesi)
Prof. Dr. Sevinç GÜLSEÇEN	(İstanbul Üniversitesi)
Prof. Dr. Ülkü BAYKAL	(İstanbul Üniversitesi)
Prof. Dr. Yılmaz KILIÇASLAN	(Adnan Menderes Üniversitesi)
Prof. Dr. Mustafa ŞEKELİ	(Kahramanmaraş Sütçü İmam Üniversitesi)
Prof. Dr. Yusuf Ekrem AKBAŞ	(Adıyaman Üniversitesi)
Doç. Dr. Ercan BULUŞ	(Tekirdağ Namık Kemal Üniversitesi)
Doç. Dr. Erdinç UZUN	(Tekirdağ Namık Kemal Üniversitesi)
Doç. Dr. İlhan UMUT	(Trakya Üniversitesi)

**Adres (Address)**

Bandırma Onyediy Eylül Üniversitesi, Gönen Meslek Yüksekokulu  
10900 Balıkesir / TÜRKİYE

**E-mail :** [ybsbb.info@gmail.com](mailto:ybsbb.info@gmail.com)

**Web :** <https://dergipark.org.tr/tr/pub/uybisbbd>

---

**Bu Sayının Hakemleri**

---

**Doç. Dr. Veli BATDI**  
(Gaziantep Üniversitesi)

**Dr. Öğr. Üyesi Ebru Tomris AYDOĞAN**  
(Yeditepe Üniversitesi)

**Doç. Dr. Selahattin KOŞUNALP**  
(Bandırma Onyedli Eylül Üniversitesi)

**Dr. Öğr. Üyesi Fatih Çağatay BAZ**  
(Osmaniye Korkut Ata Üniversitesi)

**Doç. Dr. Özüm EĞİLMEZ**  
(Bilecik Şeyh Edebali Üniversitesi)

**Dr. Öğr. Üyesi Mustafa Mikail ÖZÇİLOĞLU**  
(Kilis 7 Aralık Üniversitesi)

**Doç. Dr. Halil İbrahim CEBECİ**  
(Sakarya Üniversitesi)

**Dr. Öğr. Üyesi Cemal AKTÜRK**  
(Gaziantep İslam Bilim ve Teknoloji Üniversitesi)

**Dr. Öğr. Üyesi Durmuş ÖZDEMİR**  
(Dumlupınar Üniversitesi)

**Dr. Öğr. Üyesi İsa AVCI**  
(Karabük Üniversitesi)

**Doç. Dr. Aşkın DEMİRAĞ**  
(Yeditepe Üniversitesi)

**Dr. Durmuş KOÇ**  
(Uşak Üniversitesi)

**Dr. Öğr. Üyesi Selma BÜYÜKGÖZE**  
(Kırklareli Üniversitesi)

**Dr. Tarık İÇTEN**  
(Gazi Üniversitesi)

---

## YAYIN POLİTİKASI

Uluslararası Yönetim Bilişim Sistemleri ve Bilgisayar Bilimleri Dergisi yılda iki kez Haziran ve Aralık aylarında yayınlanan uluslararası hakemli bir dergidir. Dergide yer alan yazılar kaynak gösterilmeksizin kısmen ya da tamamen iktibas edilemez. Bu dergide yayınlanan çalışmaların bilim ve dil sorumluluğu yazarlarına aittir.

Dergimize gönderilen çalışmalar, alanında uzman iki ayrı hakem tarafından incelendikten sonra uygun görülenler yayınlanmaktadır. Yazım kurallarına ilişkin bilgilere dergimizin web adresinde yer verilmiştir. Bu derginin tüm hakları saklıdır. Önceden yazılı izin almaksızın hiçbir iletişim ve kopyalama sistemi kullanılarak yeniden kopyalanamaz, çoğaltılamaz ve satılamaz.

---

International Journal of Management Information Systems and Computer Science is an international peer-reviewed journal which is published two times a year in June and December. The articles cannot be cited partly or entirely without showing resources. The responsibility about scientific and grammatical issues is belong to authors.

The papers sent to the journal are reviewed by two referees and after their approval, they will be sent to edit before being published. Writing & Publishing Policies can be found in the journal's website. All rights reserved. No part of this publication may be reproduced, stored or introduced into a retrieval system without prior written permission.

## Makaleler / Articles

### **Veri Madenciliği: Makine Öğrenme Algoritmaları ile Türkiye'nin İşsizlik Oranı Tahminini Etkileyen Faktörlerin Tespit Edilmesi**

*Data Mining: Identifying Factors Affecting Turkey's Unemployment Rate Estimation with Machine Learning Algorithms*

Makale Türü: Araştırma Makalesi / Paper Type: Research Paper

**İncilay YILDIZ 78-91**

### **Covid-19 Kısıtlamaları Sırasında Duyguların Müzikal Parametrelerle İfade Edilmesi: Filipinler Spotify Verileri Üzerine Bir Duygu Analizi**

*The Expression of Emotions Through Musical Parameters During the Covid-19 Restrictions: A Sentiment Analysis on Philippines Spotify Data*

Makale Türü: Araştırma Makalesi / Paper Type: Research Paper

**Ahmet YÜCEL 92-116**

### **Siber Güvenlik Uzmanın Çok Kriterli Karar Verme Yöntemleri ile Seçilmesi**

*Selecting A Cyber Security Specialist with Multi-Criteria Decision-Making Methods*

Makale Türü: Araştırma Makalesi / Paper Type: Research Paper

**Rabia YUMUŞAK & Tamer EREN 117-130**

### **Sanal Gerçekliği ve Artırılmış Gerçekliği Bir Sonraki Seviyeye Taşımak: Karma Gerçeklik ile Yapay Zeka**

*Taking Virtual Reality and Augmented Reality to The Next Level: Artificial Intelligence with Mixed Reality*

Makale Türü: Derleme Makalesi / Paper Type: Review Paper

**Ahmet EFE 131-148**

### **Yazılım Sektöründe Kişilik Tiplerinin Bir Uzman Sistem Kullanılarak Belirlenmesi: Pilot Çalışma**

*Determination Of Personality Types in Software Industry by Using an Expert System: A Pilot Study*

Makale Türü: Araştırma Makalesi / Paper Type: Research Paper

**Emine KURUOĞLU & Çağla OZEN & Mustafa KAMASAK 149-161**

### **Covid-19 Pandemisi Döneminde Öğrencilerin Uzaktan Eğitimde Yaşadıkları Sorunların Kelime Bulutu ve Gizli Anlamsal Dizinleme Yöntemleriyle Tespit Edilmesi**

*Determination of the Problems Experienced by Students in Distance Education During the Covid-19 Pandemic Period with Word Cloud and Latent Semantic Indexing Methods*

Makale Türü: Araştırma Makalesi / Paper Type: Research Paper

**Halit IRMAK & Zümrüt ECEVİT SATI 162-171**

# Veri Madenciliği: Makine Öğrenme Algoritmaları ile Türkiye'nin İşsizlik Oranı Tahminini Etkileyen Faktörlerin Tespit Edilmesi

\*\*\*

## Data Mining: Identifying Factors Affecting Turkey's Unemployment Rate Estimation with Machine Learning Algorithms

DOI:10.33461/uybisbbd.1129013

İncilay YILDIZ<sup>1</sup> 

### Öz

#### Makale Bilgileri

##### Makale Türü:

Araştırma Makalesi

##### Geliş Tarihi:

10.06.2022

##### Kabul Tarihi:

12.09.2022

©2022 UYBİSBBD  
Tüm hakları saklıdır.



*İşsizlik oranı, bir ülkenin ekonomik verimini ölçen en önemli değişkenlerden biridir. İş gücü ve verimi toplumun gelişmesi açısından çok önemli olup, işsizlik oranının düşürülmesi ve istihdam sağlanması bir devletin büyümesi açısından kilit öneme sahiptir. İşsizlik oranının tahmin edilmesi alınacak önlemler ve planlamalar açısından önem arz etmektedir. Bu çalışmada, işsizlik oranının tahmini için Türkiye İstatistik Kurumunun (TÜİK) İllerde Yaşam Endeksi verilerinden ve TÜİK tarafından oluşturulan veri gruplarından yararlanılmıştır. Tahmini etkileyen değişkenlerin tespiti için veri madenciliği sürecinden ve makine öğrenme algoritmalarından yararlanılmıştır. Bu çalışma kapsamında, işsizlik oranının tahmininde hangi tür veri gruplarının daha etkili olduğu ve hangi makine öğrenme yöntemlerinin daha etkin olduğu sorgulanmıştır. Elde edilen sonuçlar tahmin duyarlılık analizi kapsamında incelenmiş ve hata oranlarına göre yorumlanmıştır. Elde edilen sonuçlara göre işsizlik oranı tahmininde, Genel Regresyon Sinir Ağı (GRNN) yöntemi Destek Vektör Makineleri (SVM) ve Gen İfadesi Programlama (GEP) yöntemlerine göre daha etkin bir yöntemdir. Ayrıca, eğitim konulu veri grubunun diğer veri gruplarından daha etkili olduğu gözlemlenmiştir.*

**Anahtar Kelimeler:** Veri Madenciliği, Makine Öğrenmesi, Tahmin, İşsizlik Oranı.

#### Article Info

##### Paper Type:

Research Paper

##### Received:

10.06.2022

##### Accepted:

12.09.2022

©2022 UYBİSBBD  
All rights reserved.



### Abstract

*The unemployment rate is one of the most important variables that measure the economic efficiency of a country. Labor force and efficiency are very important for the development of society, reducing the unemployment rate and providing employment are key to the growth of a state. Estimating the unemployment rate is important in terms of measures and plans to be taken. In this study, for the estimation of the unemployment rate, data from the Life Index in Provinces of the Turkish Statistical Institute (TUIK) and data groups created by TUIK were used. Data mining process and machine learning algorithms were used to determine the variables affecting the estimation. Within the scope of this study, it has been questioned which types of data groups are more effective in estimating the unemployment rate and which machine learning methods are more effective. Obtained results were analyzed within the scope of estimation sensitivity analysis and interpreted according to error rates. According to the results obtained, General Regression Neural Network (GRNN) method is a more effective method than Support Vector Machines (SVM) and Gene Expression Programming (GEP) methods in estimating unemployment rate. In addition, it was observed that the education-themed data group was more effective than the other data groups.*

**Keywords:** Data Mining, Machine Learning, Prediction, Unemployment Rate.

**Atıf/ to Cite (APA):** Yıldız İ., (2022). Veri Madenciliği: Makine Öğrenme Algoritmaları ile Türkiye'nin İşsizlik Oranı Tahminini Etkileyen Faktörlerin Tespit Edilmesi. Uluslararası Yönetim Bilişim Sistemleri ve Bilgisayar Bilimleri Dergisi, 6(2), 78-91

<sup>1</sup> Arş. Gör., Yönetim Bilişim Sistemleri, Uygulamalı Bilimler Fakültesi, Altınbaş Üniversitesi, İstanbul, Türkiye, incilay.yildiz@altinbas.edu.tr.



## 1. GİRİŞ

Günümüzde büyük hacimlerdeki veriler bankalar, işletmeler, Türkiye İstatistik Kurumu (TÜİK) ve bazı kuruluşlar tarafından depolanmaktadır. Bu veriler içindeki değerli bilgileri ortaya çıkarmak ya da yeni bir bilgi keşfetmek işletme karlılığı açısından oldukça önemlidir. Ayrıca, hedeflenen bir planlamanın yapılması için elde edilen bilgiler yol gösterici olarak kullanılabilir. Bu verilerden en doğru şekilde faydalanmak için, üzerinde doğru analizlerin yapılması gerekmektedir.

Verileri veri tabanlarında saklayan kurumların en önemli amaçlarından biri ham verileri bilgiye dönüştürmektir. Veriyi bilgiye dönüştürme işlemi veri madenciliği olarak adlandırılmaktadır. Veri toplama araçları ve veri tabanı teknolojilerindeki gelişmeler, bilgi depolarında çok miktarda bilginin depolanmasını ve çözümlenmesini gerektirmektedir (Dener vd., 2009).

Tahmin (Öngörü) gelecekte gerçekleşecek olan olayların önceden saptanması, hesaplanması ya da kestirimidir. Bir tahmin çalışmasının duyarlılığı, yapılan tahminlerle gerçek değerler arasındaki sapmaların analizi sonucu ortaya çıkmaktadır. Duyarlılık analizi gerçek değerler bilindikten sonra yapılabilir. Bu çalışmada TÜİK veri havuzu kullanılarak tahmin çalışması yapılmıştır.

Veri madenciliği ve uygulamaları tahminleme çalışmalarında en sık faydalanılan çalışma alanlarından biridir. Veri madenciliği uygulamasının verimini arttırmak için, çalışılan verilerin çeşitliliği, veri madenciliği algoritmasının etkinliği, sonuçların yararlılık düzeyi, kesinlik ve anlamlılık unsurlarının sağlanması, elde edilen örüntülerin çeşitli biçimlerde gösterimi, gizliliğin ve veri güvenliği özelliklerinin sağlanması gereklidir. Veri madenciliği bilgi keşfi sürecinin bir parçası şeklinde düşünülebilir (Dener vd., 2009).

Bir veri madenciliği sürecinin çeşitli fonksiyonları vardır. Bu çalışmada ilişkilendirme, regresyon ve sınıflandırma fonksiyonlardan yararlanılmıştır.

İlişkilendirme: bir veri tabanındaki verilerin mevcut ya da potansiyel ilişkilerini ortaya çıkarmayı hedefler (Gong, vd., 2012) İstatistikler ve apriori algoritmaları örnek olarak gösterilebilir (Ngai, vd., 2009).

Sınıflandırma: veri tabanı kayıtlarını belirli kriterlere göre önceden tanımlanmış sınıflara ayırmaktır. Bu sınıflarda verileri toplayarak tahmine dayalı modelleri keşfetmeye yarar (Chen ve Liu, 2004). Sinir ağları ve karar ağaçları örnek olarak gösterilebilir (Ngai, vd., 2009).

Regresyon, iki ya da daha çok nicel değişken arasındaki ilişkiyi ölçmek için kullanılan analiz metodudur. Sık görülen model eğilimlerini yakalamak için kullanılır. Doğrusal regresyon ve lojistik regresyon analizi örnek olarak gösterilebilir.

Veri madenciliği kullanılarak yapılan tahmin çalışmalarında hem hızlı hem de yetkin bir metot olan makine öğrenme yöntemleri sıkça kullanılmıştır. Makine öğrenmesi (yapay öğrenme) bir problemi başka bir probleme ait girdi ve çıktı verilerine göre modelleyen bilgisayar algoritmaları olarak tanımlanabilir. Yapay zekânın ana kollarından biri olan makine öğrenmesi, birçok akademik alandaki çalışmalara entegre edilmiştir. Makine öğrenmesi, bilgisayar sisteminin kendi kendine anlayabilme ve öğrenebilme kabiliyetini sağlamayı amaçlar. Bilimsel hedefi, alternatif öğrenme mekanizmalarını araştırmak ve bulmaktır. Ayrıca makine öğrenmesi sisteminden, öğrenici ile iletişim kurması ve çalışma verilerinin sisteme uygulanabilmesi beklenmektedir. Sistemin kurulması ve test edilmesiyle bu süreçlerin etkinliği ve limitleri belirlenmektedir (Carbonell ve Gil, 1987).

Bu çalışmada pazar araştırması, iş gücü araştırması, insan kaynakları alanı ve eğitim araştırmalarında kullanılabilecek bir veri olan Türkiye'deki işsizlik oranının, tahminini etkileyen değişkenlerin analizi makine öğrenme algoritmalarından faydalanılarak yapılmıştır. İşsizlik oranı tahmininde kullanılan bağımlı değişken TÜİK verilerinden elde edilen "işsizlik oranı" değeridir. Bu çalışmada, işsizlik oranı tahmininde etkin olan girdi değişkenlerini (bağımsız değişkenleri) saptamak ve etkin olan makine öğrenmesi yöntemini tespit etmek hedeflenmiştir.



## 2. LİTERATÜR TARAMASI

Literatür araştırması kapsamında 2015-2020 yılları arasında Türkiye'deki işsizlik oranının tahmini yapan çalışmalar incelenmiştir.

İçen ve Günay (2015), Türkiye'deki işsizlik oranı tahmini için LINGO programını kullanmışlardır. Analiz sonucunda bulanık regresyon modeli kullanılarak daha güvenilir tahminler elde edildiği gözlemlenmiştir. Çalışmada kullanılan bulanık model parametrelerinin önemi bulanık hipotez testi ile test edilmiştir. Bulanık regresyon analizi doğrusal programlama yaklaşımı ile yapılmış olup farklı h düzeyleri ile çalışılmıştır. Verilen amaç fonksiyonu bulanık katsayıların toplam yayılımlarını minimize etmek için kullanılmış ve regresyon katsayılarının merkez ve yayılım değerleri elde edilmiştir. Elde edilen model parametrelerinin önemli olup olmadığı Buckley (2004) çalışmasında önerilen bulanık hipotez testi ile test edilmiştir

Gülenay Chadwick ve Şengül (2015) Google tarama sonuçlarının verilerini kullanarak Türkiye'de üç aylık bir gecikme ile yayınlanan "aylık tarım dışı işsizlik" dönem içi tahminini geliştirip geliştirilmediği araştırılmıştır. Çalışmanın yöntemi olarak, lineer tahmin modelleri ve Bayes tabanlı model ortalaması kullanılmıştır. Veri analizinde Ocak 2005 ve Kasım 2011 tarihleri arası baz alınmıştır. Elde edilen sonuçlara göre tarım dışı işsizlik oranı dönem içi tahminlerini iyileştirmektedir. Tahmin modellerinin kare hataların ortalamalarının karekökü karşılaştırıldığında en iyi performansı gösteren modelin Google tarama sonuçlarının ana bileşenlerini içermekte olduğu tespit edilmiştir. Google tarama verilerini içeren model örneklem içinde %47,7, örneklem dışında ise %38,4 daha doğru tahminler verdiği gözlemlenmiştir. Google tarama sonuçları verilerini içermeyen en iyi model örneklem içinde %34,1, örneklem dışında ise %29,4 daha doğru tahminler verdiği gözlemlenmiştir.

Çiççi (2016) çalışmasında, Markov zincirlerini stokastik süreç temelli bir tahminleme aracı olarak kullanmıştır. Bu çalışmada, Markov zincirleri yöntemi kullanılarak Türkiye'nin 2016 yılı kurumsal olmayan işsizlik oranı tahminlemesi yapılmıştır. Veri seti olarak, TÜİK'in yayınladığı 1996 ile 2014 yıllarına ait kurumsal olmayan işsizlik verileri kullanılmıştır. Elde edilen sonuçlara göre 2016 yılında Türkiye'de işsizlik oranı %94 olasılıkla 6,5 puanın üstünde bir değer alacağı, %79 olasılıkla 9,9 puanın altında bir değer alacağı, %80 olasılıkla 10 puan altı bir değer olacağı ve %88 olasılıkla 11 puan altı bir değer olacağı öngörülmüştür. Genel anlamda 2016 yılı işsizlik oranının tek haneli bir değer alacağı tahmin edilmektedir. Bu değerlerin 6,5 ile 9,9 aralığında kalması beklenmektedir.

Yolcu ve Baş (2016) çalışmasında işsizlik oranını ve işgücüne katılım oranını tahmin etmiştir. Polonya ve Türkiye'de farklı bulanık zaman serileri yöntemleriyle işsizlik oranları analiz edilmiştir. Tüm analizler sonucunda, bulanık zaman serileri tahmin modellerinin işsizlik oranı ve işgücüne katılım oranı gibi zaman serilerinin tahmininde üstün tahmin performansına sahip olduğu ve bu tür tahminler için etkili bir tahmin aracı olarak kullanılabilmesi belirlenmiştir. Ayrıca, önerilen bulanık zaman serisi yöntemleri geliştirildiğinde daha da iyi tahmin sonuçları verebilecekleri söylenmiştir.

Tüzemen ve Yıldız (2017) Türkiye işsizlik oranını, Toplamsal ve Çarpımsal-Mevsimsel Holt-Winters yöntemleri ile tahmin edilmiştir. Çalışmada merkez bankasının veri setinden yararlanılmıştır. Veriler, 01/2005-10/2016 dönemi Türkiye işsizlik oranlarından oluşturulmuştur. Yapılan analizlerin performans değerlendirmeleri hata kareler ortalaması ve ortalama mutlak yüzde hata değerleri ile değerlendirilmiştir. Yapılan analizde, Toplamsal-Mevsimsel HW yöntemi ile  $MSE = 0,073711$  ve  $MAPE=0,021695$ , Çarpımsal-Mevsimsel HW yöntemi ile de  $MSE = 0,8618929$  ve  $MAPE = 0,02217202$  değerleri saptanmıştır. Ayrıca, Toplamsal-Mevsimsel HW yöntemi ile yapılan analizde işsizlik oranlarının 2017 yılında artarak devam edeceği öngörülmüştür.

Yıldırım ve Başgeçmez (2017) bu çalışmada, tek değişkenli zaman serisi verilerini tahmin etmede yararlı olan teknikleri kullanmışlardır. Bu çalışmanın amacı, Türkiye İşsizlik Oranını tahmin eden en etkin yöntemi saptamaktır. Türkiye İşsizlik oranına ilişkin veriler, 1996-2015 tarihleri arası baz alınarak TÜİK'den alınmıştır. Öncelikle yıllık işsizlik oranlarının karakterizasyonu için otokorelasyon analizi kullanılmıştır. Otokorelasyon analizi sonuçlarına göre Türkiye işsizlik oranları

arasında korelasyon olmadığı gösterilmiştir. Ayrıca bu serilerin mevsimselliğinin olduğu da tespit edilmiştir. Bu nedenle en iyi tahmin modellerinin basit hareketli ortalama, ağırlıklı hareketli ortalama, basit üstel düzeltme, uyarlanabilir yanıt oranı üstel düzeltme, Holt Lineer Tahmin ve Holt-Winter trend ve mevsimsellik yöntemi olduğuna karar verilmiştir. Tüm modeller için tahmine en uygun tekniğin seçilmesinde kullanılan Theil'in U istatistikleri hesaplanmıştır. Bu istatistiklere göre, Türkiye işsizlik oranı için en uygun yöntem Holt'un Doğrusal Tahmin yöntemidir. Bu tekniğe göre, Türkiye işsizlik oranınının 10,5 olacağı öngörülmüştür. Elde edilen sonuçlara göre, Uyarlanabilir Tepki Oranı Üssel Düzeltme (ARRES) yönteminin basit hareketli ortalamalar için 4.016, ağırlıklı hareketli ortalamalar için 2.653, çift hareketli ortalamalar için 4.940 olan en düşük Ortalama Kareler Hatası değerini hesapladığı için Türkiye işsizlik oranı için en uygun yöntem olduğu sonucuna varılabileceği belirtilmiştir.

Gil-Alina vd. (2019) Türkiye işsizlik oranı serilerini uzun hafızalı modeller ve kısmi bütünleşik teknikler kullanarak incelemiştir. Çalışmanın amacı, literatürde var olan hipotezlerden (NAIRU, yapısalıcı, kalıcılık ve histeresis) hangi hipotezin incelenen veriler için daha uygun olduğunu belirlemek amacıyla zaman serilerinin entegrasyon sırasını tahmin etmektir. Çalışmada, 1988 yılı 3.çeyrek ile 2013 yılı 4.çeyrek arası Türkiye işsizlik oranı verileri kullanılmıştır. Sonuçlarımız, Türkiye'de işsizliğin yüksek oranda kalıcı olduğunu ve vakaların çoğunda entegrasyon sıralarının 1'e eşit veya daha yüksek olduğunu göstermektedir. Ampirik sonuçlarından önemli politika çıkarımları ortaya çıkmıştır; işgücü ve makroekonomik politikaların Türkiye'deki işsizlik oranları üzerinde uzun süreli etkilerinin olması beklenmektedir.

Bağcı (2020) Türkiye'deki işsizlik oranını tahmin etmek için GM (1.1) modeli kullanıldıktan sonra, Markov zincirleri GM (1.1) modeline entegre edilerek Gri Markov modeli oluşturulmuş ve bu yöntemde işsizlik oranı yeniden tahmin edilmiştir. Başlangıç olarak, Gri sistem modellerinden biri olan GM (1,1) modeli kullanılarak işsizlik oranlarına ilişkin istatistikler hesaplanmıştır. Daha sonra her bir gözlem değerinin hata oranları hesaplanarak belirlenmiştir. Hem hata oranları hem de verilerin orijinal durumu kullanılarak Markov analizi yapmak için durum geçişleri matrisi elde edilmiştir. Tüm gözlemler, bu matris ve durum geçişleri yardımıyla Gri Markov analizi kullanılarak yeniden tahmin edilmiştir. İki GM arasında karşılaştırma yapabilmek için (1,1) yeni tahmin değerlerine ait hata oranları da hesaplanmıştır ve Gri Markov analizinin daha doğru tahmin sonuçları ürettiği tespit edilmiştir.

Bu çalışmada, literatürden farklı olarak tahmin modelleri oluşturulurken veri madenciliğinin sınıflandırma fonksiyonundan yararlanılmıştır. TÜİK tarafından yapılan sınıflandırma yardımıyla tahmin modelleri oluşturulmuş ve işsizlik oranı tahmininde hangi özelliklere sahip girdi değişkenlerinin yani hangi veri gruplarının daha etkin olduğu araştırılmıştır. Ayrıca işsizlik oranı 3 farklı (GRNN, SVM, GEP) makine öğrenme algoritması ile tahmin edilerek yöntemlerin performansları karşılaştırılmıştır. Regresyon ve sinir ağı tabanlı bir yöntem olan GRNN yönteminin, SVM ve GEP yöntemlerinden daha etkin sonuçlar verdiği tespit edilmiştir.

### 3. VERİ SETİ VE YÖNTEM

#### 3.1 Veri Seti

Bu çalışmada kullanılan veri seti TÜİK tarafından ilk defa 2016 yılında yayımlanan (2015 verilerini içermektedir) İllerde Yaşam Endeksi (İYE) çalışmasından alınmıştır. TÜİK tarafından veri seti hakkında verilen bilgiler şu şekildedir; "Veri seti, endeks objektif ve subjektif göstergeleri içermektedir. Subjektif göstergelerinin temel veri kaynağı il düzeyinde uygulanması planlanan Yaşam Memnuniyeti Araştırmasının (YMA) sonuçlarıdır. Objektif göstergeler ise il düzeyinde TÜİK verileri ve diğer kurum ya da kuruluşlardan derlenen verilerden oluşmaktadır." (Resmi İstatistik, 2016).

YMA verileri, 4500 cevaplı örnek hanede 9397 cevaplı bireyle gerçekleştirilen anket çalışmasından elde edilmiştir. Anketin amacı TÜİK tarafından "Türkiye'deki bireylerin öznel mutluluk algısı, sağlık, sosyal güvenlik, örgün eğitim, çalışma hayatı, gelir, kişisel güvenlik ve adalet

hizmetleri, kişisel gelişim gibi temel yaşam alanlarındaki memnuniyetleri ölçmek ve bunların zaman içindeki değişimi takip etmek” şeklinde açıklanmıştır (TÜİK, 2016a).

Veri anketinin kapsamı, “Türkiye Cumhuriyeti sınırları içinde bulunan hanelerde yaşayan 18 ve daha yukarı yaştaki T.C. vatandaşları ile yabancı uyruklu kişilerdir. Araştırmada, kurumsal nüfus (üniversite öğrenci yurdu, huzurevi, bakımevi, ceza ve tutukevi, ıslahevi, yetiştirme yurdu, askeri birlik ve kışla vb.) kapsam dışı tutulmaktadır.” Anket sonuçları Türkiye’nin 81 ilini kapsamakta ve oranlar illere göre ayrılmaktadır (TÜİK, 2016b).

Bu çalışmada TÜİK tarafından yayımlanan İYE göstergelerinin; işsizlik oranı % (IO) iller (ID), Kişi başına düşen tasarruf mevduatı TL (KDT), Orta ve üstü gelir grubundaki hanelerin oranı % (OUG), Temel ihtiyaçlarını karşılayamadığını beyan eden hanelerin oranı % (TIK), Hekim başına düşen müracaat sayısı (HDM), Sağlığından memnuniyet oranı % (SMO), Kamunun sağlık hizmetlerinden memnuniyet oranı % (KSM), Okul öncesi eğitimde (3-5 yaş) net okullaşma oranı % (OOE), TEOG sistemi yerleştirmeye esas puan ortalaması (TEOG), YGS puan ortalaması (YGS), Fakülte veya yüksekokul mezunlarının oranı % (FYM), Kamunun eğitim hizmetlerinden memnuniyet oranı (%) (EHM), Cinayet oranı (CO), Ölümlü ve yaralanmalı trafik kazası sayısı (TKO), Gece yalnız yürürken kendini güvende hissedenenlerin oranı (GGO), Kamunun asayiş hizmetlerinden memnuniyet oranı (KAM), Mahalli idareler seçimlerine katılım oranı (SKO), Siyasi partilere üyelik oranı (SPU), Sendika/dernek faaliyetleri ile ilgili olanların oranı (SDU), İnternet abone sayısı (IAS), Kanalizasyon ve şebeke suyuna erişim oranı % (SEO), Havalimanına erişim oranı (HEO), Belediyenin toplu taşıma hizmetlerinden memnuniyet oranı (BHM), Sinema ve tiyatro seyirci sayısı (STS), Bin kişi başına düşen alışveriş merkezi alanı (m<sup>2</sup>) (AMO), Sosyal ilişkilerinden memnuniyet oranı (SIM) ve Sosyal hayatından memnuniyet oranı (SHM) verileri kullanılmıştır. IO’yu tahmin etmek için IO bağımlı değişken çıktı verisi olarak, diğer 27 bağımsız değişken ise girdi verisi olarak kullanılmıştır.

TÜİK tarafından hazırlanan İYE çalışmasının verileri niteliklerine göre sınıflara ayrılmıştır. Bu sınıflandırma sonucunda oluşan veri grupları ile tahmin modelleri oluşturulmuştur. Oluşturulan 8 tahmin modeli ve içerdiği değişkenler Tablo 1’de sunulmuştur. Veri setinin istatistiği Tablo 2’de sunulmuştur.

**Tablo 1.** Tahmin Modellerinin Veri Grupları

Model No	Model İsmi	Model Değişkenleri
1	Gelir ve Servet	ID IO KDT OUG TIK
2	Sağlık	ID HDM SMO KSM
3	Eğitim	ID OOE TEOG YGS FYM EHM
4	Güvenlik	ID CO TKO GGO KAM
5	Sivil Katılım	ID SKO SPU SDU
6	Alt Yapı Hizmetlerine Erişim	ID IAS SEO HEO BHM
7	Sosyal Yaşam	ID STS AMO SIM SHM
8	Tüm Değişkenleri İçeren Model	ID IO KDT OUG TIK HDM SMO KSM OOE TEOG YGS FYM EHM CO TKO GGO KAM SKO SPU SDU IAS SEO HEO BHM STS AMO SIM SHM

**Tablo 2.** Veri Setinin İstatistik Değerleri

Değişkenler	Min	Max	Ortalama	Standart Sapma
IO	4,2	23,4	8,801	3,878

ID	1	81	41	23,527
KDT	616	18131	4342,033	2821,418
OUG	16,27	58,91	34,382	8,633
TIK	32,78	74,95	50,95	10,19
HDM	2763	8067	5834,383	1245,116
SMO	59,15	80,76	71,997	4,448
KSM	54,55	89,13	77,468	7,305
OOE	23,53	53,16	35,275	6,247
TEOG	215,33	337,99	295,936	27,672
YGS	178,6	207,95	197,64	5,512
FYM	8,56	22,65	13,118	2,348
EHM	48,18	88,89	74,089	8,46
CO	4,47	69,34	25,499	11,118
TKO	0,71	4,59	2,436	0,776
GGO	45,1	87,23	67,598	9,242
KAM	58,88	94,86	84,236	7,161
SKO	77,1	93,1	88,157	3,194
SPU	12,44	34,73	21,243	3,935
SDU	3,54	22,08	6,735	2,582
IAS	2,16	17,66	8,68	3,866
SEO	31,11	100	74,37	15,827
HEO	0	9874,8	669,867	1571,486
BHM	23,46	78,81	58,394	11,581
STS	0,29	147,44	45,353	34,005
AMO	0	284,01	69,54	68,29
SIM	78,23	96,16	88,831	3,793
SHM	21,5	80,88	54,265	11,355

### 3.2 Yöntem

#### 3.2.1. Genel Regresyon Sinir Ağı

Örnek sayısı arttıkça optimal regresyon yüzeyine yakınsama prensibi ile çalışan GRNN yönetimi, hızlı öğrenmeye dayalı bir algoritmadır. Seyrek veriye sahip bir set üzerinde çalışıldığında GRNN yöntemi avantajlıdır. Örnek olarak, veri hiper yüzeyi, her yeni numune noktasının eklenmesiyle artarak, karmaşıklaşır. GRNN yöntemi, verileri düşük ve yüksek olarak ikiye bölmek için, hiper uzay aralarında geçişe izin verir (Dtreg, 2021a).

GRNN ağı dört katmanlı bir yapıdan oluşmaktadır. Giriş katmanı, her tahmin değişkeni için giriş katmanında bir nöron bulunmaktadır. N (kategori sayısı), Kategorik değişkenler ile çalışırken N-1 nöronları kullanılır. “Giriş nöronları, medyanı çıkararak ve çeyrekler arası aralığa bölerek değerlerin aralığını standartlaştırır. Sonrasında bu değerleri gizli katmandaki nöronların her birine iletir” (Dtreg, 2021a). Gizli katman, eğitim veri kümesindeki her durum için bir nörona sahiptir. Nöron, hedef değerle birlikte durum için öngörücü değişkenlerin değerlerini depolamaktadır. Gizli bir nöron, girdi katmanından gelen girdi değerlerinin x vektörü ile sunulduğunda, test senaryosu, nöronun merkez noktasından Öklid mesafesini hesaplar ve ardından sigma değerlerini kullanarak RBF çekirdek işlevini uygular. Elde edilen sonuç değeri, desen katmanındaki nöronlara iletilir. Toplama katmanı, GRNN ağları için, toplama katmanında yalnızca iki nöron vardır; “Bir nöron payda toplama birimi, diğeri pay toplama birimidir. Payda toplama birimi, gizli nöronların her birinden gelen ağırlık değerlerini toplar. Pay toplama birimi, her bir gizli nöron için gerçek hedef değerle çarpılan ağırlık değerlerini toplar.” Karar katmanı, pay toplama biriminde biriken değeri payda toplama birimindeki değere böler ve sonucu tahmin edilen hedef değer olarak kullanır (Dtreg, 2021a).

### 3.2.2. Destek Vektör Makineleri

Temel olarak, SVM yönteminin arkasındaki ana fikir, doğrusal olarak ayrılabilir modeller için sınıflandırma için kullanılacak bir optimal hiper düzlemin oluşturulmasıdır. Optimal hiper düzlem; “hiper düzlemin marjını, yani hiper düzlemden her modelin en yakın noktasına olan mesafeyi maksimize eden kalıpları sınıflandırmak için hiper düzlemler kümesinden seçilen bir hiper düzlemi” şeklinde tanımlanmıştır. SVM yönteminin temel amacı, marjı maksimize etmektir, böylece verilen desenleri doğru bir şekilde sınıflandırabilir, yani marj boyutu büyüdükçe desenleri daha doğru sınıflandırabilir. Hiper düzlem,  $aX + bY = C$  ile formülize edilir (Pradhan, 2012).

SVM tabanlı modeller üç farklı hattan oluşmaktadır.  $W.x - b = 0$  çizgisi, ayırma marjı veya marjinal çizgi olarak bilinir.  $W.x - b = 1$  ve  $w.x - b = -1$  hatları, kenar boşluğu çizgisinin her iki tarafındaki çizgilerdir. Bu üç çizgi birlikte, verilen desenleri ayıran hiper düzlemi oluşturur ve hiper düzlemin kenarlarında yatan desene destek vektörleri denir. Kenar boşluğu çizgisi ile hiper düzlemin kenarları arasındaki dikey mesafe, kenar boşluğu olarak bilinir. Doğru sınıflandırma için SVM'nin amaçlarından biri, daha iyi sınıflandırma için bu marjı maksimize etmektir. Kenar boşluğunun değeri veya dikey mesafe ne kadar büyükse, sınıflandırma süreci o kadar iyidir ve dolayısıyla hata oluşumunu en aza indirir (Pradhan, 2012).

SVM tabanlı modeller, klasik çok katmanlı algılayıcı sinir ağlarına benzerdir. Bir çekirdek işlevi kullanan SVM yöntemi, polinom, radyal temelli işlev ve çok katmanlı algılayıcı sınıflandırıcılar gibi ağırlıklarının konveks olmayan bir denklem yerine doğrusal kısıtlamalarla ikinci dereceden bir programlama problemini çözen alternatif bir eğitim yöntemi kullanır (Dtreg, 2021b).

### 3.2.3 Gen İfadesi Programlama

GEP, ifade-mutasyon problemine zarif ve etkili bir çözüm olarak, 1999 yılında Cândida Ferreira tarafından önerilmiştir. Ferreira, ifadeleri kodlamak için, çok çeşitli mutasyon ve melezleme tekniklerinin hızlı uygulanmasına izin verirken ortaya çıkan ifadenin her zaman sözdizimsel olarak geçerli olacağını garanti eden bir sistem tasarlamış ve bu yaklaşıma GEP adı verilmiştir (Dtreg, 2021c).

GEP yöntemi biyolojik evrimi taklit eden bir bilgisayar algoritmasıdır. Karar ağaçları, sinir ağları ve polinom yapılar dahil olmak üzere birçok farklı model türü oluşturmak için kullanılabilir. DTREG programında uygulanan GEP türü, sembolik regresyondur. Sembolik matematiksel veya

mantıksal bir işlev oluşturduğu için bu şekilde adlandırılmaktadır. Girdi verileri fonksiyonlara ve değişkenlere ek olarak, ifadeler sabitler içerebilir (Dtreg, 2021c). DTREG tarafından GEP şu şekilde tarif edilmektedir; “Sabit uzunluktaki doğrusal kromozomlarda kodlanmış bilgisayar programlarını geliştiren tam teşekküllü bir genotip / fenotip sistemidir”

### 3.3 Tahmin Duyarlılık Analizi

Bir tahmin yönetimin duyarlılığı, tahmin edilen değer ile gerçek değerler arasındaki sapmaların analizi sonucu ortaya çıkmaktadır. Tahmin duyarlılık analizinde, hata tahmin ölçüleri ve bunları hesaplamaya yarayan yöntemlerden yararlanılmaktadır (Kobu, 2017). Hata değeri, çalışmadan elde edilen sonuçlar ile gerçek sonuçlar arasındaki farkın alınmasıyla hesaplanır. En düşük hataya sahip tahmini yapan yöntem, en iyi sonuçları veren yöntemdir. Bu çalışmada, tahmin sonuçlarını değerlendirmek için duyarlılık analizinin hata tahmin ölçülerinden faydalanılmıştır.

#### 3.3.1 Kolerasyon Katsayısının Karesi ( $R^2$ )

Kolerasyon katsayısı, “iki sayısal ölçüm arasında doğrusal bir ilişki olup olmadığını, varsa bu ilişkinin yönünün ve değerinin ne olduğunu belirlemek için kullanılan bir istatistiksel yöntem” olarak tanımlanmaktadır. Korelasyon katsayısı negatif ise iki değişken arasında ters ilişki vardır, pozitif ise değişkenlerden biri artarken diğeri de artmaktadır  $R^2$  Formülü (1)'de gösterilmiştir (Kobu, 2017).

$$R^2 = 1 - \frac{\sum(Y - Y')^2}{\sum(Y - \bar{Y})^2} \quad (1)$$

Bulgularda  $R^2$  değeri yorumlanırken;  $\%80 < R^2$  çok iyi,  $\%60 < R^2 < \%80$  iyi,  $\%40 < R^2 < \%60$  orta,  $\%20 < R^2 < \%40$  kabul edilebilir,  $R^2 < \%20$  hatalı olarak değerlendirilmiştir.

#### 3.3.2. Hata kareler ortalamasının karekökü (Root Mean Square Error - RMSE)

RMSE, kareler toplamının bölünmüş tahmin sayısının kare kökü olarak tanımlanır. Gerçekte gözlenen değerler ile tahmin modeli tarafından öngörülen değerler arasındaki farkı ölçmektedir. Dolayısıyla düşük RMSE değeri, modelin daha doğru tahmin ürettiğini göstermektedir. RMSE Formülü (2)'de gösterilmiştir (Kobu, 2017).

$$RMSE = \sqrt{\frac{1}{n} \sum_{i=1}^n (Y - Y')^2} \quad (2)$$

Bulgularda RMSE değeri yorumlanırken düşük RMSE değerine sahip (hata değeri daha düşük olan) modelin diğer modelden daha doğru (daha iyi) tahmin gerçekleştirdiği düşünülmüştür.

Denklemler 1'de,  $Y$  gerçek IO,  $Y'$  tahmin edilen IO,  $\bar{Y}$  gerçek IO değerlerinin ortalamasını,  $n$  ise test edilen alt küme örneklerinin sayısını göstermektedir.

#### 3.3.3. Ortalama Mutlak Yüzde Hata (Mean Absolute Percentage Error - MAPE)

MAPE, genellikle hata değerlerinin birim değerleri farklılık gösteriyorsa tercih edilmektedir. MAPE ile farklı birim değerlerin karşılaştırılmasında ortaya çıkabilecek dezavantajlar elimine edilmektedir. MAPE Formülü (3)'de gösterilmiştir (Kobu, 2017).

$$MAPE = 100 * \frac{\sum_{i=1}^n \frac{|Y-Y'|}{Y}}{n} \quad (3)$$

Denklemlerde  $Y$  gerçek IO,  $Y'$  tahmin edilen IO,  $\bar{Y}$  gerçek IO değerlerinin ortalamasını,  $n$  ise test edilen alt küme örneklerinin sayısını göstermektedir.

Bulgularda MAPE değeri yorumlanırken;  $MAPE < \%10$  çok iyi,  $\%10 < MAPE < \%20$  iyi,  $\%20 < MAPE < \%50$  kabul edilebilir,  $\%50 < MAPE$  hatalı olarak değerlendirilmiştir.

Bu çalışmada kullanılan makine öğrenme yöntemleri DTREG programı yardımıyla çalıştırılmıştır. GRNN, SVM ve GEP algoritmaları 10 katlı çapraz doğrulama yapılarak tahmin modellerine uygulanmıştır. Tablo 3’de makine öğrenme yöntemlerinde kullanılan parametreler ve değerleri verilmiştir. Bu çalışmada veri setinin hem girdi hem çıktı değerleri ağa tanıtılmış önce eğitim sonra test sonuçları hesaplatılmıştır. Elde edilen sonuçlar Tablo 4’de gösterilmiştir. Sonuçlar incelendiğinde sınıflandırma sonuçlarına göre düzenlenmiş 8 farklı tahmin modeli görülmektedir. Bu tahmin modellerinin GRNN, SVM ve GEP yöntemlerinden elde edilen RMSE değerleri bulunmaktadır.

Şekil 1’de ise GRNN tabanlı modeller ile elde edilen sonuçların SVM ve GEP tabanlı modeller ile elde edilen sonuçlara göre RMSE değerlerinin yüzde iyileştirme oranlarının grafiği verilmiştir.

Tablo 5’de GRNN yöntemine göre veri gruplarının RMSE oranları sıralanmıştır.

Tablo 6’de ise GRNN yöntemine göre veri gruplarının etkinliği sunulmuştur. Analiz yapılırken RMSE oranının yanı sıra  $R^2$ , MAPE ve programın çalışma süresi hesaplanmıştır.

**Tablo 3.** Yöntemlerin Parametreleri

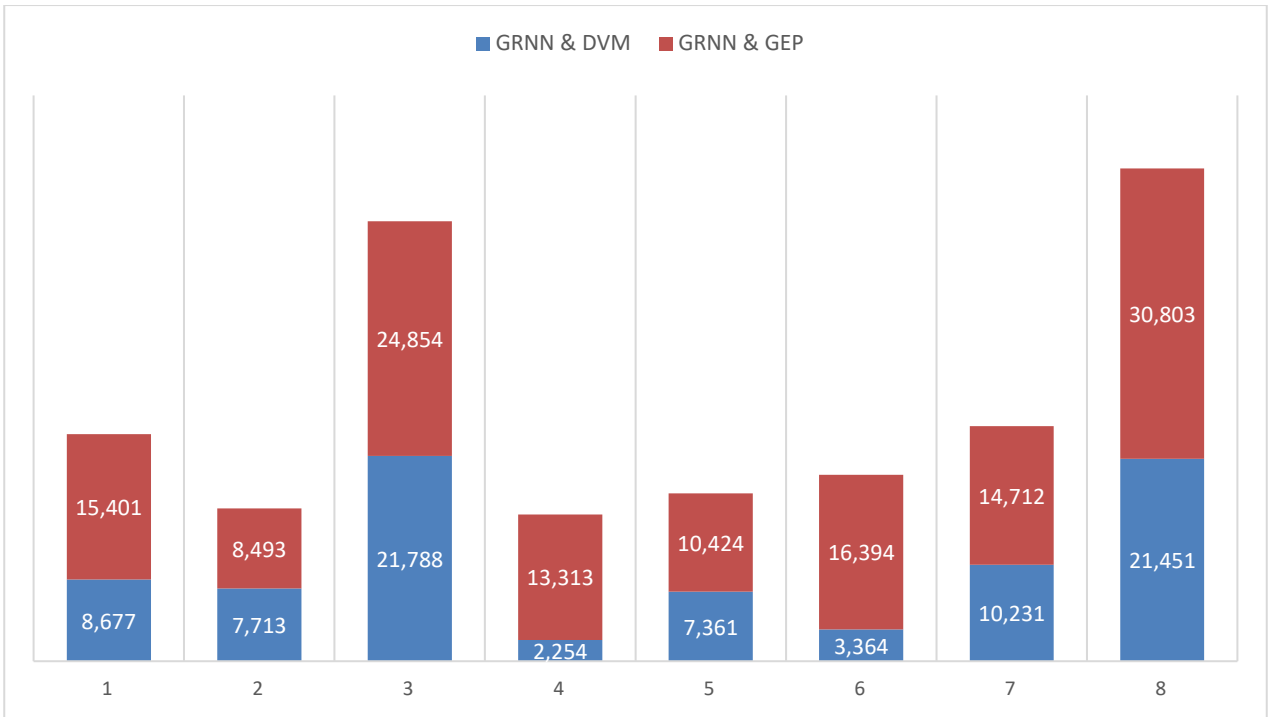
Yöntemler	Parametreler	Değerler
GRNN	Sigma	[0,0001 - 10]
	Arama Sayısı	20
	Çekirdek Fonksiyonu	Gaussian
SVM	Cost (C)	[0,1 - 5000]
	Gamma ( $\gamma$ )	[0.001 - 50]
	Çekirdek Fonksiyonu	RBF
GEP	Populasyon Değeri	[10-60]
	En Yüksek Gen Sayısı	2000
	Kromozom Başına Gen Sayısı	4

**Tablo 4.** Yöntemlerin Karşılaştırmalı Sonuçları

No	Veri Grubu	GRNN		DVM		GEP	
		Eğitim	Test	Eğitim	Test	Eğitim	Test
		RMSE	RMSE	RMSE	RMSE	RMSE	RMSE



1	Gelir ve Servet	2,927	3,143	2,770	3,442	3,178	3,715
2	Sağlık	1,423	3,438	3,668	3,725	3,436	3,757
3	Eğitim	1,485	2,614	2,401	3,342	3,482	3,478
4	Güvenlik	2,397	3,123	3,122	3,195	3,215	3,602
5	Sivil Katılım	2,634	3,315	2,954	3,578	3,554	3,700
6	Alt Yapı Hizmetlerine Erişim	1,186	2,831	1,740	2,929	2,958	3,386
7	Sosyal Yaşam	1,303	2,845	2,550	3,169	3,019	3,336
8	Tüm Değişkenleri İçeren Model	0,257	2,112	1,832	2,688	2,340	3,052



Şekil 1. Yöntemlerin Karşılaştırmalı Sonuçları

Tablo 5. GRNN Yöntemine Göre Veri Gruplarının Etkinlik Sıralaması

Model No	Veri Grubu	Değişkenler	Eğitim RMSE	Test RMSE
8	Tüm Değişkenleri İçeren Model	ID IO KDT OUG TIK HDM SMO KSM OOE TEOG YGS FYM EHM CO TKO GGO KAM SKO SPU SDU IAS SEO HEO BHM STS AMO SIM SHM	0,257	2,112
3	Eğitim	ID OOE TEOG YGS FYM EHM	1,485	2,614

6	Alt Yapı Hizmetlerine Erişim	ID IAS SEO HEO BHM	1,186	2,831
7	Sosyal Yaşam	ID STS AMO SIM SHM	1,303	2,845
4	Güvenlik	ID CO TKO GGO KAM	2,397	3,123
1	Gelir ve Servet	ID IO KDT OUG TIK	2,927	3,143
5	Sivil Katılım	ID SKO SPU SDU	2,634	3,315
2	Sağlık	ID HDM SMO KSM	1,423	3,438

**Tablo 6.** GRNN Tabanlı Modellerin Veri Grupları Bazında Detaylı Analizi

GRNN Yöntemi	Eğitim			Test			
	R <sup>2</sup>	RMSE	MAPE	R <sup>2</sup>	RMSE	MAPE	ZAMAN
Gelir ve Servet	0,4234	2,927	23,841	0,335	3,143	24,801	00:01:36
Sağlık	0,86364	1,423	9,344	0,2044	3,438	24,671	00:01:07
Eğitim	0,85151	1,485	11,698	0,5401	2,614	20,126	00:01:53
Güvenlik	0,61324	2,397	18,459	0,34356	3,123	22,241	00:01:37
Sivil Katılım	0,5331	2,634	22,007	0,26038	3,315	28,211	00:01:27
Alt Yapı Hizmetlerine Erişim	0,90532	1,186	9,044	0,46061	2,831	18,848	00:01:09
Sosyal Yaşam	0,88568	1,303	10,142	0,4551	2,845	22,548	00:01:13
Tüm Değişkenleri İçeren Model	0,99555	0,257	15,281	0,69984	2,112	16,932	00:04:16
Ortalama Sonuçlar	0,75893	1,701471	14,97687	0,412374	2,927368	22,2971	00:01:47

#### 4. BULGULAR

Tablo 4, 5, 6 ve Şekil 1 ile elde edilen bulgular şu şekildedir;

- Tablo 4 incelendiğinde, IO tahmininde tüm modeller için GRNN yönteminin en iyi (en düşük RMSE hata değerli) sonuçları verdiği ve SVM yönteminin ise GEP yönteminden daha iyi sonuçlar verdiği gözlemlenmiştir.
- Tablo 4 incelendiğinde, GRNN tabanlı modellerin ortalama RMSE değerleri SVM tabanlı modellere göre %10,355 & GEP tabanlı modellere göre %16,800 daha düşüktür.
- Tablo 4 incelendiğinde, Tüm yöntemler için en iyi tahmin modelinin tüm değişkenleri içeren 8.model olduğu gözlemlenmektedir

- Tablo 4 incelendiğinde, En iyi model olan 8.model (en düşük hata değerleri sahip) ile en fazla hata değerine sahip 2.model sağlık veri grubu arasındaki RMSE değerlerinin farkı (iyileşme oranı) 1,022'dir.
- Tablo 5 incelendiğinde GRNN yöntemine göre en iyi tahmin modelinin en düşük test RMSE değerine sahip olan 8 model olduğu gözlemlenmiştir.
- Tablo 5 incelendiğinde GRNN yöntemine göre en iyi tahmin model sıralamasının 8,3,6,7,4,1,5,2 olduğu gözlemlenmektedir. Buna göre IO tahmininde veri gruplarının etkinliği şu şekilde sıralanabilir; öncelikle eğitim, alt yapı hizmetleri, sosyal yaşam, güvenlik, gelir ve servet, sivil katılım ve son olarak sağlık veri grubu.
- Şekil 1'den elde edilen verilere göre en fazla 8.model ve 3.modelde en az 4.model ve 2.modelde iyileşme gözlemlenmiştir.
- Tablo 6 incelendiğinde,  $R^2$  değerinin ortalama %41 olarak görülmektedir. Bu oran orta düzeyde uyumluluk olarak yorumlanmıştır. En iyi sonucu veren tüm değişkenleri içeren modelin %69  $R^2$  değeri ile iyi düzeyde uyumluluğa sahip olduğu söylenebilir.
- Tablo 6 incelendiğinde ortalama MAPE değerinin %22 olması iyiye yakın kabul edilebilir hata oranının olduğunu. Tüm değişkenleri içeren modelin ise %16 hata değerine sahip olduğu yani iyi bir tahmin modeli olduğu ortaya konulmuştur.
- Tablo 6 incelendiğinde tahmin modellerinin GRNN yönteminde çalışma zamanlarının ortalama 1 dakika 47 saniye olduğu görülmektedir. Bu durum çok kısa sürede kabul edilebilir oranda sonuçlar elde edildiğini göstermektedir.

## 5. SONUÇ

Veri madenciliği, veri tabanında bulunan verilerin bilgiye dönüştürülmesi kullanılmasının yanı sıra, tahmin için uygun verilerin tespitinde de kullanılmaktadır. Bu çalışmada veri madenciliği yöntemlerinin sınıflandırma (TÜİK'in özelliklerine göre veri grupları), ilişkilendirme ve regresyon (makine öğrenme algoritmaları) fonksiyonlarından yararlanılmıştır.

Bu çalışmada, TÜİK'in İYE araştırmasının veri seti ve veri grupları kullanılarak 8 farklı tahmin modelleri geliştirilmiştir. 8 tahmin modeli GRNN, DVM ve GEP makine öğrenme yöntemleriyle analiz edilmiştir. Elde edilen sonuçların duyarlılığı RMSE oranı kullanılarak karşılaştırılmıştır. Tüm modeller için kullanılan yöntemden bağımsız olarak en iyi sonuçlar tüm değişkenleri içeren 8.modelden elde edilmiştir ve en düşük performans ise sağlık verilerini içeren 2.modelden elde edilmiştir. 8 tahmin modelinde de tutarlı olarak en iyi sonuçlar GRNN yönteminden elde edilmiştir. Ayrıca SVM yöntemi ile GEP yönteminden daha iyi sonuçlar elde edilmiştir. GRNN yönteminin performansı incelendiğinde tüm tahmin modellerinde orta derece kolerasyon ve kabul edilebilir MAPE doğruluk oranı olduğu gözlemlenmiştir. GRNN yöntemi ile 8.modelden elde edilen sonuçlarda ise iyi derecede kolerasyon ve doğruluk oranı gözlemlenmiştir.

Gelecek çalışmalarda, IO tahmininde kullanılan yapay öğrenme yöntemleri farklılaştırılabilir. Derin öğrenme yöntemleri ile sonuçlar analiz edilebilir. Ayrıca girdi değişkenlerinin sayısı artırılarak hata oranları düşürülebilir.

## KAYNAKÇA

Bağcı, B. (2020) "Gri Markov Modeli ile Türkiye'de İşsizlik Oranı Tahmini", Sosyal Güvenlik Dergisi, 10 (2) ss. 259-272

Buckley J.J, (2004) Fuzzy Statistics, Springer, Germany. ISBN 3-540-21084-9

- Carbonell, G.J., Gil, Y. (1987, Haziran) “Learning by Experimentation” Proceedings of the Fourth International Workshop on Machine Learning, 256-266.
- Chen S.Y., Liu X. (2004) “The contribution of data mining to information science” Journal of Information Science, vol 30, no 6. pp. 550–558
- Çiftçi, C. (2016) “Forecasting of Unemployment Rate For Turkey: Markov Chains Approach”, Eurasian Academy of Sciences Eurasian Business & Economics Journal, Vol. S2, 657-665, <http://dx.doi.org/10.17740/eas.econ.2016-MSEMP-140>
- Dener, M., Dörterler M., Orman A. (2009, Şubat)“Açık Kaynak Kodlu Veri Madenciliği Programları: WEKA’da Örnek Uygulama” Akademik Bilişim’09- XI. Akademik Bilişim Konferansı Bildirileri, Harran Üniversitesi, Şanlıurfa, ss. 787-796.
- DTREG, Gene Expression Programming [Çevrimiçi]. <https://www.dtreg.com/methodology/view/gene-expression-programming> Erişim: 2021a
- DTREG, General Regression Neural Networks [Çevrimiçi]. <https://www.dtreg.com/solution/probabilistic-and-general-regression-neural-networks> Erişim: 2021b
- DTREG, Support Vector Machines, [Çevrimiçi]. <https://www.dtreg.com/solution/support-vector-machines> Erişim: 2021c
- Gil-Alina L.A., Ozdemir Z.A., Tansel A. (2019) “Long Memory in Turkish Unemployment Rates”, Emerging Markets Finance and Trade (35) 201-217 (Published online: 27 Jul 2018), <https://doi.org/10.1080/1540496X.2018.1425837>
- Gong, L., Lei, H., & Zhu, L. (2012, May). The application of association rules of data mining in book-lending service. In 2012 9th International Conference on Fuzzy Systems and Knowledge Discovery (pp. 761-764). IEEE.
- Günay S., İcen D. (2015) “Türkiye’deki İşsizlik Oranının Bulanık Doğrusal Regresyon Analiziyle Tahmini”, İstatistikçiler Dergisi: İstatistik & Aktüerya, IDIA 8, 10-26
- Kobu, B. (2017) Talep Tahminlerinde Duyarlılık, Üretim Yönetimi, 18.baskı, Beta Basım Yayım Dağıtım, İstanbul, Türkiye, 120-125
- M. Gülenay Chadwick and G. Şengül, (2015) “Nowcasting The Unemployment Rate in Turkey: Let’s Ask Google”, Central Bank Review (15) 15-40, ISSN 1303-0701
- Ngai, E.W.T., Xiu, L., Chau, D.C.K. (2009) “Application of data mining techniques in customer relationship management: A literature review and classification” Expert Systems with Applications, 36, (2). 2592–2602
- Pradhan A., (2012) “Support Vector Machine-A Survey” International Journal of Emerging Technology and Advanced Engineering, 2 (8), 82-85
- Resmi İstatistik, İllerde Yaşam Endeksi, [Çevrimiçi]. Erişim: <http://www.resmiistatistik.gov.tr/detail/subject/illerde-yasam-endeksi/>
- Türkiye İstatistik Kurumu, TÜİK (2016a), Yaşam Memnuniyeti Araştırması amaç, [Çevrimiçi]. Erişim: ][https://tuikweb.tuik.gov.tr/MicroVeri/YMA\\_2016/metaveri/amac/index.html](https://tuikweb.tuik.gov.tr/MicroVeri/YMA_2016/metaveri/amac/index.html)
- Türkiye İstatistik Kurumu, TÜİK (2016b) Yaşam Memnuniyeti Araştırması kapsam [Çevrimiçi]. Erişim: [https://tuikweb.tuik.gov.tr/MicroVeri/YMA\\_2016/metaveri/kapsam/index.html](https://tuikweb.tuik.gov.tr/MicroVeri/YMA_2016/metaveri/kapsam/index.html)
- Tüzemen A., Yıldız Ç. (2018) “Holt Winters Tahminleme Yöntemlerinin Karşılaştırmalı Analizi: Türkiye İşsizlik Oranları Uygulaması” Atatürk Üniversitesi İktisadi ve İdari Bilimler Dergisi, 32, (1), 1-18

- Yıldırım H., Başeğmez H. (2017) “Analysis and forecast of Turkey unemployment rate”, Global Journal of Mathematical Analysis, vol 1, no 1. 11-15,. doi: 10.14419/gjma.v5i1.6841
- Yolcu U., Eren B. (2016) “The Forecasting Of Labour Force Participation And The Unemployment Rate In Poland And Turkey Using Fuzzy Time Series Methods Comparative” Economic Research, 19, (2), doi: 10.1515/er-2016-0010

# Covid-19 Kısıtlamaları Sırasında Duyguların Müzikal Parametrelerle İfade Edilmesi: Filipinler Spotify Verileri Üzerine Bir Duygu Analizi

\*\*\*

## The Expression of Emotions Through Musical Parameters During the Covid-19 Restrictions: A Sentiment Analysis on Philippines Spotify Data

DOI:10.33461/uybisbbd.1139568

Ahmet YÜCEL<sup>1</sup> 

### Öz

#### Makale Bilgileri

##### Makale Türü:

Araştırma Makalesi

##### Geliş Tarihi:

02.07.2022

##### Kabul Tarihi:

11.10.2022

©2022 UYBİSBBD  
Tüm hakları saklıdır.



Spotify, dinleyicilerin dinleme kalıplarına göre çok çeşitli çalma listeleri sunan dünyanın en büyük çevrimiçi müzik paylaşım platformudur. Bu makale, Filipinler verisi üzerine yapılan inceleme göre, müzik tercihinin duygusal durumla yüksek oranda ilişkili olduğunu ve müziğin pandemi sırasında bir duygu düzenleyici araç olduğunu ortaya koymaktadır. Popüler makine öğrenimi yöntemleri (sınıflandırma ve regresyon ağaçları, güçlendirilmiş ağaçlar, rastgele ormanlar, Destek Vektör Makineleri ve Yapay Sinir Ağları), pandeminin şiddetiyle orantılı olarak farklı zaman periyotlarında insanların müzik tercihlerini sınıflandırmak için, 5 katlı çapraz doğrulama ile birlikte kullanılmıştır. Algoritmaların öğrenme süreçlerinde günlük resmi covid-19 istatistikleri ve Spotify verileri ana değişkenler olarak kullanılmaktadır. SVM, %98,01 doğruluk oranı elde ederek ortalama doğruluk oranında diğer alternatif modeller içinde en iyi performansı göstermiştir. Ek olarak, ANN, özellikle tek bir modelde elde edilen doğruluk açısından diğer alternatiflerden daha iyi performans göstererek %99,30 doğruluk oranına ulaşmıştır. Ayrıca (azalan sırada) en büyük (mutlak) değişime sahip değişkenler *ST\_Instrumentalness* (%26,45), *ST\_Acousticness* (%19,03), *ST\_Liveness* (%16,11) ve *ST\_Valence* (%14,1)'dir. Pandemi kaynaklı stres ve konserlerin iptali göz önüne alındığında, *ST\_Valence* (müzikal pozitiflik) ve *ST\_Liveness* değişkenlerinin böyle bir oranda değişeceği sezgisel bir beklenti olmuştur. Sonuçlar, müzik tercihinin duygusal durumun önemli bir göstergesi olduğunu doğrulamaktadır.

#### Article Info

##### Paper Type:

Research Paper

##### Received:

02.07.2022

##### Accepted:

11.10.2022

©2022 UYBİSBBD  
All rights reserved.



**Anahtar Kelimeler:** Nicel Karar Yöntemleri, Veri Analitiği, Makine Öğrenimi, Duygu Analizi, Müzik.

### Abstract

Spotify is the world's largest online music streaming platform that offers a tremendous variety of playlists based on listeners' listening patterns. This paper proposes that music preference is highly associated with emotional state, and music is an emotion regulator tool during the pandemic in the Philippines. Well-known machine learning methods (i.e., classification and regression trees, boosted trees, random forests, Support Vector Machines, and Artificial Neural Networks) in combination with 5-fold cross-validation are used to classify periods in proportion to the severity of the pandemic and people's musical preferences. Daily official covid-19 statistics and Spotify data are used as main variables during the algorithms' learning processes. SVM outperformed the other alternatives in average accuracy rate by achieving a 98.01% accuracy rate. Additionally, ANN outperformed the other alternatives in terms of accuracy achieved specifically in a single model, achieving an accuracy rate of 99.30%. Moreover, the variables with the largest (absolute) change (in descending order) are *ST\_Instrumentalness* (26,45%), *ST\_Acousticness* (19,03%), *ST\_Liveness* (16,11%), and *ST\_Valence* (14,1%). Given pandemics-related stress and cancelation of concerts, it would be an intuitive expectation that the variables *ST\_Valence* (musical positivity) and *ST\_Liveness* would change at such a rate. The results confirm that musical preference is a significant indicator of emotional state.

**Keywords:** Quantitative Decision Methods, Data Analytics, Machine Learning, Sentiment Analysis, Music.

**Atıf/ to Cite (APA):** Yücel, A., (2022). The Expression of Emotions Through Musical Parameters During the Covid-19 Restrictions: A Sentiment Analysis on Philippines Spotify Data. International Journal of Management Information Systems and Computer Science, 6(2), 92-116

<sup>1</sup> Dr. Öğr. Üyesi, Bankacılık ve Finans Bölümü, Şereflikoçhisar Uygulamalı Bilimler Fakültesi, Ankara Yıldırım Beyazıt Üniversitesi, Ankara, Türkiye, ayucel@ybu.edu.tr.

## 1. INTRODUCTION

The new coronavirus disease (COVID-19), which first emerged in Wuhan, China, in December of 2019, has not only caused worldwide distress (Jena, et al. 2021) but also induced a big global worry and psychological shock (Giordano, et al. 2020). World Health Organization (WHO) declared it as a pandemic in January 2020 (Jena, et al. 2021).

Contagious diseases are always serious threats to humanity (Wang, et al. 2020). Because of the scant available medical options and the severity of symptoms that may cause to patients' death, COVID-19 is particularly dangerous. Also, some patients infected with COVID-19 may not be exhibiting any indication of disease but pass it on other people (Karasmanaki and Tsantopoulos 2021). Pyrexia, insufficiency of breathing, and cough, are among the most common somatic signs of the Covid-19 (Arslan, Yıldırım and Tanhan, et al. 2020) whereas psychological and neurological consequences of COVID-19 is not well-documented (Rahman, et al. 2020).

The COVID-19 pandemic is a strong warning that pandemics are repeating phenomena and it is probable that such calamities will repeat in the future (Donthu and Gustafsson 2020). The worry and dread about the disease have also caused to a societal stigmatization (Rahman, et al. 2020). The COVID-19 is not only an epidemiological catastrophe but also a health crisis by causing a vast variety of psychological issues such as anxiety, fear and confusion (Arslan, et al. 2021).

Daily social life contains substantial physical activities such as meeting with friends, participating concerts and intellectual events (Karasmanaki and Tsantopoulos 2021). However, daily life has been rearranged under the pandemic circumstances. Governments around the world have put into practice numerous protective policies to prevent the rapid boosting spread. All sports activities have been cancelled or suspended to reduce the spread of the virüs (Nicola, et al. 2020). The best intervention model to contain the disease was found to confine people in their homes (Sood 2020). All the governments have closed their country borders, restricted the movement of people (Donthu and Gustafsson 2020). These kind of limitations have caused to deterioration in economic courses (Jena, et al. 2021) and in the supply chains (Nicola, et al. 2020). The quarantine policies that have been performed by many countries impacted businesses, employment rates, and main services (Sharif, Aloui and Yarovaya 2020) and also have led to stock market instability (Bayraktar 2020) and bankruptcies for many popular corporations in different sectors (Donthu and Gustafsson 2020).

The present pandemic has had intense economic outcomes across the world (Donthu and Gustafsson 2020). Under the pandemic circumstances, the demand for manufactured commodities has decreased (Nicola, et al. 2020), investors have lost their interest in investing and started saving capital (Donthu and Gustafsson 2020) resulting a worldwide panic that led the global economy to a stop (Bayraktar 2020). The ambiguity and financial damage caused by covid-19 have caused the business sectors to become extremely unreliable (Bayraktar 2020). In a nutshell, the pandemic has not only caused health crises, but it has also created chaos in the all economies on a measure not observed since the Great Depression (Laing 2020).

The tourism sector is dealing with an unprecedented shock of cancellations and a dramatic reduction in demand. The food and farming sectors are also struggling with excessive demand because of panic-buying and stocking up on goods (Nicola, et al. 2020). Also global stock markets have been impacted seriously by oil price drop (Sharif, Aloui and Yarovaya 2020). These striking decrease in prices and production across all mining industry, demonstrate the present catastrophic situation in the industry (Laing 2020). The strong impact on the economy has caused fiscal insecurity and distress in the society which will implicitly cause health problems (Sood 2020).

Although there is very scarce data on the mental health course of past outbreak cases (Sood 2020), there is sufficient finding that negative psychological factors can cause substantial increase in inflammatory diseases (Wang, et al. 2020). The ambiguity of the pandemic is a main factor for psychological distress among people (Sood 2020). The psychological consequences of the distress might exceed the health consequences of the pandemic (Xie, et al. 2020). Temporary shutting down



restaurants, cafes, recreation centers, and the all similar public places has induced abrupt transformation on usual dietary and social activity models (Flanagan, et al. 2021). Under the quarantine circumstances, people are consuming more junk food and having less personal care (Donthu and Gustafsson 2020). Therefore covid-19 has the potential to lead non-contagious diseases like obesity (Flanagan, et al. 2021). In the study of Flanagan et al., the changes in daily dietary styles, physical exercises, sleep patterns, mental health, and sedentary lifestyle before and in the course of the COVID-19 was observed (Flanagan, et al. 2021).

The mental health effect of the pandemic have been proved to be much more severe for those who are prone to psychological problems (Simon, et al. 2021). Psychological health is negatively correlated with fear of death, and positively correlated with good health (Kasapoğlu 2020). The new form of daily life including restricted social activity, irregular sleeping hours, and social distancing concept have adversely affected the mental well-being (Chaturvedi, Vishwakarma and Singh 2021). As people stay at home for longer periods of time, the possibility of experiencing anxiety and claustrophobia is increasing (Sood 2020). In the study of Tull et al., findings show the significance of social interaction to mitigate adverse psychological effects of the COVID-19 (Tull, et al. 2020).

An unusually interesting process is going on, people are being fined for not staying at home (Donthu and Gustafsson 2020). All educational institutions have ceased face to face education and began giving online education (Arslan, et al. 2021). The findings show that the shutdown of educational institutions has been extremely effective prevention to slow down the spread of the virüs (Karasmanaki and Tsantopoulos 2021). These worldwide closure measures have impacted 1.4 billion students (UNESCO 2020). Students have become physically away from their classmates and their major social environment. Schools and colleges are gainful environments where students can enjoy the feeling of belongingness that is a crucial component for better education process and mental well-being. During the pandemic, a substantial increase has been observed in the number of mental problems among students (Arslan, et al. 2021).

The essential role of art in education and entertainment is undeniable, and it is a means of relaxation in times of economic difficulties. The need for artistic content has escalated during the lockdown. Economic indications show that art and culture will be one of the most damaged industries, and most likely one of the slowest to recuperate. Online access methods are highly preferred, as it is temporarily not possible to physically participate in artistic and cultural activities (Radermecker 2020). Under the lockdown conditions, there is a rapid increase in the usage of Internet and social media. Also, online shopping and (Donthu and Gustafsson 2020) gaming platforms have reached to an unprecedented growth in the number of users (Nicola, et al. 2020).

Depression is a extremely widespread psychiatric (Sakka and Juslin 2018) and chronic illness that can affect all age groups (Castillo-Pérez, et al. 2010). People with depression have trouble controlling their emotions. Music has always been utilized as an alternative healing tool in all societies (Martín, et al. 2021). Music listening is a beneficial activity for mitigating depression symptoms, promoting positive emotions and diminishing negative emotions (Sakka and Juslin 2018). According to reports from people, music has been a substantial support for coping with isolation, an effective tool for emotional balancing, and a cure for loneliness (Martín, et al. 2021). Aljanaki et al., put forth that there is adequate variety and expressiveness in music to transmit and stimuli emotions (Aljanaki, Wiering and Veltkamp 2016). Bogt et al. (2011) reveal the protective function of music towards stress and conclude that listening to music stimulates positive emotions and supports coping with stress and anxiety (Bogt, et al. 2011). The findings of Pothoulaki et al. (2008) show that music is an excellent mood stimulant which ensures that attention is not focused on stressful situations; the drop in anxiety level is especially evident when listening to music pleasant to the listener (Pothoulaki, et al. 2008). In the study of Juslin et al. (2008), negative moods such as irritation, anxiety or worry are not mostly experienced while listening to music (Juslin, et al. 2008).

Spiritual self-regulation methods give substantial benefits for social and educational performance, as well as for personal health (Martín, et al. 2021). Music has the potential to evoke

strong emotions in people (Aljanaki, Wiering and Veltkamp 2016). Also music can affect all brain activities connected with emotions (Koelsch, 2018). There are many studies in the field of neuropsychology that examine brain responses to musical parameters (Bresin and Friberg 2011). According to this studies, emotions that arise with music are real and can be extremely intense. In addition the vast majority of such music studies declare considerable transformation in the anterior hippocampal (Koelsch, 2018). Moreover, music affects the functioning of heart, the action of breathing, and blood pressure. Koelsch & Jäncke concluded that pulse and breathing rates are higher in reaction to energetic music compared with soothing music (Koelsch and Jäncke, Music and the heart 2015).

For different age groups, the favorite genres and streaming frequencies are different. The most popular genres are Pop, rock, and jazz (Gurgen 2016). People between the ages of 18-40 use music as a means of coping with loneliness (Martín, et al. 2021). There are several studies investigating the effect of music on the brain. For example, Mozart's Sonata has been proven to increase brain capacity by activating neural signal pathways. This is referred to as the "Mozart effect" (Castillo-Pérez, et al. 2010). Understanding music, which has been an important part of all cultures throughout human history, will make it possible to correctly understand the emotions aroused by music (Koelsch, 2018). Emotions evoked by music are usually the outcomes of dynamic emotional regulation (Eerola, et al. 2018). Music can be considered both as a way of expressing emotions and as a factor that causes emotional reactions in listeners (Eerola, et al. 2018). In general, the most common purpose of listening to music is to cultivate positive emotions (Sakka and Juslin 2018).

According to the results of a number of studies in the literature, music therapy can alleviate pain, regulate sleeping pattern, and reduce anxiety without medication (Giordano, et al. 2020). Giordano et al. proves the effectiveness of taking music therapy as a supplementary intervention to alleviate anxiety and improve comfort (Giordano, et al. 2020). In the study of Martín et al., participants were asked to use music to express their feelings and support some social difficulties such as loneliness. The findings verify that music is one of the most preferred methods to deal with loneliness (Martín, et al. 2021). According to the results of a meta-analysis, loneliness is an element of risk for all-cause mortality (Sood 2020). Also Castillo-Pérez et al. proposes that people with low or medium level depression can use music to improve the efficacy of psychological support. No side effects have been detected for music therapy, whereas music therapy can apply an extremely positive effect on psychological health. In the music therapy session, the depression rates of the music-listening group have decreased significantly (Castillo-Pérez, et al. 2010). Music-evoked happiness is provided by dopamine release (Eerola, et al. 2018). Also, music is very effective for mitigating cancer pain and (Castillo-Pérez, et al. 2010) muscle spasm (Faus, Matas and Elósegui 2019).

The convenience in exchanging information with others and the sense of social connection convince people to be involved in social media. Social media is a very prevalent platform among younger age groups (Arslan, et al. 2021). In a study on the subject, it has been revealed that the social media usage is related to the concepts of belongingness, self-expression, and receiving personal support (Arslan, et al. 2021). Even though social media provide people with so many benefits during the quarantine, it also has some potential harms such as misleading rumors or incorrect information (Sood 2020).

Providing instant access to countless music libraries, advances in digital technology are giving people an unprecedented variety of music (Anderson, et al. 2021). In order to receive the highest possible quality and speed service from the databases of musical sources, automatic evaluation and classification algorithms are required. The methods applied by musical content sharing systems are based on emotion-based analysis algorithms (Aljanaki, Wiering and Veltkamp 2016). Spotify is a Swedish online music streaming service and is described as the world's largest streaming platform with over a hundred million active users. According to user reviews, compared to competitors like Apple or Amazon music streaming services, Spotify offers more convenient and reliable service. Spotify's success is often seen as a consequence of its economic and cultural effect (Vonderau 2019).

Spotify is usually described as the “new radio” for the effect it has on hit musics or albums, and for its function in music consumption and music discovery. Spotify offers a growing variety of playlists based on listener-centered driven data (Prey, Valle and Zwerwer 2020). Prey et al. investigated the mediator function of Spotify on the music industry and how Spotify is changing the way we consume music (Prey, Valle and Zwerwer 2020). Vonderau examined Spotify as a social media corporation that operates at the intersection of technology, marketing, and music (Vonderau 2019).

People's preferences can be predicted with high accuracy thanks to computer-based algorithms. Music preferences are an important indicator of people's psychological states. By using online services, we can perform our daily activities, from social communication to information retrieval. Recent social and psychological research suggests that it is possible to successfully predict people's personality traits from digital records. The information obtained from this process is called a "digital footprint". Therefore, the music choices on Spotify can be considered as an important digital source of information about personality traits. Listening to music is a significant individual activity and therefore has the potential to contain qualified information regarding the personal preferences. Anderson et al. examined the correlation between personality and music listening patterns on Spotify. The findings show that the music listening pattern on Spotify provides significant information on the personality (Anderson, et al. 2021).

In this study, it was examined whether people's musical preferences changed under the Covid-19 pandemic conditions. Several musical parameters shared by Spotify and statistical data on the course of the pandemic have been brought together on the daily basis. Since this is a descriptive study, we do not have an obvious design of hypotheses; however, we do have a general assumption of the patterns between emotional reactions and music parameters. Our research revealed that there is a significant change in music listening habits during the pandemic.

## 2. EXPERIMENTAL STUDY

### 2.1. Data

The data used in this study is a combination of two data sets obtained from different sources. The first dataset has been provided by a *kaggle.com* member *JC Albert Peralta*. That is a *Spotifycharts.com* originated data shared within the scope of *Spotify Developers API*. The second dataset has been obtained from another similar data-sharing platform, *Our World in Data* (*ourworldindata.org*). The first dataset contains details about the top 200 most listened daily music tracks on *Spotify* between the dates 01.01.2017 and 15.01.2021 in the Philippines. Details of the variables are given below (Table 1).

**Table 1.** Description of variables (*Spotify*)

Variables	Type	Description
<b>Date</b>	<b>Ordinal</b>	The date on which the information of the track was collected
<b>Track Id</b>	<b>Nominal</b>	Unique <i>Spotify</i> id of the track
<b>Track Name</b>	<b>Text</b>	Name of the track
<b>Position</b>	<b>Numerical</b>	Chart rank (from 1 to 200) in terms of number of streams.
<b>Artist Name</b>	<b>Text</b>	Artist of the track
<b>Acousticness</b>	<b>Numerical</b>	The ratio of the acoustics of the track. It can take a value between 0 and 1.
<b>Danceability</b>	<b>Numerical</b>	Danceability ratio in terms of a combination of musical elements such as tempo, rhythm stability, beat strength, and general regularity. It can take a value between 0 and 1.
<b>Energy</b>	<b>Numerical</b>	Energy is a perceptual measure of intensity and activity. It can take a value between 0 and 1.
<b>Instrumentalness</b>	<b>Numerical</b>	It can take a value between 0 and 1. The closer the instrumentals value is to 1, the greater likelihood the track contains no vocal content.
<b>Key</b>	<b>Numerical</b>	The note or chord that the track is centered around. It can take a value between 0 and 11. If no key is detected, the value is -1.

<b>Liveness</b>	<b>Numerical</b>	The presence of an audience in the recording and the probability that the track was performed live are measured. It can take a value between 0 and 1.
<b>Loudness</b>	<b>Numerical</b>	Overall loudness of a track in decibels (dB). Values typically range between -60 and 0 dB.
<b>Mode</b>	<b>Numerical</b>	The type of scale from which its melodic content is derived. Major is represented by 1 and minor is 0.
<b>Speechiness</b>	<b>Numerical</b>	The ratio of the presence of spoken words in a track. It can take a value between 0 and 1.
<b>Tempo</b>	<b>Numerical</b>	The overall estimated tempo of a track in beats per minute (BPM).
<b>Valence</b>	<b>Numerical</b>	Describes the musical positiveness (e.g. happy, cheerful, euphoric) conveyed by a track. Tracks with low valence sound more negative (e.g. sad, depressed, angry). It can take a value between 0 and 1.
<b>Streams</b>	<b>Numerical</b>	Number of streams in the specified date

The second dataset contains daily covid-19 statistics in the Philippines between the dates 30.01.2020 when the first coronavirus case was detected in the Philippines and 15.01.2021. Details of the variables are given below (Table 2):

Table 2: Description of variables (Covid-19)

Variable	Type	Description
<b>date</b>	Ordinal	Date of the day when daily data on the pandemic were obtained.
<b>total_cases</b>	Numerical	The total number of coronavirus patients detected until the day of data collection.
<b>new_cases</b>	Numerical	The total number of coronavirus patients detected on the day the data was collected.
<b>new_cases_smoothed</b>	Numerical	Smoothed number of daily new covid-19 cases.
<b>total_deaths</b>	Numerical	The number of deaths due to the total coronavirus detected until the day the data was collected.
<b>new_deaths</b>	Numerical	The number of deaths due to the total coronavirus detected on the day the data was collected.
<b>new_deaths_smoothed</b>	Numerical	Smoothed number of daily new deaths caused by covid-19.
<b>reproduction_rate</b>	Numerical	Daily calculated coronavirus transmission or contagiousness coefficient rate.
<b>positive_rate</b>	Numerical	The rate of detecting positive in the total Covid-19 tests applied daily.
<b>stringency_index</b>	Numerical	It provides a picture of the stage at which any country enforced its strongest measures. The Stringency Index is a number from 0 to 100 that reflects these indicators. A higher index score indicates a higher level of stringency.

The *Spotify* dataset has a total of 294,600 cases, as it contains information on the 200 most-streamed tracks daily for more than 4-years. On the other hand, covid-19 data consists of a total of 352 cases. These two data sets were combined by averaging the musical characteristics of the 200 most listened-to music tracks on a daily basis. Later, this data was combined with the covid-19 data. The combined data includes all of the variables listed in Table 1 and 2 above. Besides, the combined data consists of 1473 cases, and only the last 352 cases contain both *Spotify* and covid-19 data.

Both StatSoft Statistica 14 and IBM SPSS 22 statistical package programs have been used in data analysis.

## 2.2. Methodology

In this study, we examine the effect of musical parameters on human psychology. We conduct our observations in two different time periods, pre-pandemic when social life goes on routinely, and pandemic when human psychology is heavily exposed to negative effects. *Spotify* music data consists of a period of 1473 days, and pandemic conditions were effective only in the last 352 days of that period. Therefore, in order to observe the effects of the pandemic, the 352-day data needs to be compared with the rest. In addition, by using only pandemic period (352-day) data, period-specific situations can be observed.

In addition to considering the pandemic period as a whole, considering it according to sub-periods will also provide important observation opportunities. The transmission rate of the infection, and mortality or recovery rates of the patients vary during the pandemic. The stringency rates of the measures implemented by governments also change in proportion to the variability in the rates. Therefore, both the fear of getting sick and the severity of the social restrictions that people are exposed to differ in the process. All these situations cause periodic differences in the psychological effect of the pandemic on people. In this respect, it would be a valid approach to divide the pandemic

period into sub-periods according to the course of the pandemic and examine people's music listening tendencies according to these sub-periods.

Another notable issue is the way *Spotify* data is handled. While Covid-19 data consists of a single line of information daily, *Spotify* data consists of 200 rows of information, corresponding to the 200 most-streamed tracks. The 200-rows data need to be reduced to a single row and matched with the pandemic data of that date. The reduction has been performed by taking an average of 200 rows.

The content of the "streams" variable in Spotify music data is different from other variables. "danceability," "loudness," "acousticness," etc. variables provide information about the musical aspects of the tracks, while the variable "streams" gives the total number of plays of the tracks daily. From this point of view, the "streams" variable alone does not provide satisfactory information. On the other hand, in this study, the relationship between the musical aspects of the tracks and the emotional states of people is investigated, and it is clear that the number of streaming is also a measure of emotional orientation and therefore should be taken into account. The streaming number of a song shows not only the listeners' interest in that song but also the tendency towards the musical characteristics of that song. Therefore, the weighting of the musical parameters was applied using the variable "streams". In this way, the weighting of both the song and the musical parameters in the general music preferences of the listeners was calculated. As a notation, each 'variable' takes the form "ST\_Variable" which is generated based on the weight calculation based on the number of streams.

After the data pre-processing step, qualified analyzes have been operated on the data. It has been investigated whether there is a statistically significant change between people's musical preferences in the periods of before (pre-pandemic) and after (post-pandemic) the first Covid-19 case was diagnosed. A t-test was used to test whether there was a significant difference between the arithmetic means of the musical parameter scores of the two groups (pre- and post-pandemic). In addition, Levene's test was applied to test whether the groups had equal variances.

In addition, the value Pearson Correlation has been calculated to determine the degree of correlation between musical and pandemic variables in the post-pandemic data only.

It is an undeniable fact that the change in fatality and transmission rates during the pandemic and the severity of the measures taken accordingly have direct and negative effects on human and societal psychology, just like the curfews, the compulsory wearing of masks, etc. However, these rates have not remained constant throughout the pandemic. There are sub-periods in which the rates decreased or increased in the process. Similarly, it would be a natural outcome that the effects of the pandemic on the people's psychology in these sub-periods would differ. The pandemic period data were divided into sub-periods with a k-means clustering application to observe the differences between sub-periods using the variables "new\_cases," "new\_deaths," "reproduction\_rate," "positive\_rate," and "stringency\_index," which contain important information about the daily course of the pandemic. In cases where the number of groups (categories) is more than two, ANOVA is applied to compare the means of the groups. However, when "homogeneity of variances" is violated, "Robust Tests of equality of means" such as *Welch* and *Brown-Forsythe* tests are applied. Thus, if no statistically significant violation is detected, then it would be appropriate to operate "Multiple Comparison". As a result, *Games-Howell* multiple comparison tests were applied.

Having identified a significant relationship between the pandemic and musical preferences, the next step will be to uncover the pattern of relationships. Is it possible to determine whether a randomly selected day is before or after the pandemic by looking only at people's musical preferences, thanks to a model created using music data? If such a pattern of relationships is recognized, how successful is that pattern in explaining the data? For the recognition of such a pattern, widely used machine learning algorithms were preferred. Thanks to this method, it is possible to reveal which pandemic variable is effective in musical preferences and to what extent.

One of the main issues of all machine learning algorithms is deciding when to stop, such as how to prevent the algorithm from adjusting abstruse aspects of the training data that are unlikely to increase the predictive power of the model in question. This problem is also known as overfitting. That is a common problem for many machine learning algorithms used in predictive models.

In the data, the pre-pandemic period covers a period about three times as long as the post-pandemic period. For this reason, a model built on all the data may have an overfitting problem. An equal amount of data was taken from pre-pandemic and post-pandemic data to overcome this problem. A balanced distribution of data was created covering the last 352 days before the pandemic and the first 352 days after the pandemic, making a total of 654 days. In addition, 5-fold cross-validation was applied to test the success of the models on balanced, stratified, and mutually disjoint data sets. The models were built using machine learning algorithms CART, BT, RF, SVM, and ANN. The modeling process was repeated for each of the subsets.

### **2.2.1. Classification And Regression Trees (CART)**

CART is a nonparametric and distribution-free decision tree method used for classification and prediction. Prior association detection and modeling are not required for CART, as it spontaneously forms mutually disjoint groups that provide comprehensible insight into the interactive pattern of significant independent variables (Ma 2018; Zhang and Singer 1999). Moreover, CART is a robust statistical method that can handle many types of data and outliers (Perry, et al. 2013).

### **2.2.2. Random Forest (RF)**

The random forest method generates so many decision trees and brings randomness into the composition, which can provide a model of the most informative variables (Perry, et al. 2013). In each child node (leaf node), RF first randomly selects a set of variables and then selects the ideal variable from the selected set. The process of forest construction requires repeated cycles of tree composition (He and Ding 2019). Random forest is a grouping method that uses various classification trees to strengthen predictive models (Rudy, et al. 2019).

### **2.2.3. Stochastic Gradient Boosting Trees (BT)**

Stochastic gradient boosting decision trees is a classification and regression algorithm based on bootstrapped subsets that generates child nodes by repeated splittings (Safavian and Landgrebe 1991).

### **2.2.4. Support Vector Machine (SVM)**

The support vector machine (SVM) is a supervised algorithm that operates by selecting a splitting function that divides the training set into two categories. If no splitting function can linearly split the data, a kernel function is used as a linear classifier to solve the non-linearity (Boyle 2011). In addition, SVM can also be used as a regression algorithm (Nell and Shawe-Taylor 2000).

### **2.2.5. Artificial Neural Network (ANN)**

The main goal of artificial neural networks (ANN) is to model complicated systems. The principal structure of the neural network of a biological system can manage non-linear, insufficient, and incomprehensible information. ANN is a composite process of mathematical algorithms designed to extract information by mimicking the learning process of the human brain (Cain 2017). Also, ANN can be characterized as a composition of analogically distributed computational models. The basic processing component, neurons, are deeply interconnected, dynamically reshaping themselves during the learning process (Zohuri and Moghaddam 2017).

### **2.2.6. K-Means Clustering**

K-means clustering is a simple unsupervised machine learning method that aims to partition observations into k clusters. The basic concept is to assign k centers corresponding to the k clusters. The centers should be as far apart as possible. Each point is placed in a cluster corresponding to the nearest center. After all points are distributed into clusters, the first iteration is complete. Next, k new

centroids corresponding to the k new clusters are recomputed. According to the k new centroids, a new partitioning must be performed for the same observation points. The centroids of the clusters can move to a new location after each iteration. The process is repeated as long as the centroids move (Rajaguru and Prabhakar 2017).

**2.2.7. Levene’s Test**

As a nonparametric procedure, Levene's test is used to assess the variance homogeneity. If the hypothesis that the variances of the two populations are equal is rejected, then an investigation of the reason for the inequality should be conducted (Levene 1960).

**2.2.8. Welch Test**

Welch's test is used to test the hypothesis that two populations have equal means. The test performs better than the F-test and is quite good for normally distributed populations (Wilcox 2003).

**2.2.9. Brown-Forsythe Test**

The Brown-Forsythe test is used to test the hypothesis that two groups in a population have equal variances. The test is based on an analysis of variance (ANOVA). The test statistic of the Brown-Forsythe test, which is a number calculated from a statistical test of a hypothesis, is the F statistic obtained from an ordinary one-way ANOVA (Brown and Forsythe 1974).

**2.2.10. Games-Howell Post-Hoc Test**

The Games-Howell (Games and Howell 1976) post hoc test is an extended adaptation of the Tukey-Kramer test, and it is suitable for cases where equality of variances is violated (Lee and Lee 2018).

**3. RESULTS AND DISCUSSION**

The data covering a total period of 1473 days, before and after the pandemic (B/A), were considered as a whole and it was observed whether there was a statistically significant difference between listeners' musical preferences according to the periods. A t-test was applied in which (Before/After Pandemic) the BA\_CVD19 variable is the grouping factor. Although some of the variables have normality problems, we can ignore them when we consider the number of cases (Lumley, et al. 2002).

Levene's test was used to check for equality of variances. Group statistics are given in Table 3. The number of cases before the pandemic is 1121, and the number of cases after the pandemic is 352. The percentage change for the variables Streams, ST\_Danceability, ST\_Energy, ST\_Key, ST\_Loudness, ST\_Mode, ST\_Speechiness, ST\_Acousticness, ST\_Instrumentalness, ST\_Liveness, ST\_Valence, and ST\_Tempo during the pandemic are 8.35, 9.31, 6.01, 10.07, 11.88, 5.69, 19.03, -26.45, 16.11, 14.1, and 9.97, respectively. The four variables with the greatest (absolute) change (in descending order) are ST\_Instrumentalness, ST\_Acousticness, ST\_Liveness, and ST\_Valence. Given the stress that occurred during the pandemic and the cancellation of concerts, it would be an intuitive expectation that the variables ST\_Valence (the musical positivity rate) and ST\_Liveness would change at such a rate.

**Table 3.** Group Statistics

	Group	N	Mean	Std. Dev.	Std. Err. Mean
streams	B	1121	49742,23	6214,47	185,61
	A	352	53897,31	7025,29	374,45
ST_Danceability	B	1121	30939,59	3730,02	111,41
	A	352	33818,79	4374,15	233,14
ST_Energy	B	1121	27169,30	2858,32	85,37
	A	352	28801,16	3965,37	211,36
ST_Key	B	1121	266319,01	39064,52	1166,76
	A	352	293131,83	36941,06	1968,97



ST_Loudness	B	1121	-355363,61	53026,33	1583,76
	A	352	-397566,79	50694,21	2702,01
ST_Mode	B	1121	38593,82	5067,50	151,35
	A	352	41999,37	5221,55	278,31
ST_Speechiness	B	1121	3543,27	560,46	16,74
	A	352	3745,05	652,28	34,77
ST_Acousticness	B	1121	16462,94	2964,14	88,53
	A	352	19596,50	2427,38	129,38
ST_Instrumentalness	B	1121	197,07	126,90	3,79
	A	352	144,95	115,11	6,14
ST_Liveness	B	1121	7087,14	1045,95	31,24
	A	352	8229,20	1096,88	58,46
ST_Valence	B	1121	21298,78	2477,92	74,01
	A	352	24301,91	3308,63	176,35
ST_Tempo	B	1121	5793387,47	736362,84	21993,23
	A	352	6371048,97	831038,61	44294,51

According to Levene's test (Table 4) for equality of variances, some variables are not assumed to have equality of variances (Sig. < 0.05), which means that the variability in the two groups is not equal. Equality of variance is also assumed for some variables (Sig. > 0.05), i.e., the variability in the two groups is not significantly different. However, looking at the t-test for equality of means, we find that the significance (Sig. 2-tailed) value is less than 0.05 for both cases. This situation leads to the conclusion that there is a statistically significant difference between the two groups for all variables.

**Table 4.** Independent Samples Test

		Levene's Test for Equality of		t-test for Equality of Means				
		F	Sig.	t	df	Sig. (2-tailed)	Mean Difference	Std. Error Difference
streams	Equal variances assumed	5,71	0,02	-10,60	1471,00	0,00	-4155,09	392,08
	Equal variances not assumed			-9,94	534,56	0,00	-4155,09	417,93
ST_Danceability	Equal variances assumed	10,39	0,00	-12,10	1471,00	0,00	-2879,20	237,88
	Equal variances not assumed			-11,14	521,08	0,00	-2879,20	258,39
ST_Energy	Equal variances assumed	64,87	0,00	-8,46	1471,00	0,00	-1631,87	192,94
	Equal variances not assumed			-7,16	470,95	0,00	-1631,87	227,95
ST_Key	Equal variances assumed	9,59	0,00	-11,38	1471,00	0,00	-26812,82	2356,46
	Equal variances not assumed			-11,72	616,94	0,00	-26812,82	2288,70
ST_Loudness	Equal variances assumed	3,10	0,08	13,16	1471,00	0,00	42203,18	3206,38
	Equal variances not assumed			13,48	611,01	0,00	42203,18	3131,96
ST_Mode	Equal variances assumed	1,05	0,31	-10,92	1471,00	0,00	-3405,55	311,89
	Equal variances not assumed			-10,75	573,60	0,00	-3405,55	316,80
ST_Speechiness	Equal variances assumed	13,47	0,00	-5,66	1471,00	0,00	-201,78	35,66
	Equal variances not assumed			-5,23	523,79	0,00	-201,78	38,59
ST_Acousticness	Equal variances assumed	21,73	0,00	-18,03	1471,00	0,00	-3133,56	173,84

	Equal variances not assumed			-19,99	708,01	0,00	-3133,56	156,77
ST_Instrumentalness	Equal variances assumed	0,51	0,48	6,87	1471,00	0,00	52,12	7,59
	Equal variances not assumed			7,23	640,74	0,00	52,12	7,21
ST_Liveness	Equal variances assumed	1,65	0,20	-17,66	1471,00	0,00	-1142,06	64,66
	Equal variances not assumed			-17,23	565,60	0,00	-1142,06	66,29
ST_Valence	Equal variances assumed	49,59	0,00	-18,21	1471,00	0,00	-3003,13	164,93
	Equal variances not assumed			-15,70	480,85	0,00	-3003,13	191,25
ST_Tempo	Equal variances assumed	4,24	0,04	-12,44	1471,00	0,00	-577661,50	46436,08
	Equal variances not assumed			-11,68	535,21	0,00	-577661,50	49454,08

Correlation rates between Covid-19 and music variables were calculated, considering only data from the pandemic period (last 352 days). Statistically significant correlation values are marked with an asterisk (\*). From the results, there is a significant overall correlation between the pandemic and music variables. The strengency\_index variable, which expresses the severity rates of the measures implemented by governments in the name of protection against the pandemic, which directly affects daily life and therefore the psychology of everyone who is affected or not affected with the disease during the pandemic, has a very strong correlation with all musical variables. Correlation values are given in Table 5.

**Table 5. Pearson Correlation**

	total_cases	new_cases	new_deaths_smoothed	total_deaths	new_deaths	new_deaths_smoothed_per_million	reproduction_rate	positive_rate	stringency_index	
streams	**0,279	-0,006	0,055	**0,26	0,002	0,055	**-,0596	**-,0276	**-,0733	<b>Pearson Corr.</b>
	0	0,913	0,302	0	0,964	0,302	0	0	0	<b>Sig. (2-tailed)</b>
ST_Danceability	**0,241	-0,044	0,020	**0,225	-0,023	0,020	**-,0601	**-,0302	**-,0736	<b>Pearson Corr.</b>
	0	0,414	0,705	0	0,671	0,705	0	0	0	<b>Sig. (2-tailed)</b>
ST_Energy	**0,343	0,039	*0,114	**0,327	0,039	*0,114	**-,0549	**-,0222	**-,0699	<b>Pearson Corr.</b>
	0	0,470	0,033	0	0,464	0,033	0	0	0	<b>Sig. (2-tailed)</b>
ST_Key	0,048	-0,099	-0,077	0,026	-0,083	-0,077	**-,0631	**-,0360	**-,0762	<b>Pearson Corr.</b>
	0,372	0,064	0,150	0,624	0,120	0,150	0	0	0	<b>Sig. (2-tailed)</b>
ST_Loudness	**-,020	0,035	0,001	**-,0177	0,034	0	**0,619	**0,312	**0,740	<b>Pearson Corr.</b>
	0	0,507	0,992	0,001	0,522	0,993	0	0	0	<b>Sig. (2-tailed)</b>
ST_Mode	**0,238	0,080	0,087	**0,215	0,027	0,087	**-,0572	**-,0184	**-,0705	<b>Pearson Corr.</b>
	0	0,133	0,102	0	0,612	0,102	0	0,001	0	<b>Sig. (2-tailed)</b>
ST_Speechiness	**0,201	**-,0221	-0,093	**0,188	-0,096	-0,093	**-,0619	**-,0493	**-,0735	<b>Pearson Corr.</b>
	0	0	0,082	0	0,071	0,082	0	0	0	<b>Sig. (2-tailed)</b>
ST_Acousticness	**0,236	*0,113	0,076	**0,209	0,019	0,076	**-,0571	**-,0159	**-,0698	<b>Pearson Corr.</b>
	0	0,034	0,157	0	0,729	0,156	0	0,003	0	<b>Sig. (2-tailed)</b>
ST_Instrumentalness	**-,041	**-,0107	**-,0276	**-,0421	**-,0199	**-,0276	**-,0334	**-,0172	**-,0402	<b>Pearson Corr.</b>
	0	0,044	0	0	0	0	0	0,001	0	<b>Sig. (2-tailed)</b>
ST_Liveness	**0,457	*0,121	**0,212	**0,445	*0,106	**0,212	**-,0451	*-,0113	**-,0612	<b>Pearson Corr.</b>
	0	0,023	0	0	0,046	0	0	0,035	0	<b>Sig. (2-tailed)</b>
ST_Valence	**0,381	0,046	0,098	**0,372	0,028	0,098	**-,0518	**-,0212	**-,0677	<b>Pearson Corr.</b>
	0	0,394	0,066	0	0,596	0,066	0	0	0	<b>Sig. (2-tailed)</b>
ST_Tempo	**0,313	0,037	0,098	**0,293	0,034	0,098	**-,0574	**-,0239	**-,0718	<b>Pearson Corr.</b>
	0	0,495	0,067	0	0,526	0,067	0	0	0	<b>Sig. (2-tailed)</b>

\*\* . Correlation is significant at the 0.01 level (2-tailed). \* . Correlation is significant at the 0.05 level (2-tailed).

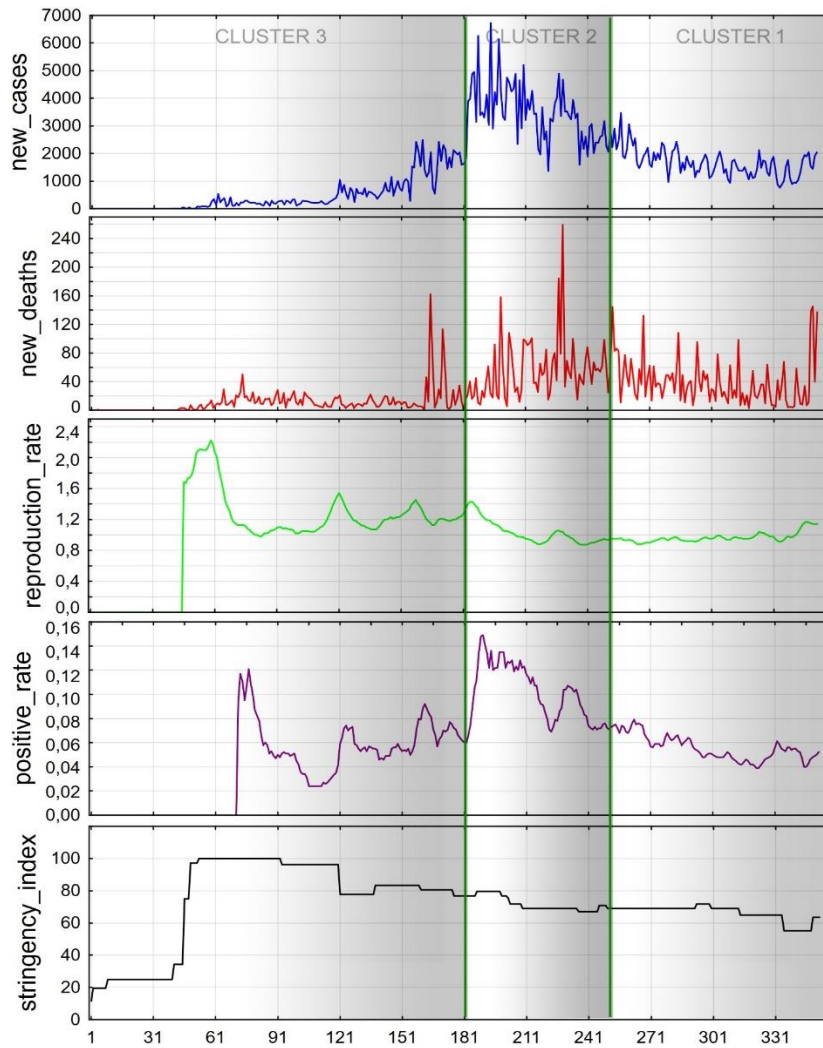
During the pandemic, the severity of the disease and the stringency rate of the measures taken in proportion, increase and decrease spontaneously. The pandemic period data (352 days) were

grouped using a k-means clustering model applied with five pandemic variables. In this way, it was possible to observe the musical variables concerning the changes in the severity of the pandemic. According to the results, the whole pandemic period was divided into 3 clusters. Then, the degree of difference between the music trends in the clusters was calculated using the multiple comparison method. Descriptive statistics for the clusters are given in Table 6.

**Table 6.** Descriptive Statistics for Clusters (Covid-19 statistics)

Cluster		new_cases	new_deaths	reproduction_rate	positive_rate	stringency_index
1	Mean	1679,853	39,608	,9675	,05575	66,1881
	N	102	102	102	102	102
	Std. Deviation	493,7987	32,6797	,06435	,010440	4,92366
	Minimum	764,0	3,0	,88	,039	55,09
	Maximum	3475,0	145,0	1,17	,079	71,76
	Variance	243837,176	1067,964	,004	,000	24,242
2	Mean	3526,691	56,971	1,0297	,10263	71,8469
	N	68	68	68	68	68
	Std. Deviation	1060,2803	42,5515	,15215	,023400	4,52522
	Minimum	1362,0	6,0	,87	,066	67,13
	Maximum	6725,0	259,0	1,43	,149	79,63
	Variance	1124194,366	1810,626	,023	,001	20,478
3	Mean	469,703	10,802	,9698	,03588	73,6879
	N	182	182	182	182	182
	Std. Deviation	612,8900	17,1059	,61695	,033082	29,31355
	Minimum	,0	,0	,00	,000	11,11
	Maximum	2486,0	162,0	2,22	,121	100,00
	Variance	375634,121	292,613	,381	,001	859,284
Total	Mean	1410,926	28,068	,9807	,05453	71,1590
	N	352	352	352	352	352
	Std. Deviation	1351,3193	34,0430	,44996	,036460	21,55411
	Minimum	,0	,0	,00	,000	11,11
	Maximum	6725,0	259,0	2,22	,149	100,00
	Variance	1826063,915	1158,924	,202	,001	464,580

Governments have taken several measures to combat the Covid 19 pandemic around the world. In cities that have entered a state of emergency, many measures have been taken, such as curfews, bans on crowded events such as cinemas, concerts, etc., and the obligation to wear masks. The majority of people have never been socially controlled by such restrictions at any time in their lives. Because of the deaths in our environment caused by the pandemic and the anxiety caused by the constant threat of death to the people themselves, they have experienced great psychological depression in the process. On the other hand, it is an indisputable fact that the measures taken by governments to slow down the death rate and the spread of disease are highly effective. However, the transmission and mortality rates have not remained constant throughout the pandemic, but there were interim periods when the rates decreased or increased. Parallel to the change in rates, the strictness of the measures implemented has also changed. Therefore, the pressure of the measures on the people also varies to the same extent. From this point of view, it is a natural result that the effects of the pandemic on people's psychology vary according to the sub-periods. To divide the pandemic period into sub-periods and thus observe the differences between the periods, a k-means clustering model was constructed with the variables "new\_cases", "new\_deaths", "reproduction\_rate", "positive\_rate" and "stringency\_index" containing important information about the entire period.



**Figure 1.** K-means clustering

The clustering model divides the data into three periods. Each period is represented as a cluster. The distribution of the data according to the 3 clusters is shown in Figure 1 for each variable. The sequence of cases being in the same cluster according to the days in a row is a strong indicator that the clustering model is successful in detecting the similarities between the cases according to the variables present. The clusters are arranged so that the most recent period is cluster 1 and the oldest period is cluster 3. When the distribution of dates within clusters is examined, cluster 3 represents the period of pandemic onset and progression, cluster 2 represents the period when the effects of the pandemic generally peak, and cluster 1 represents the period when the pandemic slows and begins to normalize. In cases where the number of groups (categories) is more than two, ANOVA is applied to compare the means between groups. Therefore, the ANOVA test was applied. According to the result of Levene's test, for all variables, the hypothesis that the variances of the two populations are equal is rejected. In other words, the problem of "homogeneity of variances" has been identified (Table 7).

**Table 7.** Test of Homogeneity of Variances

Dep. Variables	Levene Statistic	df1	df2	Sig.
streams	22,174	2	349	,000
ST_Danceability	21,831	2	349	,000
ST_Energy	17,403	2	349	,000
ST_Key	27,182	2	349	,000
ST_Loudness	27,832	2	349	,000
ST_Mode	22,934	2	349	,000
ST_Speechiness	22,461	2	349	,000

<b>ST_Acousticness</b>	23,621	2	349	,000
<b>ST_Instrumentalness</b>	45,457	2	349	,000
<b>ST_Liveness</b>	8,178	2	349	,000
<b>ST_Valence</b>	18,317	2	349	,000
<b>ST_Tempo</b>	20,089	2	349	,000

In this case, robust tests for equality of means (Table 8) such as Welch and Brown-Forsythe are applied. The test results are statistically significant (at Sig. 0.05). Based on the results, it is explicit that the 'Multiple Comparison' method can be applied.

**Table 8.** Robust Tests of Equality of Means

Robust Tests of Equality of Means		Statistic	df1	df2	Sig.
streams	<b>Welch</b>	24,650	2	197,390	,000
	<b>Brown-Forsythe</b>	25,058	2	330,481	,000
ST_Danceability	<b>Welch</b>	22,225	2	195,295	,000
	<b>Brown-Forsythe</b>	20,630	2	328,132	,000
ST_Energy	<b>Welch</b>	31,228	2	188,976	,000
	<b>Brown-Forsythe</b>	32,763	2	312,846	,000
ST_Key	<b>Welch</b>	4,866	2	202,098	,009
	<b>Brown-Forsythe</b>	3,419	2	341,763	,034
ST_Loudness	<b>Welch</b>	14,402	2	203,873	,000
	<b>Brown-Forsythe</b>	14,704	2	333,707	,000
ST_Mode	<b>Welch</b>	12,104	2	202,131	,000
	<b>Brown-Forsythe</b>	15,457	2	326,167	,000
ST_Speechiness	<b>Welch</b>	46,674	2	195,528	,000
	<b>Brown-Forsythe</b>	35,862	2	330,546	,000
ST_Acousticness	<b>Welch</b>	11,633	2	210,108	,000
	<b>Brown-Forsythe</b>	16,526	2	307,177	,000
ST_Instrumentalness	<b>Welch</b>	68,924	2	163,926	,000
	<b>Brown-Forsythe</b>	52,059	2	241,710	,000
ST_Liveness	<b>Welch</b>	42,123	2	179,089	,000
	<b>Brown-Forsythe</b>	50,044	2	282,666	,000
ST_Valence	<b>Welch</b>	36,588	2	194,176	,000
	<b>Brown-Forsythe</b>	41,069	2	314,797	,000
ST_Tempo	<b>Welch</b>	26,397	2	193,858	,000
	<b>Brown-Forsythe</b>	28,396	2	323,483	,000

Looking at the pandemic data suggests that the periods in clusters 2 and 3 are the periods when the pandemic is most intense. In this regard, the fact that value movements in musical variables move synchronously with this periodicity will be a factor supporting the hypothesis that people's musical preferences vary according to pandemic conditions. Examining the results of Multiple Comparison (Table 9), it was found that there is a statistically significant difference between the mean values of the distributions of the variables ST\_Speechiness, ST\_Acousticness, and ST\_Instrumentalness according to the three periods. Except for these three variables, there is no statistically significant difference between the means of the Cluster 2 and 3 distributions. On the other hand, it was determined that there were statistically significant differences between the Cluster 1 distribution means and the Cluster 2 and 3 distribution means, for all variables. Only for the variable, ST\_Key was the difference between the means of the data distribution of the cluster 3 and cluster 1 periods not statistically significant. From this point of view, it can be concluded that the distribution of musical variables, in general, varies significantly according to the pandemic conditions.

**Table 9.** Multiple comparison

Dependent Variable	(I) Cluster	(J) Cluster	Mean Difference (I-J)	Std. Error	Sig.
streams	1	2	4992,37	812,48	** ,000
		3	4915,24	810,35	** ,000

	2	3	-77,13	805,56	0,995
ST_Danceability	1	2	3174,73	514,22	** ,000
		3	2642,3	504,4	** ,000
	2	3	-532,43	515,13	0,557
ST_Energy	1	2	3046,58	481,7	** ,000
		3	3272,78	449,34	** ,000
	2	3	226,2	466,71	0,879
ST_Key	1	2	12994,11	4165,75	** ,006
		3	5632,94	4350,54	0,399
	2	3	-7361,17	4396,73	0,218
ST_Loudness	1	2	-28708,99	5797,5	** ,000
		3	-27067,66	6076,92	** ,000
	2	3	1641,34	5666,48	0,955
ST_Mode	1	2	2054,53	612,55	** ,003
		3	3089,6	633,59	** ,000
	2	3	1035,08	579,91	0,177
ST_Speechiness	1	2	694,54	72,21	** ,000
		3	374,25	71,32	** ,000
	2	3	-320,29	75,54	** ,000
ST_Acousticness	1	2	674,11	275,5	** ,041
		3	1447,86	303,78	** ,000
	2	3	773,75	246,87	** ,005
ST_Instrumentalness	1	2	-75,53	12,42	** ,000
		3	-115,39	10,55	** ,000
	2	3	-39,87	14,91	** ,022
ST_Liveness	1	2	898,11	141,12	** ,000
		3	1153,33	126,62	** ,000
	2	3	255,22	122,45	0,097
ST_Valence	1	2	2891,76	390,69	** ,000
		3	2962,46	382,4	** ,000
	2	3	70,7	360,39	0,979
ST_Tempo	1	2	572799,97	98303,5	** ,000
		3	640326,8	95352,19	** ,000
	2	3	67526,83	96131,18	0,762

\*\* Significant at the 0.05 level

After determining that the relationship between the musical and pandemic variables was qualified and significant, machine learning algorithms were used to determine the pattern of relationships. If a randomly selected day can be determined whether to be from the pandemic period thanks to a model created with the musical data, this model would show the most qualified relationship pattern available between the two datasets. For the detection of the expressed pattern, widely used machine learning algorithms were preferred. These algorithms are CART, BT, RF, SVM, and ANN. Since the pre-pandemic period covers a period approximately three times longer than the post-pandemic period, a model that ignores this unbalanced distribution will have problems with overfitting. Therefore, modeling was performed using an equal number of data from the pre-pandemic and post-pandemic periods. We used 352-days pre-pandemic data to balance the 352-days post-pandemic data. In addition, 5-fold cross-validation was used to assess the predictive performance of the models. To do this, the dataset was divided into five subsets that are stratified, mutually disjoint, and have balanced distributions. Each model was reproduced for each of the five subsets.

**Table 10.** C&RT Tree Structure (\* terminal node)

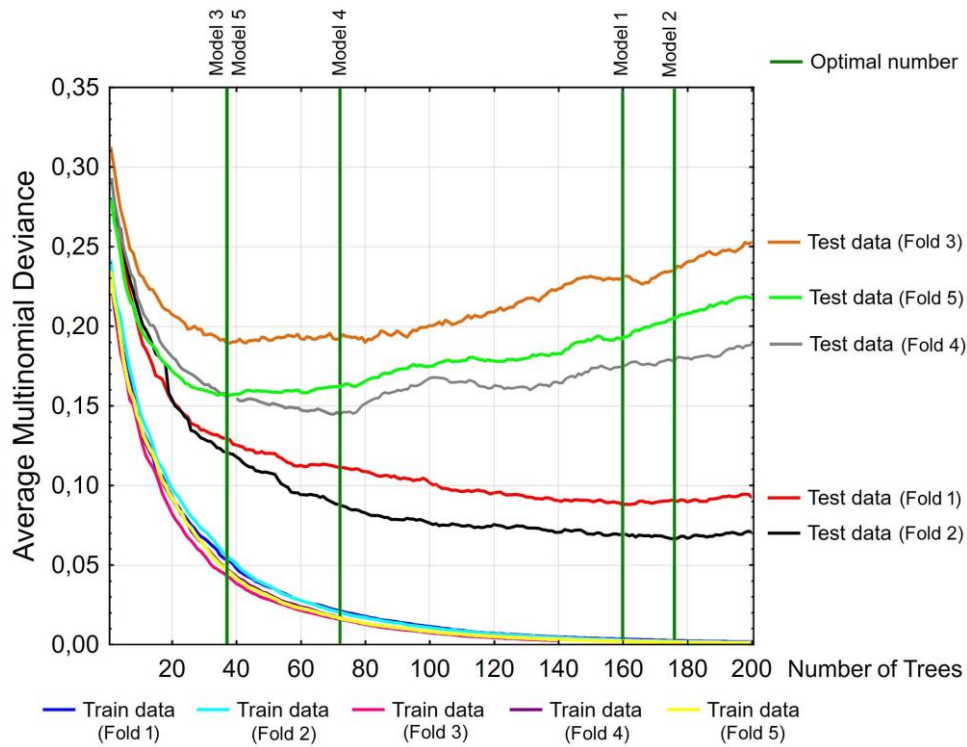
	Node ID	Num. of Nodes	Size of Node	N in Class (0)	N in Class (1)	Selected Category	Split Variable	Criterion (Child 1)	Criterion (Child 2)	Child Node (1)	Child Node (2)
Fold-I	1	2	562	281	281	0	ST_Instrumentalness	$x \leq 124,87$	$x > 124,87$	2	3
	*2		169	4	165	1					
	3	2	393	277	116	0	ST_Speechiness	$x \leq 3630,86$	$x > 3630,86$	4	5
	*4		65	2	63	1					
	5	2	328	275	53	0	ST_Valence	$x \leq 26901,34$	$x > 26901,34$	6	7
	6	2	268	257	11	0	ST_Speechiness	$x \leq 3802,85$	$x > 3802,85$	8	9

Fold-1	8	2	15	11	4	0	ST_Danceability	$x \leq 31681,23$	$x > 31681,23$	10	11	
	*10		10	10	0	0						
	*11		5	1	4	1						
	9	2	253	246	7	0	ST_Energy	$x \leq 27439,69$	$x > 27439,69$	12	13	
	12		11	8	3	0						
	13	2	242	238	4	0	ST_Valence	$x \leq 26334,43$	$x > 26334,43$	14	15	
	14	2	227	226	1	0	ST_Valence	$x \leq 24689,77$	$x > 24689,77$	16	17	
	*16		165	165	0	0						
	*17		62	61	1	0						
	*15		15	12	3	0						
	7	2	60	18	42	1	ST_Instrumentalness	$x \leq 535,15$	$x > 535,15$	18	19	
	*18		43	1	42	1						
	*19		17	17	0	0						
	Fold-2	1	2	562	281	281	0	ST_Instrumentalness	$x \leq 123,30$	$x > 123,30$	2	3
		2	2	164	3	161	1	ST_Key	$x \leq 338282,99$	$x > 338282,99$	4	5
		*4		157	1	156	1					
		*5		7	2	5	1					
		3	2	398	278	120	0	ST_Speechiness	$x \leq 3610,82$	$x > 3610,82$	6	7
		*6		67	1	66	1					
7		2	331	277	54	0	ST_Valence	$x \leq 26901,34$	$x > 26901,34$	8	9	
8		2	272	259	13	0	ST_Speechiness	$x \leq 3802,85$	$x > 3802,85$	10	11	
10		2	18	13	5	0	ST_Loudness	$x \leq -373351$	$x > -373351$	12	13	
*12			5	0	5	1						
*13			13	13	0	0						
11		2	254	246	8	0	ST_Energy	$x \leq 27418,08$	$x > 27418,08$	14	15	
*14			13	9	4	0						
15		2	241	237	4	0	ST_Valence	$x \leq 26334,43$	$x > 26334,43$	16	17	
16		2	226	225	1	0	ST_Energy	$x \leq 29267,23$	$x > 29267,23$	18	19	
*18			59	58	1	0		$x \leq 29267,23$				
*19			167	167	0	0		$x \leq 29267,23$				
*17			15	12	3	0		$x \leq 29267,23$				
9		2	59	18	41	1	ST_Instrumentalness	$x \leq 534,42$	$x > 534,42$	20	21	
*20			42	1	41	1						
*21			17	17	0	0						
Fold-3	1	2	564	282	282	0	ST_Instrumentalness	$x \leq 125,84$	$x > 125,84$	2	3	
	*2		170	5	165	1						
	3	2	394	277	117	0	ST_Speechiness	$x \leq 3630,86$	$x > 3630,86$	4	5	
	*4		65	2	63	1						
	5	2	329	275	54	0	ST_Valence	$x \leq 26901,34$	$x > 26901,34$	6	7	
	6	2	268	256	12	0	ST_Energy	$x \leq 27312,19$	$x > 27312,19$	8	9	
	8	2	20	15	5	0	ST_Key	$x \leq 285709,02$	$x > 285709,02$	10	11	
	*10		15	15	0	0						
	*11		5	0	5	1						
	9	2	248	241	7	0	ST_Speechiness	$x \leq 3810,84$	$x > 3810,84$	12	13	
	*12		9	5	4	0						
	13	2	239	236	3	0	ST_Valence	$x \leq 26334,43$	$x > 26334,43$	14	15	
	14	2	228	227	1	0	ST_Energy	$x \leq 29267,87$	$x > 29267,87$	16	17	
	*16		61	60	1	0						
	*17		167	167	0	0						
	*15		11	9	2	0						
	7	2	61	19	42	1	ST_Instrumentalness	$x \leq 325,48$	$x > 325,48$	18	19	
	*18		40	0	40	1						
	*19		21	19	2	0						
Fold-4	1	2	564	282	282	0	ST_Instrumentalness	$x \leq 123,30$	$x > 123,30$	2	3	
	2	2	158	1	157	1	ST_Speechiness	$x \leq 4700,70$	$x > 4700,70$	4	5	
	*4		152	0	152	1						
	*5		6	1	5	1						
	3	2	406	281	125	0	ST_Speechiness	$x \leq 3611,23$	$x > 3611,23$	6	7	
	*6		71	1	70	1						
	7	2	335	280	55	0	ST_Valence	$x \leq 26906,01$	$x > 26906,01$	8	9	
	8	2	272	261	11	0	ST_Speechiness	$x \leq 3802,85$	$x > 3802,85$	10	11	
	10	2	21	16	5	0	ST_Danceability	$x \leq 31681,23$	$x > 31681,23$	12	13	
	*12		15	15	0	0						



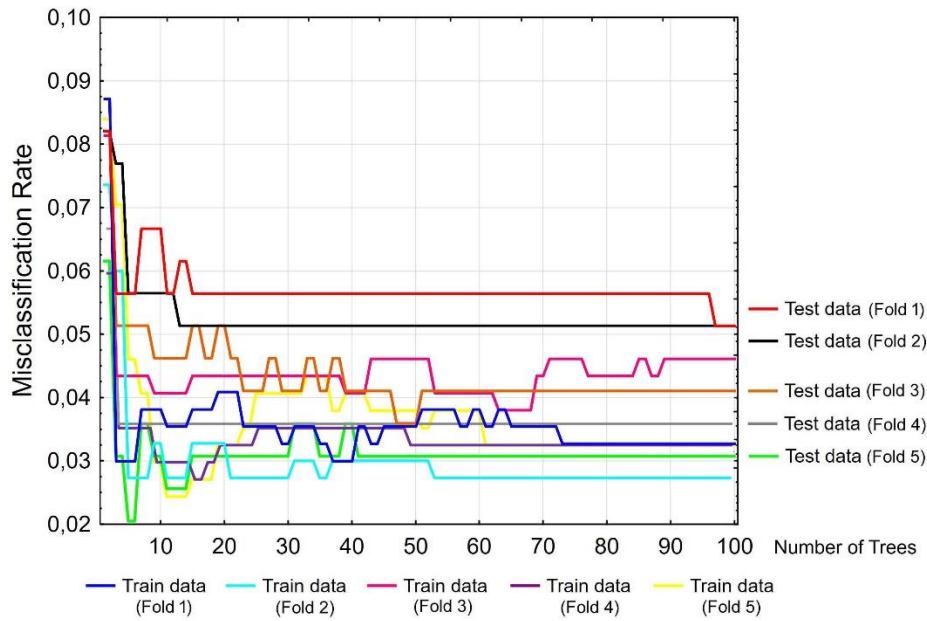
	*13		6	1	5	1					
	11	2	251	245	6	0	ST_Energy	$x \leq 27439,69$	$x > 27439,69$	14	15
	*14		13	9	4	0					
	15	2	238	236	2	0	ST_Valence	$x \leq 26474,14$	$x > 26474,14$	16	17
	16	2	228	227	1	0	ST_Energy	$x \leq 29267,23$	$x > 29267,23$	18	19
	*18		59	58	1	0					
	*19		169	169	0	0					
	*17		10	9	1	0					
	9	2	63	19	44	1	ST_Instrumentalness	$x \leq 325,48$	$x > 325,48$	20	21
	*20		42	0	42	1					
	*21		21	19	2	0					
Fold-5	1	2	564	282	282	0	ST_Instrumentalness	$x \leq 134,76$	$x > 134,76$	2	3
	2	2	191	12	179	1	ST_Acousticness	$x \leq 15449,12$	$x > 15449,12$	4	5
	*4		9	7	2	0					
	5	2	182	5	177	1	ST_Acousticness	$x \leq 16138,24$	$x > 16138,24$	6	7
	*6		17	4	13	1					
	7	2	165	1	164	1	ST_Key	$x \leq 338282,99$	$x > 338282,99$	8	9
	*8		156	0	156	1					
	*9		9	1	8	1					
	3	2	373	270	103	0	ST_Speechiness	$x \leq 3699,77$	$x > 3699,77$	10	11
	*10		56	3	53	1					
	11	2	317	267	50	0	ST_Valence	$x \leq 26901,34$	$x > 26901,34$	12	13
	12	2	257	249	8	0	ST_Energy	$x \leq 27312,19$	$x > 27312,19$	14	15
	14	2	18	13	5	0	ST_Acousticness	$x \leq 18336,67$	$x > 18336,67$	16	17
	*16		13	13	0	0					
	*17		5	0	5	1					
	15	2	239	236	3	0	ST_Valence	$x \leq 26334,43$	$x > 26334,43$	18	19
	*18		225	225	0	0					
	*19		14	11	3	0					
	13	2	60	18	42	1	ST_Instrumentalness	$x \leq 534,42$	$x > 534,42$	20	21
	*20		44	2	42	1					
	*21		16	16	0	0					

C& RT model structures are given separately for each fold in Table 10. The models created in Fold-1 and 3 have 10 terminal nodes each, while the others have 11 terminal nodes. The training sets created based on random and balanced selection consist of 562 cases in Fold-1 and 2, while the others consist of 564 cases. ST\_instrumentalness was the first split variable for all models. The total number of terminal nodes for the five models is 52. In determining the terminal nodes, the lowest 55.5% and the highest 100% confidence values were reported. The mean confidence value was 92.5%, and the standard deviation was 10.7%. These results demonstrate the high performance of the C& RT models.



**Figure 2.** Summary of Boosted Trees Models

Overfitting is one of the main problems for most machine learning algorithms used in predictive data analytics applications. A general solution to avoid this problem is to balance the distribution of the data according to the categories of the dependent variable when constructing (training) the corresponding model, and to compare the available observed values with the model-based predicted values in an unused test subset. In this way, it is possible to measure the accuracy performance of the respective model and control potential overfitting. In model building, each successive simple tree is built on a randomly selected subset of the dataset. In other words, each consecutive tree is built independently on a randomly assembled subset. In this way, together with the 5-fold cross-validation, it will be a strong shield against the overfitting problem. This approach, i.e., performing sequential boosting calculations on subsamples taken independently, is called stochastic gradient boosting. In Figure 2, one can see the point where the potential overfitting problem of the model starts. It can be observed that the estimation error for the training data gradually approaches zero as more trees are added to the model. It can be observed that the estimation error for the training data gradually approaches zero as more trees are added to the model. This point, expressed as the optimal point, clearly shows the point where the overfitting problem begins. At this point, the process of adding a new tree to the model should be stopped. The number of trees added up to this point is called the optimal number of trees. The optimal number of trees of the models is 160, 176, 37, 72, and 37, respectively. The maximum tree size is 13 for all five models. The value of Average Multinomial Deviance shows how far a prediction value is from the observation value.



**Figure 3.** Summary of Random Forest Models

For both CT and RF, misclassification rates (out-of-bag error (OOB)) were used to measure predictive success in models based on data generated with a random Bootstrap sample. Since the data was artificially generated based on the bootstrapping algorithm, it would be more appropriate to pay attention to model performance, even though both models have a tree structure similar to C& RT. Although tiny random changes that may occur during the data generation process result in changes in the bootstrap models, there is no large change in the final performance. Random Forest is an algorithm that generates black-box models like Boosted Trees. A similar overfitting problem exists in Random Forest models. In this model, instead of specifying an optimal number of trees, an optimal misclassification range is determined and when the misclassification values reach this range, the model stops performing new iterations. It was observed that the models (in the test sets) are fixed in the range of 0.03 - 0.06 misclassifications (Figure 3).

**Table 11.** SVM Model specifications

	Num. independents	of SVM Classification	Kernel type	Num. of SVs	Num. of SVs ( 0 )	Num. of SVs ( 1 )
<b>Fold-1</b>	12	type 1	RBF	75 ( 42 bounded)	37	38
<b>Fold-2</b>	12	type 1	RBF	63 ( 33 bounded)	33	30
<b>Fold-3</b>	12	type 1	RBF	76 ( 41 bounded)	37	39
<b>Fold-4</b>	12	type 1	RBF	68 ( 30 bounded)	34	34
<b>Fold-5</b>	12	type 1	RBF	58 ( 20 bounded)	26	32

The next step was to create SVM models. All available musical variables (12 independent variables) were used in the models. In addition, the type-1 SVM classification and the RBF kernel function (depending on the error function) were used in all models, as this increases the resistance of the model to the overfitting problem. The number of SVs used in the models is given in Table 11.

**Table 12.** Summary of Active Artificial Neural Networks

	Net. Name	Test Perf. (%)	Training Algorithm	Error Function	Hidden Activation	Output Activation
<b>Fold-1</b>	MLP 12-5-2	96,43	BFGS 35	SOS	Tanh	Identity
<b>Fold-2</b>	MLP 12-10-2	99,11	BFGS 66	SOS	Exponential	Exponential
<b>Fold-3</b>	MLP 12-4-2	98,21	BFGS 40	SOS	Exponential	Exponential
<b>Fold-4</b>	MLP 12-7-2	100,00	BFGS 35	SOS	Exponential	Identity
<b>Fold-5</b>	MLP 12-10-2	100,00	BFGS 75	Entropy	Exponential	Softmax

The algorithm ANN performs a random %80:%20 split of the dataset, trains on the 80% portion, and tests the models on the 20% portion. Here, ANN has trained five neural networks (Table 12), each with 12 predictors and 2 outputs (prediction), and several nodes in the middle layer varying from 4 to 10 (Net. Name). The term MLP stands for multilayer perceptron, another name for a neural network. As the complexity in the association between the independent variables and the dependent variables increases, the number of processing components in the middle layer should also increase. If the output pattern is distinctly nonlinear, then several middle layers should be implemented to capture the nonlinear relationships. If the output is not highly nonlinear, then the excessive layers will result in overtraining (Nisbet, Elder and Miner 2009). The five networks also differed in the activation functions used to pass data through a node. Data is collected and transmitted to the subsequent node by the activation function. Various activation functions produce slightly distinct solutions of the model for a given case. The models that perform best in the algorithm report are models 4 and 5.

**Table 13.** 5-Fold Cross Validation

			C&RT		RF		BT		SVM		ANN	
			Predicted									
			0	1	0	1	0	1	0	1	0	1
Fold-1	Observed	0	69	2	68	3	71	0	70	1	71	0
		1	2	69	2	69	3	68	2	69	6	65
Fold-2		0	68	3	65	6	69	2	68	3	70	1
		1	1	70	1	70	2	69	0	71	0	71
Fold-3		0	68	2	69	1	68	2	70	0	70	0
		1	3	67	3	67	4	66	2	68	5	65
Fold-4		0	67	3	66	4	67	3	69	1	70	0
		1	4	66	5	65	2	68	2	68	2	68
Fold-5		0	67	3	69	1	67	3	69	1	70	0
		1	4	66	4	66	3	67	2	68	2	68

Table 13 provides the most detailed information about the performance of the five models in each fold. The number of correct and incorrect predictions made by the models with respect to the dependent variable categories (0/1) are given categorically in the table. There is an average of 70 (or 71) cases per category. The number of incorrect predictions for each category ranges from 0 to 6.

**Table 14.** Model Evaluation Summary (Accuracy Rates) (Testing Set)

	C&RT	RF	BT	SVM	ANN
Fold-1	97,18	96,48	97,89	97,89	95,77
fold-2	95,07	97,18	97,18	97,89	99,30
Fold-3	96,43	97,14	95,71	98,57	96,43
Fold-4	95,00	93,57	96,43	97,86	98,57
Fold-5	95,00	96,43	95,71	97,86	98,57
Mean	95,73	96,16	96,58	98,01	97,72

Table 14 shows the prediction accuracy rates provided by the five models in each fold. The average values of the accuracy rates of the models vary from 95.73% (C% RT) to 98.01% (SVM). In terms of accuracy obtained specifically for one model, the lowest value is 93.57% (RF) and the highest is 99.30% (ANN).

#### 4. CONCLUSION

COVID -19, which was first detected in China in late 2019, was recognized as a global pandemic by the World Health Organization (WHO) on March 11, 2020 (Kasapoğlu 2020). The impact of the pandemic is unprecedented, as it has evolved from a health crisis to a global economic and social disaster (Simon, et al. 2021). In addition, the pandemic has caused public anxiety and psychological problems (Wang, et al. 2020). The pandemic has led to the implementation of policies such as social distancing and containment worldwide to slow the spread of the disease (Tull, et al.

2020). The restrictions have affected the physical and psychological well-being of people around the world (Javed, et al. 2020). COVID-19-associated ailments and associated quarantines have had a significant negative impact on mental health (Arslan, Yildirim and Zangeneh, Coronavirus Anxiety and Psychological Adjustment in College Students: Exploring the Role of College Belongingness and Social Media Addiction 2021). Some studies have examined factors of happiness during quarantine, while other studies have focused on psychological well-being and related topics (Greyling, Rossouw and Adhikari 2021).

Internet is a novel tool to improve the professional skills needed for today's modern industries, as well as an alternative way to communicate and obtain information (McDool, et al. 2020). In particular, the Internet, which is used extensively during the pandemic for professional, commercial, educational, social, etc. purposes, has become more indispensable. Because of the social distancing rules that were in place during the pandemic, it has become common to use the Internet as a means of entertainment and relaxation in addition to these purposes, and thus the Internet has become a natural part of daily life. All the old habits have gained a new form based on the Internet and the computer. The effects of this change can also be seen in the form of consumption of art and cultural content. Artistic and cultural activities, used as a kind of break in hectic and stressful periods of life, have become more integrated with Internet access in the new period. From music albums to concert halls, all the musical content we access physically has also gained an almost entirely online form. Music is a simple and inexpensive therapy (Koelsch and Jäncke, Music and the heart 2015). Music is a natural and important part of both society and individual life (Hizlisoy, Yildirim and Tufekci 2021).

Data on music listening provide a special opportunity to study and understand aspects of human psychology (Anderson, et al. 2021). This study examined the relationship between the pandemic and *Spotify* music data in the Philippines from different perspectives and in detail. Following statistically different data analysis approaches, it was found that people's music listening behavior differs significantly under pandemic conditions. In addition, significant differences in people's consumption patterns of musical content were observed depending on the decreasing or increasing severity of the pandemic. The results confirm that people's musical preferences are significantly shaped depending on their emotional state. Given the above results, the ultimate implication is that if policymakers want to raise happiness levels and increase the likelihood of achieving happiness levels, they must consider the factors that matter most to people's happiness. Music and musical activities can be used as an effective tool to positively influence people's emotions.

This study investigated people's musical preferences under the pandemic's depressive conditions. Data from *Spotify*'s contact profiles and musical catalog were merged with epidemic information. Instead of offering a specific hypothesis, the models in the descriptive study we presented were constructed on the presumption that there is a linear or non-linear link between people's emotional states and musical preferences. According to our research, during the pandemic time, music listening patterns significantly changed. Also, our study had several limitations. First, the sample consisted solely of Filipino users of *Spotify*. Therefore, it is not clear to what extent our conclusions can be generalized to other geographical regions and cultures. Moreover, while recent empirical evidence suggests that music is universal in form and function (Mehr, et al. 2019), because our results are based on streaming services that require users to have Internet-enabled devices, we cannot generalize our results to regions where there is limited Internet connectivity or online music streaming services.

## REFERENCES

Aljanaki, Anna, Frans Wiering, and Remco C. Veltkamp. "Studying emotion induced by music through a crowdsourcing game." *Information Processing & Management* 52, no. 1 (2016): 115-128.

- Anderson, Ian, et al. "'Just the Way You Are': Linking Music Listening on Spotify and Personality." *Social Psychological and Personality Science* 12, no. 4 (May 2021): 561–72.
- Arslan, Gökmen, Murat Yıldırım, Ahmet Tanhan, Metin Buluş, and Kelly-Ann Allen. "Coronavirus Stress, Optimism-Pessimism, Psychological Inflexibility, and Psychological Health: Psychometric Properties of the Coronavirus Stress Measure." *International Journal of Mental Health and Addiction*, 2020: 1557-1882.
- Arslan, Gökmen, Murat Yıldırım, and Masood Zangeneh. "Coronavirus Anxiety and Psychological Adjustment in College Students: Exploring the Role of College Belongingness and Social Media Addiction." *International Journal of Mental Health and Addiction*, 2021: 1557-1882.
- BAYRAKTAR, Ahmet. "COVID 19 Pandemisinin Finansal Etkileri: BİST İmalat Sektörü Uygulaması." *Journal of Turkish Studies* 15, no. 8 (2020): 3415-3427.
- Bogt, Tom F.M. Ter, Juul Mulder, Quinten A.W. Raaijmakers, and Saoirse Nic Gabhainn. "Moved by music: A typology of music listeners." *Psychology of Music* 39, no. 2 (2011): 147-163.
- Boyle, Brandon H. *Support Vector Machines:: Data Analysis, Machine Learning and Applications (Computer Science, Technology and Applications)*. Nova Science Publishers Inc, 2011.
- Bresin, Roberto, and Anders Friberg. "Emotion rendering in music: Range and characteristic values of seven musical variables." *Cortex* 47, no. 9 (2011): 1068-1081.
- Brown, Morton B., and Alan B. Forsythe. "Robust Tests for the Equality of Variances." *Journal of the American Statistical Association* (Taylor & Francis) 69, no. 346 (1974): 364-367.
- Cain, Gayle. *Artificial Neural Networks: New Research (Computer Science, Technology and Applications)*. Nova Science Pub Inc, 2017.
- Castillo-Pérez, Sergio, Virginia Gómez-Pérez, Minerva Calvillo Velasco, Eduardo Pérez-Campos, and Miguel-Angel Mayoral. "Effects of music therapy on depression compared with psychotherapy." *The Arts in Psychotherapy* 37, no. 5 (2010): 387-390.
- Chaturvedi, Kunal, Dinesh Kumar Vishwakarma, and Nidhi Singh. "COVID-19 and its impact on education, social life and mental health of students: A survey." *Children and Youth Services Review* 121 (2021): 105866.
- Donthu, Naveen, and Anders Gustafsson. "Effects of COVID-19 on business and research." *Journal of Business Research* 117 (2020): 284-289.
- Eerola, Tuomas, Jonna K. Vuoskoski, Henna-Riikka Peltola, Vesa Putkinen, and Katharina Schäfer. "An integrative review of the enjoyment of sadness associated with music." *Physics of Life Reviews* 25 (2018): 100-121.
- Faus, S., A. Matas, and E. Elósegui. "Music and regaining calm when faced with academic stress." *Cogent Arts & Humanities* (Cogent OA) 6, no. 1 (2019): 1634334.
- Flanagan, Emily W., Robbie A. Beyl, S. Nicole Fearnbach, Abby D. Altazan, Corby K. Martin, and Leanne M. Redman. "The Impact of COVID-19 Stay-At-Home Orders on Health Behaviors in Adults." *Obesity* 29, no. 2 (2021): 438-445.

- Games, Paul A., and John F. Howell. "Pairwise Multiple Comparison Procedures with Unequal N's and/or Variances: A Monte Carlo Study." *Journal of Educational Statistics* (American Educational Research Association) 1, no. 2 (1976): 113-125.
- Giordano, Filippo, et al. "Receptive music therapy to reduce stress and improve wellbeing in Italian clinical staff involved in COVID-19 pandemic: A preliminary study." *The Arts in Psychotherapy* 70 (2020): 101688.
- Greyling, Talita, Stephanie Rossouw, and Tamanna Adhikari. "The good, the bad and the ugly of lockdowns during Covid-19." *PLoS ONE* 16, no. 1 (2021).
- Gurgen, Elif Tekin. "Social and Emotional Function of Music Listening: Reasons for Listening to Music." *Eurasian Journal of Educational Research*, no. 66 (2016): 229-242.
- He, Ruisi, and Zhiguo Ding. *Applications of Machine Learning in Wireless Communications 81. vol/IET telecommunications series*. Institution of Engineering and Technology, 2019.
- Hizlisoy, Serhat, Serdar Yildirim, and Zekeriya Tufekci. "Music emotion recognition using convolutional long short term memory deep neural networks." *Engineering Science and Technology, an International Journal* 24, no. 3 (2021): 760-767.
- Javed, Bilal, Abdullah Sarwer, Erik B. Soto, and Zia-ur-Rehman Mashwani. "The coronavirus (COVID-19) pandemic's impact on mental health." *The international Journal of health planning and management* 35, no. 5 (2020): 993– 996.
- Jena, Pradyot Ranjan, Ritanjali Majhi, Rajesh Kalli, Shunsuke Managi, and Babita Majhi. "Impact of COVID-19 on GDP of major economies: Application of the artificial neural network forecaster." *Economic Analysis and Policy* 69 (2021): 324-339.
- Juslin, Patrik N, Simon Liljeström, Daniel Västfjäll, Gonçalo Barradas, and Ana Silva. "An experience sampling study of emotional reactions to music: listener, music, and situation." *Emotion (Washington, D.C.)* 8, no. 5 (2008): 668–683.
- Karasmanaki, Evangelia, and Georgios Tsantopoulos. "Impacts of social distancing during COVID-19 pandemic on the daily life of forestry students." *Children and Youth Services Review* 120 (2021): 105781.
- Kasapoğlu, Figen. "Examining the Relationship between Fear of COVID-19 and Spiritual Well-Being." *Spiritual Psychology and Counseling* 5 (2020): 341-354.
- Koelsch, Stefan. "Investigating the Neural Encoding of Emotion with Music." *Neuron* 98, no. 6 (2018): 1075-1079.
- Koelsch, Stefan, and Lutz Jäncke. "Music and the heart." *European Heart Journal* 36, no. 44 (2015): 3043–3049.
- Laing, Timothy. "The economic impact of the Coronavirus 2019 (Covid-2019): Implications for the mining industry." *The Extractive Industries and Society* 7, no. 2 (2020): 580-582.
- Lee, Sangseok, and Dong Kyu Lee. "What is the proper way to apply the multiple comparison test?" *Korean J Anesthesiol* 71, no. 5 (2018): 353-360.
- Levene, Howard. "Robust tests for equality of variances." *Contributions to Probability and Statistics: Essays in Honor of Harold Hotelling* (Stanford University Press), 1960: 278–292.

- Lumley, Thomas, Paula Diehr, Scott Emerson, and Lu Chen. "The Importance of the Normality Assumption in Large Public Health Data Sets." *Annual Review of Public Health* 23 (2002): 151-169.
- Ma, Xin. *Using Classification and Regression Trees: A Practical Primer*. Information Age Publishing, 2018.
- Martín, Javier Centeno, Delfin Ortega-Sánchez, Ignacio Nieto Miguel, and Gracia María Gil Martín. "Music as a factor associated with emotional self-regulation: A study on its relationship to age during COVID-19 lockdown in Spain." *Heliyon* 7, no. 2 (2021).
- McDool, Emily, Philip Powell, Jennifer Roberts, and Karl Taylor. "The internet and children's psychological wellbeing." *Journal of Health Economics* 69 (2020): 102274.
- Mehr, Samuel A., et al. "Universality and diversity in human song." *Science* (American Association for the Advancement of Science) 366, no. 6468 (2019).
- Nell, Cristianini, and John Shawe-Taylor. *An Introduction to Support Vector Machines and Other Kernel-Based Learning Methods*. Cambridge: Cambridge University Press, 2000.
- Nicola, Maria, et al. "The socio-economic implications of the coronavirus pandemic (COVID-19): A review." *International Journal of Surgery* 78 (2020): 185-193.
- Nisbet, Robert, John Elder, and Gary Miner. *Handbook of Statistical Analysis and Data Mining Applications 1st Edition*. Academic Press, 2009.
- Perry, Walter L., et al. "Predicting Suicide Attacks : Integrating Spatial, Temporal, and Social Features of Terrorist Attack Targets." *RAND Corporation*, 2013.
- Pothoulaki, M., et al. "An Investigation of the Effects of Music on Anxiety and Pain Perception in Patients Undergoing Haemodialysis Treatment." *Journal of Health Psychology* 13, no. 7 (2008): 912-920.
- Prey, Robert, Marc Esteve Del Valle, and Leslie Zwerwer. "Platform pop: disentangling Spotify's intermediary role in the music industry." *Information, Communication & Society* (Routledge), 2020: 1-19.
- Radermecker, Anne-Sophie V. "Art and culture in the COVID-19 era: for a consumer-oriented approach." *SN Business & Economics* 1, no. 1 (2020): 4.
- Rahman, Jawaria, Abilash Muralidharan, Sohail J. Quazi, Hajra Saleem, and Safeera Khan. "Neurological and Psychological Effects of Coronavirus (COVID-19): An Overview of the Current Era Pandemic." *Cureus* 12, no. 6 (2020).
- Rajaguru, Harikumar, and Sunil Kumar Prabhakar. *KNN Classifier and K-Means Clustering for Robust Classification of Epilepsy from EEG Signals. A Detailed Analysis*. Anchor Academic Publishing, 2017.
- Rudy, A. C. A., P. D. Morse, S. V. Kokelj, W. E. Sladen, and S. L. Smith. "A New Protocol to Map Permafrost Geomorphic Features and Advance Thaw-Susceptibility Modelling." *Cold Regions Engineering 2019*. 2019. 661-669.
- Safavian, S.R., and D. Landgrebe. "A survey of decision tree classifier methodology." *IEEE Transactions on Systems, Man, and Cybernetics*, 1991: 660 - 674.



- Sakka, Laura S., and Patrik N. Juslin. "Emotion regulation with music in depressed and non-depressed individuals: Goals, strategies, and mechanisms." *Music & Science* 1 (2018).
- Sharif, Arshian, Chaker Aloui, and Larisa Yarovaya. "COVID-19 pandemic, oil prices, stock market, geopolitical risk and policy uncertainty nexus in the US economy: Fresh evidence from the wavelet-based approach." *International Review of Financial Analysis* 70 (2020): 101496.
- Simon, Judit, Timea M. Helder, Ross G. White, and Catharina van der Boor & Agata Łaszewska. "Impacts of the Covid-19 lockdown and relevant vulnerabilities on capability well-being, mental health and social support: an Austrian survey study." *BMC Public Health* 21 314 (2021).
- Sood, Sadhika. "Psychological Effects of the Coronavirus Disease-2019 Pandemic." *RHiME* 7 (2020): 23-26.
- Tull, Matthew T., Keith A. Edmonds, Kayla M. Scamaldo, Julia R. Richmond, Jason P. Rose, and Kim L. Gratz. "Psychological Outcomes Associated with Stay-at-Home Orders and the Perceived Impact of COVID-19 on Daily Life." *Psychiatry Research* 289 (2020): 113098.
- Vonderau, Patrick. "The Spotify Effect: Digital Distribution and Financial Growth." *Television & New Media* 20, no. 1 (2019): 3-19.
- Wang, Ying, et al. "Acute psychological effects of Coronavirus Disease 2019 outbreak among healthcare workers in China: a cross-sectional study." *Translational Psychiatry* 10 348 (2020).
- Wilcox, Rand R. *Applying Contemporary Statistical Techniques*. Academic Press, 2003.
- Xie, Lin, Hong Luo, Mei Li, Wenjie Ge, Bingyu Xing, and Qunfang Miao. "The immediate psychological effects of Coronavirus Disease 2019 on medical and non-medical students in China." *International Journal of Public Health* 65, no. 8 (2020): 1445–1453.
- Zhang, Heping, and Burton Singer. *Recursive Partitioning in the Health Sciences*. 1431-8776. Springer-Verlag New York, 1999.
- Zohuri, Bahman, and Masoud Moghaddam. *Neural Network Driven Artificial Intelligence : Decision Making Based on Fuzzy Logic*. Nova Science Publishers Inc, 2017

# Siber Güvenlik Uzmanın Çok Kriterli Karar Verme Yöntemleri ile Seçilmesi

\*\*\*

## Selecting A Cyber Security Specialist with Multi-Criteria Decision-Making Methods

Rabia YUMUŞAK<sup>1</sup> 

Tamer EREN<sup>2</sup> 

DOI:10.33461/uybisbbd.1179062

### Öz

#### Makale Bilgileri

##### Makale Türü:

Araştırma Makalesi

##### Geliş Tarihi:

22.09.2022

##### Kabul Tarihi:

05.12.2022

©2022 UYBİSBBD  
Tüm hakları saklıdır.



Bilgi güvenliği kişilerin, şirketlerin, kurumların ve ülkelerin maddi ve manevi zarar görmemesi için en etkin şekilde yönetilmesi gereken bir süreçtir. Endüstri 4.0 ile farklı bir boyut kazanan bilginin dijital ortama aktarılması ve dijital ortamda işlenebiliyor olması siber güvenlik konusunu kritik bir noktaya taşımaktadır. Çünkü bilginin sahibi olan kişiden ülke ölçeğine kadar güvenliğin sağlanmaması maddi kayıpların yanında geri dönüşü olmayan itibar kayıplarına da yol açmaktadır. Bu sebeplerden dolayı siber güvenlik süreçlerinin optimal bir şekilde yönetilmesi gerekmektedir. Bu kapsamda bu çalışmada siber güvenlik hizmeti veren bir güvenlik şirketine personel alım süreci incelenmiştir. Siber güvenlik uzmanların herhangi bir saldırı anında müdahale etmesi kritiktir. Bununla birlikte sektörde yetişmiş personel bulmak zordur. Sonuç olarak güvenlik şirketlerinin personel alımlarında süreci en iyi şekilde yönetmesi gerekmektedir. Bu ihtiyaç sonucunda problemin çok amaçlı ve çok kriterli yapısı göz önünde bulundurularak çok kriterli karar verme yöntemleri kullanılmıştır. Birinci aşamada Analitik Hiyerarşi Prosesi (AHP) yöntemi ile siber güvenlik uzmanı seçimi kriterleri değerlendirilmiştir. Toplamda 15 kriter dikkate alınmıştır. Problemin hiyerarşik yapısı çözüm yöntemi ile korunmuştur. Ardından sırala algoritması olan PROMETHEE yöntemi ile 10 aday değerlendirilmiştir. Çalışma ele alınan problem, değerlendirilen kriterler ve kullanılan yöntem entegrasyonu açısından literatüre yeni bir bakış açısı kazandırmaktadır.

**Anahtar Kelimeler:** Siber güvenlik, personel yönetimi, personel seçimi, AHP, PROMETHEE.

#### Article Info

##### Paper Type:

Research Paper

##### Received:

22.09.2022

##### Accepted:

05.12.2022

©2022 UYBİSBBD  
All rights reserved.



### Abstract

Information security is a process that must be managed most effectively to prevent material and moral damage to individuals, companies, institutions, and countries. The fact that the information, which has gained a different dimension with Industry 4.0, is transferred to the digital environment and can be processed in the digital domain brings the issue of cyber security to a critical point. Because the lack of security from the person who owns the information to the country scale causes irreversible reputational losses as well as financial losses, for these reasons, it is necessary to manage the cyber security processes optimally. In this context, in this study, the personnel recruitment process of a security company providing cyber security services was examined. Cyber security experts must intervene in case of any attack. However, it is not easy to find trained personnel in the sector. As a result, security companies need to manage the process in the best way in personnel recruitment. As a result of this need, multi-criteria decision-making methods were used considering the multi-purpose and multi-criteria structure of the problem. First, the criteria for cyber security expert selection were evaluated with the Analytical Hierarchy Process (AHP) method. In total, 15 criteria were taken into account. The hierarchical structure of the problem is preserved by the solution method. Then, ten candidates were evaluated with the PROMETHEE method, a sort algorithm. The study brings a new perspective to the literature regarding the problem addressed, the criteria evaluated, and the method integration used.

**Keywords:** Cyber security, personnel management, Personnel selection, AHP, PROMETHEE.

**Atıf/ to Cite (APA):** Yücel, A., (2022). Siber Güvenlik Uzmanın Çok Kriterli Karar Verme Yöntemleri ile Seçilmesi. Uluslararası Yönetim Bilişim Sistemleri ve Bilgisayar Bilimleri Dergisi, 6(2), 117-130

<sup>1</sup> Öğr. Gör. Bilişim Güvenliği Teknolojisi, Kapadokya Meslek Yüksekokulu, Kapadokya Üniversitesi, Türkiye  
rabiayumusak95@gmail.com.

<sup>2</sup> Prof. Dr. Endüstri Mühendisliği Bölümü, Mühendislik ve Mimarlık Fakültesi, Kırıkkale Üniversitesi, Türkiye  
tamereren@gmail.com.

## 1. GİRİŞ

Hızla gelişen teknolojiyle birlikte yeni fırsatlar sunan siber güvenlik ve altyapısında önemli gelişmeler olsa da, siber uzay tamamen güvenli olmaktan uzaktır (Shiva, Roy, & Dasgupta, 2010). Bu durum güvenlik hizmeti sunan firmaları sürekli olarak yeni gelişmeleri takip etmeye ve beraberinde personelinin de bu gelişmelere adapte etmeye itmektedir. Siber güvenlik hizmeti sunan şirketler hızlı bir şekilde değişen teknoloji dünyasında ürünlerini en iyi şekilde sunabilmek ve bu ürünlerin kullanımını sürdürülebilir bir şekilde sağlamak için personel yönetimi süreçlerine ayrıca önem vermek durumundadırlar. Çünkü bütün ürünler personellerin bilgi ve becerisinin teknoloji ile harmanlanmasıyla meydana gelmektedir. Bununla birlikte siber güvenlik alanında yetişmiş personel bulmak zordur. Bu sebepten dolayı şirketler için, personel seçiminden başlayıp personelin yetiştirilme süreci, performansının değerlendirilmesi ve terfi süreçlerine kadar bütün adımları en iyi şekilde yönetmek büyük önem arz etmektedir (Furnell, Fischer, & Finch, 2017). İnsan kaynakları departmanlarında yürütülen personel yönetim süreçlerinin ilk adımı personel seçimidir. Personel yönetimindeki diğer adımların hepsi ilk adım olan personel seçimine bağlıdır. Bu durum personel seçim sürecinin kritikliğini ortaya koymaktadır. Çünkü yanlış personel seçildiğinde personelin eğitim süreçlerinden kuruma olan bağlılığına kadar olan birçok adımda olumsuz süreçler yaşanabilmektedir. Bunun yanı sıra uygulama alanı bilgi güvenliği gibi kritik bir birim olması da problemin önemini kritik noktalara taşımaktadır.

İşletmelerin temel proseslerinde biri olan personel yönetimi birçok karar sürecini içerisinde barındırmaktadır. Personel ihtiyacının belirlenmesindeki sonraki aşama olan personel seçimi süreci personel yönetimi prosesinin en kritik karar süreçleri arasında yer almaktadır. Çünkü yapılan seçim sonucunda personele bir yatırım yapılmaktadır. Yanlış personel seçimi sonucunda yapılan yatırım maliyetinin yanında zaman kaybı hatta itibar zedelenmesine kadar olumsuzluklar yaşanabilmektedir. Bu sebepten dolayı personel seçimi süreci önemlidir (Danişan, Özcan, & Eren, 2022). Problemin öneminden yola çıkarak bu çalışmada bir siber güvenlik şirketi için personel seçimi problemi ele alınmıştır.

Literatürde personel seçimi konusunda yapılmış çalışmalar mevcuttur. Bu çalışmalar da gıda (Chuang, Hu, Liou, & Tzeng, 2020), tekstil (Danişan vd., 2022), lojistik (Nong & Ha, 2021), güvenlik (Kose, Kabak, & Aplak, 2013), turizm (Urosevic, Karabasevic, Stanujkic, & Maksimovic, 2017), hizmet sistemleri (Uslu, Yılmaz, & Yiğit, 2021) ve bu çalışmanın uygulama alanı olan bilişim teknolojileri (Raj Mishra, Sisodia, Raj Pardasani, & Sharma, 2020) gibi birçok sektör ele alınmaktadır. Farklı sektörlerde ele alınan personel seçim probleminin çözüm aşamasında çok kriterli karar verme yöntemleri tercih edilmektedir. Maêda vd. (Maêda vd., 2021) bu çalışmaları derleyerek bir literatür taraması sunmuşlardır. Bilişim teknolojilerini ele alan personel seçimi çalışmaları incelendiğinde AHP (Erdem, 2016), TOPSIS (Samanlioglu, Taskaya, Gulen, & Cokcan, 2018), ARAS (Raj Mishra vd., 2020), EDAS (Karabasevic, Zavadskas, Stanujkic, Popovic, & Brzakovic, 2018) ve SWARA (Heidary Dahooie, Beheshti Jazan Abadi, Vanaki, & Firoozfar, 2018) yöntemlerinin kullanıldığı tespit edilmiştir. Bu çalışmada ise AHP yöntemi ile siber güvenlik uzmanı seçimini etkileyen kriterlerin ağırlıkları hesaplanmıştır. Ardından PROMETHEE yöntemi ile adaylar önceliklendirilmiştir. Literatürde bilişim teknolojileri alanında personel seçimi problemi ele alınmış olsa da siber güvenlik özelinde yapılan bir çalışma bulunmamaktadır. Bununla birlikte çalışmada kullanılan 15 değerlendirme kriteri literatürde ilk defa birlikte kullanılmıştır.

Çalışmanın ikinci bölümünde AHP ve PROMETHEE yöntemleri anlatılmıştır. Üçüncü bölümde Siber güvenlik uzmanı seçimi problemine önerilen çözüm modelinin uygulama adımları ayrıntılı bir şekilde sunulmuştur. Son olarak dördüncü bölümde çalışmanın sonuçları değerlendirilmiştir.

## 2. YÖNTEMLER

Siber güvenlik uzmanı seçimini etkileyen birçok kriter yer almaktadır. Bununla birlikte kriterlerin eşit öneme sahip olduğunu varsayıldığında sonuçlar etkileneceği için bu çalışmada

öncelikle kriterlerin seçim sürecine etkilerini hesaplamak için AHP yöntemi kullanılmıştır. AHP yöntemi çok kriterli karar verme yöntemlerinden olup seçim problemlerinde kriter ağırlıklandırma aşamasında sıklıkla kullanıldığı tespit edilmiştir. Bununla birlikte kriterlerin hiyerarşik yapısı göz önünde bulundurularak AHP yöntemi kullanılmıştır. Ardından adayların değerlendirilmesi için hem kullanım kolaylığı sunması hem de problem yapısında yer alan nicel ve nitel değerlerin birlikte kullanılması şartını sağlayan PROMETHEE yöntemi kullanılmıştır.

## 2.1. AHP Yöntemi

AHP yöntemi ilk olarak 1980 yılında, Thomas L. Saaty (Saaty, 1980) tarafından, çok kriterli kompleks problemlerin çözümü için geliştirilmiş bir yöntemdir. AHP, karar verme prosesinde nitel ve nicel kriterleri karşılaştırabilme özelliğine sahip olduğu için sıklıkla tercih edilmektedir. AHP'nin güçlü yönü, bu yöntemin sayılabilen ve sayılamayan faktörleri sistematik bir yol ile düzenlemesi ve tüm faktörleri dikkate alarak karar verme prosesinde basit ve etkin bir çözüm yolu sunmasıdır (Handfield, Walton, Sroufe, & Melnyk, 2002). Bu sebepten dolayı literatürde kriterlerin önem ağırlıklarının bulunmasında ve çeşitli seçim problemlerinde kullanılmaktadır. Bu çalışmada ise siber güvenlik uzmanı seçimini etkileyen kriterlerin ağırlıklandırılması aşamasında AHP yöntemi kullanılmıştır. AHP yöntemi sağlıktan (Deringöz, Danışan, & Tamer, 2021; Nursena, Yapıcı, Yumuşak, & Eren, 2021) enerjiye (Özcan, Danışan, Yumuşak, & Eren, 2020; Özcan, Yumuşak, & Eren, 2021), ulaşımdan (Paçacı, Serpil, & Çubuk) bilişim teknolojilerine (Yapıcı, Oral, Yumuşak, & Eren, 2021) kadar birçok sektörde kullanılmaktadır. AHP yönteminin uygulama adımları Tablo 1'de verilmiştir (Özcan, Yumuşak, & Eren, 2019).

**Tablo 1:** AHP yönteminin uygulama adımları (Saaty, 1980)

Adım numarası	Açıklama
Adım 1	Karar vericinin amacını, bu amaca etki eden kriterleri ve alternatiflerin de eklenmesiyle bunlar arasındaki ilişkilerin belirlenip hiyerarşik yapının oluşturulmasını içerir.
Adım 2	Uzman kişiler tarafından tüm kriter ve alternatiflerin ikili olarak önem derecelerine göre karşılaştırılması ile gerçekleştirilir. Bu aşamada, Saaty tarafından geliştirilen 1-9 önem skalası kullanılır.
Adım 3	Oluşturulan karşılaştırma matrisi normalize edilir. Bu normalize işlemi her matristeki her değerın sütun toplamına bölünmesi ile oluşur. Böylece normalize matris elde edilir.
Adım 4	Normalize işlemi yapıldıktan sonra hiyerarşideki karşılaştırılan öğelere ilişkin öncelik veya ağırlık vektörlerinin hesaplanması yapılır.
Adım 5	Karar vericinin skaladaki değerleri tutarlı verip vermediğini ölçmek için her bir ikili karşılaştırma matrisindeki yargıların tutarlılık oranının hesaplanması gerekmektedir. Tutarlılığa yakınlık göstergesi olarak nitelendirilen "Tutarlılık indeksi (CI)" hesaplanır.
Adım 6	AHP skorlarının analizi aşamasında ise en yüksek değere sahip alternatif, en iyi alternatif olarak seçilir.

Bu çalışmada AHP yöntemi ile kriter ağırlıkları belirlenmiş olup ardından PROMETHEE yöntemi ile alternatifler değerlendirilmiştir.

## 2.2. PROMETHEE Yöntemi

PROMETHEE, Brans ve Vincke (Thakkar, 2021) tarafından 1985 yılında geliştirilen bir sıralama yöntemidir. PROMETHEE basit bir kavramdır ve kullanımı şu anda uygulama için mevcut olan diğer çok kriterli karar verme tekniklerinden daha kolaydır (Tuzkaya, Gülsün, Kahraman, & Özgen, 2010). Sınırlı sayıda alternatifi olan problemlerde, PROMETHEE çoklu ve karmaşık kriterlere göre sıralama yapmak için avantajlıdır (Dağdeviren, 2008). Ayrıca nitel ve nicel değerlerin birlikte değerlendirilmesine imkan sunan bir yöntemdir. Bu sebeplerden dolayı literatürde sıklıkla kullanılmaktadır. Literatürde enerji (Li, Xu, Wei, Bai, & Liu, 2022), sağlık (Eren, Danişan, Deringöz, & Aksüt, 2022), ulaşım (Sang, Yu, Chang, & Liu, 2022), tekstil (Danişan vd., 2022) ve bilişim teknolojileri (Taşkın, Gezik, Yumuşak, & Eren, 2022; Yaşar, Poyraz, Yumuşak, & Tamer, 2022) alanında uygulamaları bulunmaktadır. Bununla birlikte tedarikçi seçimi (Tong, Wang, & Pu, 2022), tesis yer seçimi (Ayough, Boshruai, & Khorshidvand, 2022), teknoloji seçimi (Taşkın vd., 2022) ve bu çalışmanın konusu olan personel seçimi (Dumnić, Mostarac, Ninović, Jovanović, & Buhmiler, 2022) problemlerinin çözümü için literatürde tercih edilmektedir. PROMETHEE yönteminin uygulama adımları Tablo 2’de özetlenmiştir.

**Tablo 2:** PROMETHEE yöntemi uygulama adımları (Thakkar,2021)

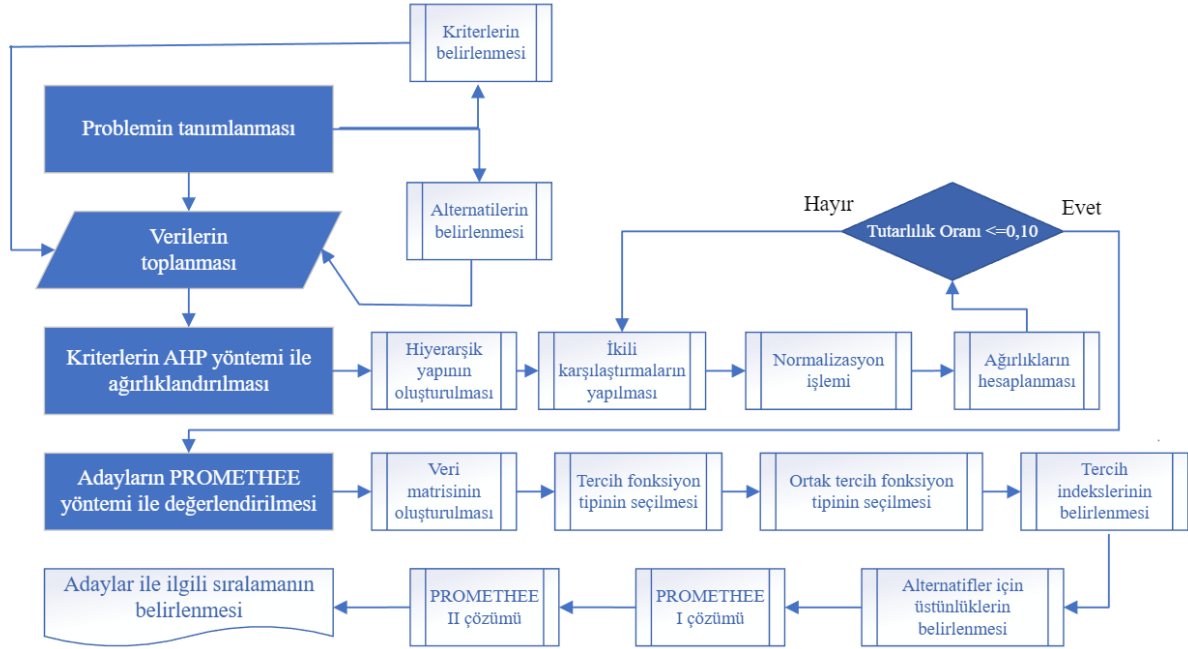
Adım numarası	Açıklama
Adım 1	Alternatifler, kriterler ve bunlara ilişkin ağırlıklar için problemin çözümünde kullanılan veri matrisi oluşturulması.
Adım 2	6 farklı tercih fonksiyonu arasından kriterlerin yapısına göre en uygun fonksiyon tipinin seçilmesi.
Adım 3	Ortak tercih fonksiyonlarının belirlenmesi.
Adım 4	Her alternatif çifti için tercih indeksleri belirlenmesi
Adım 5	Alternatifler için negatif $\Phi^-$ ve pozitif $\Phi^+$ üstünlükler belirlenmesi
Adım 6	PROMETHEE I ile en iyi alternatifin belirlenmesi ve seçimi
Adım 7	PROMETHEE II ile alternatifler için tam önceliklerin belirlenmesi.

Bu çalışmada PROMETHEE yöntemi uygulama aşamasında siber güvenlik uzmanlarının sıralanması için kullanılmıştır. Uygulama ile ilgili detaylı açıklamalar üçüncü bölümde yer almaktadır.

## 3. UYGULAMA

Endüstri 4.0 kavramı ile birlikte hızla dijitalleşen şirketlerin bilgi güvenliğini sağlamak büyük önem arz etmektedir. Bilgi güvenliğinin öneminin her geçen gün artması ve bu problemlere karşı çözüm süreçlerine olan talebin artması, nitelikli işgücüne olan ihtiyacın artması sonucunu doğurmuştur. Bu kapsamda bu çalışmada siber güvenlik alanında hizmet veren bir şirkete personel alımı süreci incelenmiştir. Ele alınan şirkette siber güvenlik uzmanı ihtiyacı ortaya çıkmış olup bu çalışma bu sürecinin optimize edilmesi amaçlanmıştır. Çalışma yedi aşamada gerçekleştirilmiş olup ilk olarak problem tanımlanması, seçimi etkileyen kriterlerin belirlenmesi ve değerlendirmeye alınacak adayların belirlenmesi aşaması gerçekleştirilmiştir. Ardından adaylar ile ilgili veriler

toplanıp uzman görüşleri alınıp AHP yöntemi ile siber güvenlik uzmanı seçimini etkileyen kriterler önceliklendirilmiştir. Kriter önceliklerinin belirlenmesinin ardından PROMETHEE yöntemi ile adaylar değerlendirilmiştir. Son olarak sonuçlar değerlendirilmiştir. Uygulamanın akış şeması Şekil 1’de verilmiştir.

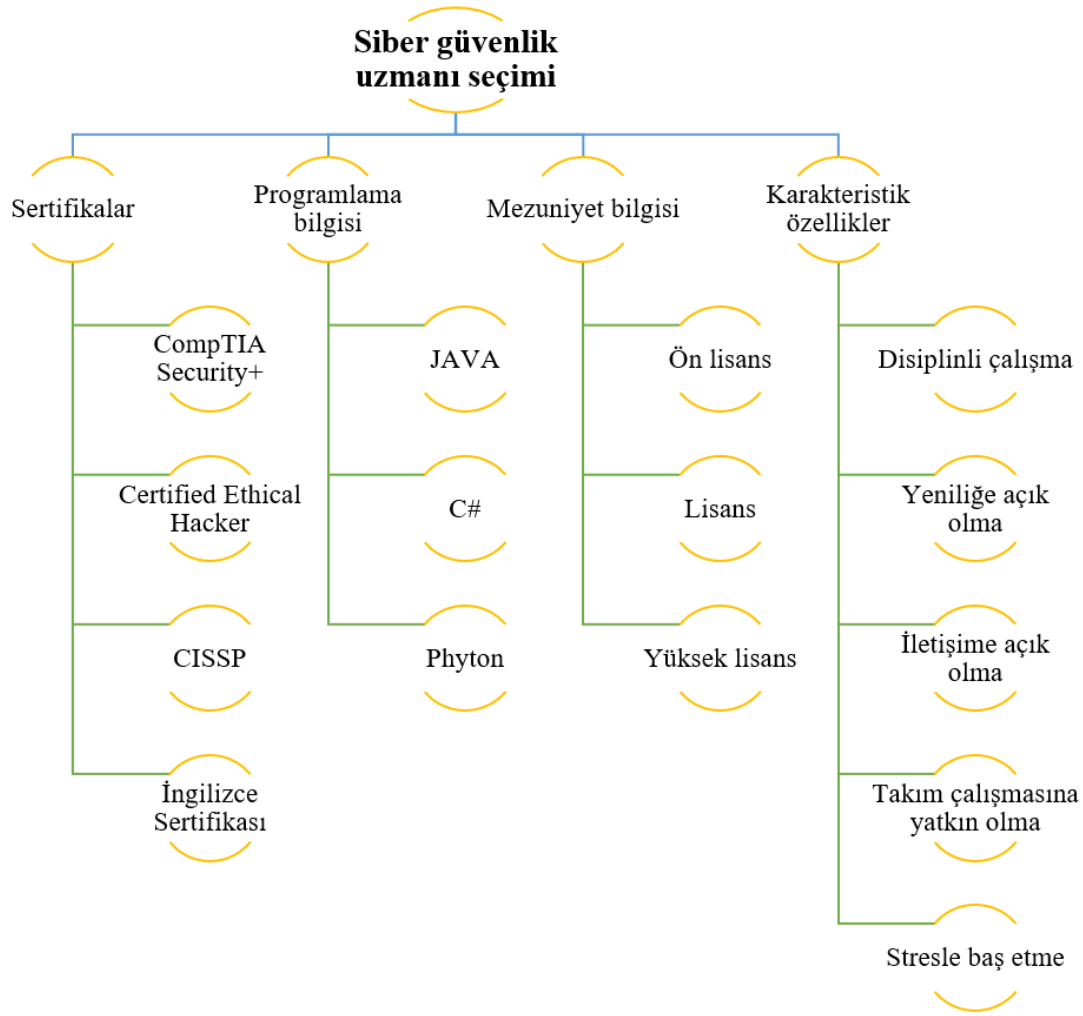


Şekil 1: Uygulama akış şeması

Uygulamanın ilk adımı problemin tanımlanmasıdır. Problemin tanımlanması Bölüm 3.1’de verilmiştir.

### 3.1. Problemin Tanımlanması

Bir siber güvenlik şirketi için personel alımı sürecinin yönetilmesi gerekmektedir. Bu süreçte ilk olarak personel seçimini etkileyen kriterlerin belirlenmelidir. Siber güvenlik uzmanı almak isteyen şirketin taleplerini açık bir şekilde ortaya koyması sürecin en etkin şekilde yönetilmesi açısından kritik öneme sahiptir. Şirketin insan kaynakları birimi ve ilgili departmanın yöneticisi ile kriterler belirlenmiş olup siber güvenlik uzmanı seçimi problemini etkileyen kriterler Şekil 2’de verilmiştir. Bununla birlikte çalışmada 10 aday değerlendirilmiştir.



Şekil 2: Problemin şematik gösterimi

Problemin tanımlanmasının ardından kriterlerin belirlenmesi aşamasına geçilmiştir. Bu aşama Bölüm 3.2’de verilmiştir.

### 3.2. Kriterlerin Belirlenmesi

Siber güvenlik gibi spesifik bir alan için uzman seçimi yapılırken değerlendirme kriterlerinin doğru belirlenmesi kritik öneme sahiptir. Bu çalışma uygulamaya yönelik olduğu için sektöre hakim ve siber güvenlik alanıyla ilgili yetkinlik düzeyini belirleyebilecek uzmanların görüşleri dikkate alınmıştır. Uzmanlardan birisi siber güvenlik alanında hizmet veren şirketin kurucu ortağı olup hem sektörde hem de akademik alanda çalışmaları bulunmaktadır. İkinci uzman ise şirketin insan kaynakları biriminde çalışmaktadır. Son olarak üçüncü uzman da siber güvenlik alanında eğitimler veren bir mühendistir. Üç uzmanın görüşleri alınarak kriterler belirlenmiştir. Kriterler, kodları ve açıklamaları Tablo 3’te verilmiştir.

Tablo 3: Kriterler ve açıklamaları

Ana Kriterler	Alt kriterler	Açıklama
Sertifikalar (K1)	CompTIA Security+ (K1.1) CEH (K1.2) CISSP (K1.3) İngilizce Sertifikası (K1.4)	Siber güvenlik gibi yeni bir alanda adayların yetkinliklerini ispatlaması için bazı sertifikalara ihtiyaç duyulmaktadır. Siber güvenlik alanında en çok tercih edilen üç sertifika ve İngilizce bilgisi sertifikalar ana kriteri altında değerlendirilmiştir.

Programlama Bilgisi (K2)	JAVA (K2.1) C# (K2.2) Phyton (K2.3)	Siber güvenlik alanında yapılacak iş paketlerinde programlama bilgisi vazgeçilmez bir unsurdur. Bu sebepten dolayı JAVA, C# ve Phyton programlama dilleri değerlendirmeye alınmıştır.
Mezuniyet Bilgisi (K3)	Ön lisans (K3.1) Lisans (K3.2) Yüksek Lisans (K3.3)	İnsan kaynakları departmanları işe alım süreçlerinde adayların mezuniyet durumlarını dikkate almaktadırlar. Siber güvenlik uzmanları alanı farklı ön lisans ve lisans programları arasından yetiştirildiği için mezuniyet ana kriteri üç başlıkta incelenmiştir.
Karakteristik Özellikler (K4)	Disiplinli çalışma (K4.1) Yeniliğe açık olma (K4.2) İletişime açık olma (K4.3) Takım çalışmasına yatkın olma (K4.4) Stresle baş etme (K4.5)	Siber güvenlik uzmanları mesleki anlamda donanıma sahip olduğu kadar karakteristik özellikleri ile süreçlerin sonuçlandırılmasında iş bitirici olmaları gerekir. Bu açıdan bakılarak beş kriter karakteristik özellikler ana kriteri altında değerlendirilmiştir.

Kriterlerin belirlenmesinin ardından alternatiflerin belirlenmesi aşamasına geçilmiştir. Bu aşama Bölüm 3.3'te verilmiştir.

### 3.3. Alternatiflerin Belirlenmesi

Uygulamanın yapıldığı güvenlik şirketi personel alım ilanında bilgisayar mühendisliği, bilişim güvenliği teknolojisi, yönetim bilişim sistemleri, bilişim sistemleri mühendisliği ve yazılım mühendisliği bölümlerinden mezun olması şartı aramaktadır. Bununla birlikte uzman kadrosuna ihtiyaç duyulduğu için en az 3 yıl siber güvenlik alanında çalışmış olma şartı da aranmaktadır. Sektörde siber güvenlik alanında yetişmiş personel bulma sorunu olduğu için bu kriterleri sağlayan sadece 10 aday başvuru yapmıştır. Bu 10 aday, siber güvenlik uzmanı seçimi probleminde alternatifler olarak belirlenmiştir.

### 3.4. Verilerin Toplanması

İlana başvuru yapan 10 aday için verilerin toplanması aşaması gerçekleştirilmiştir. Bu aşamada adayların özgeçmişlerinden ve özgeçmişlerinde yer alan referanslardan yararlanılmıştır. Sertifikalar, programlama bilgisi ve mezuniyet bilgisi için özgeçmiş dosyalarına başvurulurken Karakteristik özellikler için referans verilen kişilerden yardım alınmıştır. Tüm bu veri toplama süreçlerinin sonucunda Tablo 4 ortaya çıkmıştır.

**Tablo 4:** Veriler

	Aday 1	Aday 2	Aday 3	Aday 4	Aday 5	Aday 6	Aday 7	Aday 8	Aday 9	Aday 10
K11	var	yok	yok	yok	var	yok	yok	yok	yok	var
K12	yok	yok	yok	yok	yok	yok	var	yok	yok	yok
K13	yok	yok	yok	var	yok	yok	yok	yok	yok	yok
K14	var	yok	yok	var	var	yok	var	yok	var	var
K21	iyi	orta	orta	iyi	orta	orta	kötü	iyi	kötü	iyi
K22	iyi	kötü	iyi	iyi	çok iyi	iyi	iyi	iyi	orta	iyi
K23	çok iyi	iyi	orta	çok iyi	iyi	iyi	iyi	kötü	iyi	iyi
K31	yok	yok	var	yok	yok	var	yok	var	yok	yok
K32	var	var	yok	var	var	yok	var	yok	var	var
K33	var	yok	yok	var	yok	yok	yok	yok	yok	yok
K41	9	6	5	8	7	4	7	5	8	7
K42	7	7	9	5	3	4	3	7	5	4
K43	3	7	6	3	9	6	7	4	7	6
K44	6	4	4	7	5	5	3	9	5	2



K45 | 5 6 3 5 4 7 5 3 6 9

Tablo 4’te sunulan verilerin toplaması aşaması tamamlandıktan sonra kriterlerin AHP yöntemi ile ağırlıklandırılması aşamasına geçilmiştir. Bu aşama Bölüm 3.5’te verilmiştir.

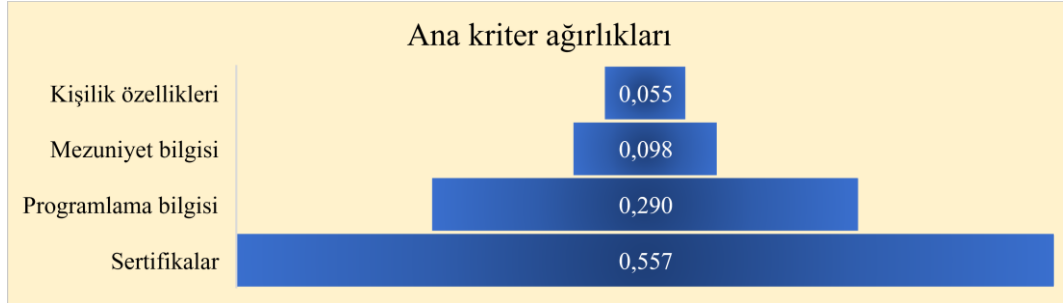
### 3.5. Kriterlerin AHP Yöntemi ile Önceliklendirilmesi

AHP uygulamasında ilk olarak hiyerarşik yapı oluşturulur. Bu çalışmada alternatifler AHP yöntemi ile değerlendirilmediği için hiyerarşik yapı Şekil 2’deki gibidir. AHP yönteminin sadece kriterlerin ağırlıklandırılması aşamaları kullanılmıştır. Bu aşamada veriler üç uzmandan alınmış olup değerlerinin geometrik ortalaması alınıp ortalaması tam sayı olmayan değerler için yuvarlama işlemi yapılmıştır. İlk olarak ana kriterler ağırlıklandırılmıştır. Ana kriterlerin değerlendirildiği ikili karşılaştırma matrisi Tablo 5’te verilmiştir.

**Tablo 5:** Ana kriterler için ikili karşılaştırma matrisi

Ana Kriterler	Sertifikalar	Programlama bilgisi	Eğitim bilgisi	Kişilik özellikleri
<b>Sertifikalar</b>	1	3	5	8
<b>Programlama bilgisi</b>	0,333	1	4	6
<b>Eğitim bilgisi</b>	0,200	0,250	1	2
<b>Kişilik özellikleri</b>	0,125	0,167	0,500	1

Tablo 5’te verilen ikili karşılaştırma matrisi üzerinde AHP adımları uygulandığında Şekil 3’te verilen sonuçlara ulaşılmıştır. Tutarlılık oranı ise 0.083 çıkmış olup sonucun tutarlı olduğu anlaşılmaktadır.



**Şekil 3 :** Ana kriter ağırlıkları

Ardından sırayla her ana kritere ait alt kriterler kendi arasında ikili karşılaştırmalara tabi tutulur. Sertifikalar ana kriterine ait alt kriterlerin ikili karşılaştırma matrisi Tablo 6’da verilmiştir.

**Tablo 6:** Sertifikalar kriterleri için ikili karşılaştırma matrisi

	CompTIA Security+	Certified Ethical Hacker (CEH)	CISSP	İngilizce Sertifikası
<b>CompTIA Security+</b>	1	3	0,333	2
<b>Certified Ethical Hacker (CEH)</b>	0,333	1	0,143	0,5
<b>CISSP</b>	3	7	1	6
<b>İngilizce Sertifikası</b>	0,5	2	0,167	1

Tablo 6’da verilen ikili karşılaştırma matrisi üzerinde AHP adımları uygulandığında Şekil 4’te verilen sonuçlara ulaşılmıştır. Tutarlılık oranı ise 0.021 çıkmış olup sonucun tutarlı olduğu anlaşılmaktadır.



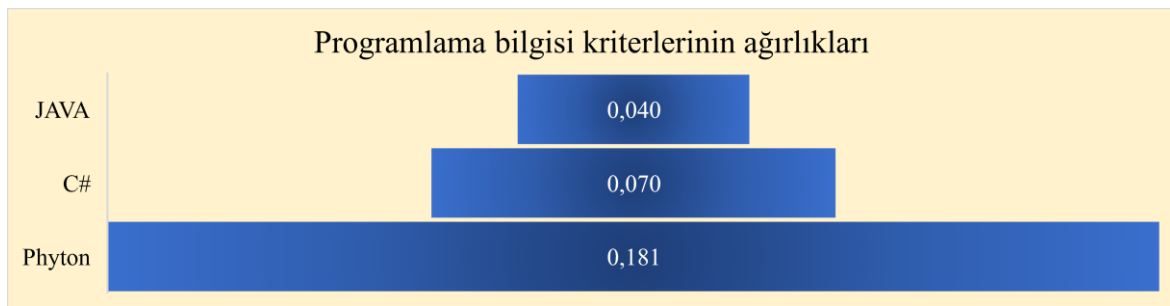
Şekil 4: Sertifikaların kriter ağırlıkları

Diğer bir ana kriter olan programlama bilgisi için ikili karşılaştırma matrisi oluşturulmuştur. Oluşturulan matris Tablo 7’te verilmiştir.

Tablo 7: Programlama bilgisi kriterlerinin ikili karşılaştırma matrisi

	JAVA	C#	Phyton
JAVA	1	0,5	0,250
C#	2	1	0,333
Phyton	4	3	1

Tablo 7’te verilen ikili karşılaştırma matrisi üzerinde AHP adımları uygulandığında Şekil 5’te verilen sonuçlara ulaşılmıştır. Tutarlılık oranı ise 0.029 çıkmış olup sonucun tutarlı olduğu anlaşılmaktadır.



Şekil 5: Programlama bilgisinin kriterlerinin ağırlıkları

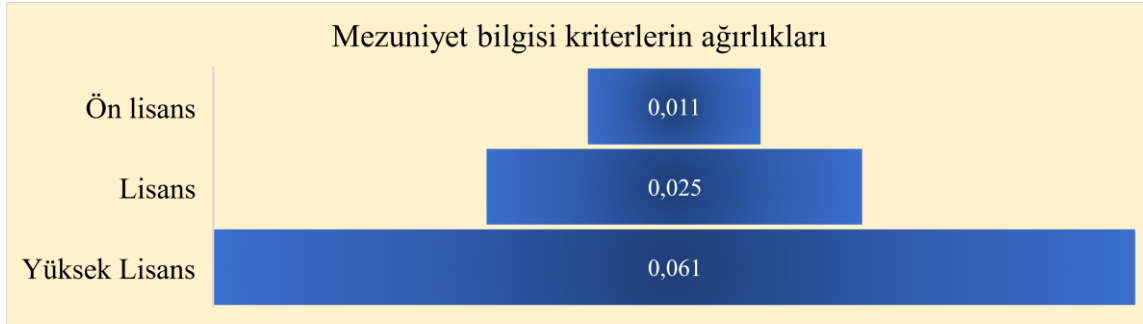
Üçüncü ana kriter olarak Mezuniyet bilgisi ele alınmıştır. Mezuniyet bilgisi ana kriterine ait alt kriterlerin ikili karşılaştırma matrisi Tablo 8’de verilmiştir.

Tablo 8: Mezuniyet bilgisi kriterinin ikili karşılaştırma matrisi

	Ön lisans	Lisans	Yüksek Lisans
Ön lisans	1	0,25	0,143

<b>Lisans</b>	4	1	0,333
<b>Yüksek Lisans</b>	7	3	1

Tablo 8’de verilen ikili karşılaştırma matrisine AHP adımları uygulandığında Şekil 6’da verilen kriter ağırlıkları ortaya çıkmıştır. Tutarlılık oranı 0.054 çıkmıştır. Sonucun tutarlı olduğu tespit edilmiştir.



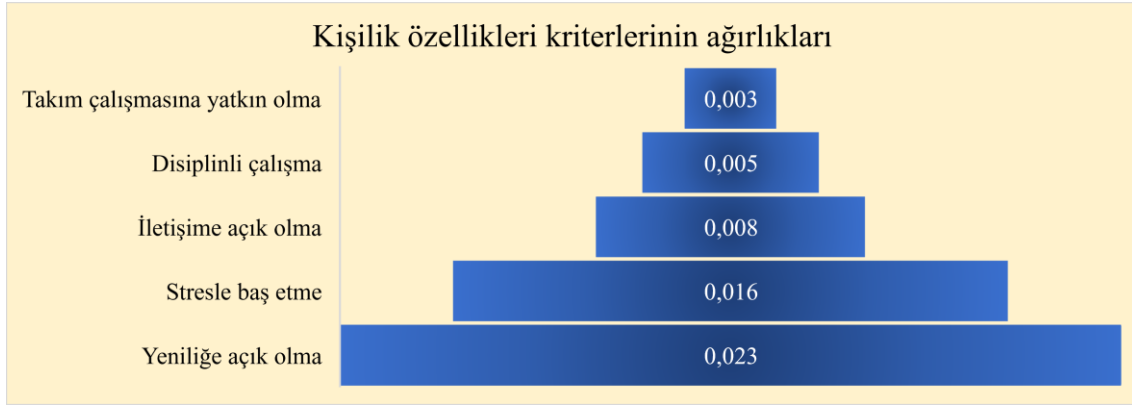
Şekil 6: Mezuniyet bilgisi kriterlerinin ağırlıkları

Son olarak kişilik özellikleri ana kriterine ait alt kriterlerin ikili karşılaştırma matrisi Tablo 9’da verilmiştir.

Tablo 9: Kişilik özellikleri kriterlerinin ikili karşılaştırma matrisi

	<b>Disiplinli çalışma</b>	<b>Yeniliğe açık olma</b>	<b>İletişime açık olma</b>	<b>Takım çalışmasına yatkın olma</b>	<b>Stresle baş etme</b>
<b>Disiplinli çalışma</b>	1	0,167	0,50	4	0,2
<b>Yeniliğe açık olma</b>	6	1	3	5	2
<b>İletişime açık olma</b>	2	0,333	1	3	0,5
<b>Takım çalışmasına yatkın olma</b>	0,25	0,2	0,333	1	0,143
<b>Stresle baş etme</b>	5	0,5	2	7	1

Tablo 9’de verilen ikili karşılaştırma matrisine AHP yönteminin adımları uygulandığında Şekil 7’de verilen sonuçlar elde edilmiştir. Tutarlılık oranı 0.093 çıkmış olup sonucun tutarlı olduğu tespit edilmiştir.



Şekil 7: Kişilik özellikleri kriterlerinin ağırlıkları

### 3.6. Adayların PROMETHEE Yöntemi ile Değerlendirilmesi

AHP yöntemi ile kriter ağırlıklarının belirlenmesi aşamasından sonra PROMETHEE yöntemi ile adaylar değerlendirilmiştir. Adaylar ile ilgili veriler Tablo 4'te yer almaktadır. Tablo 4'te sunulan veriler ve kriter ağırlıkları kullanılarak PROMETHEE yöntemi çözülmüştür. Şekil 8'de verilen bölümde çözüm için gerekli olan kriter ağırlıkları, veriler ve kriterlerin fonksiyonları girilmiştir.

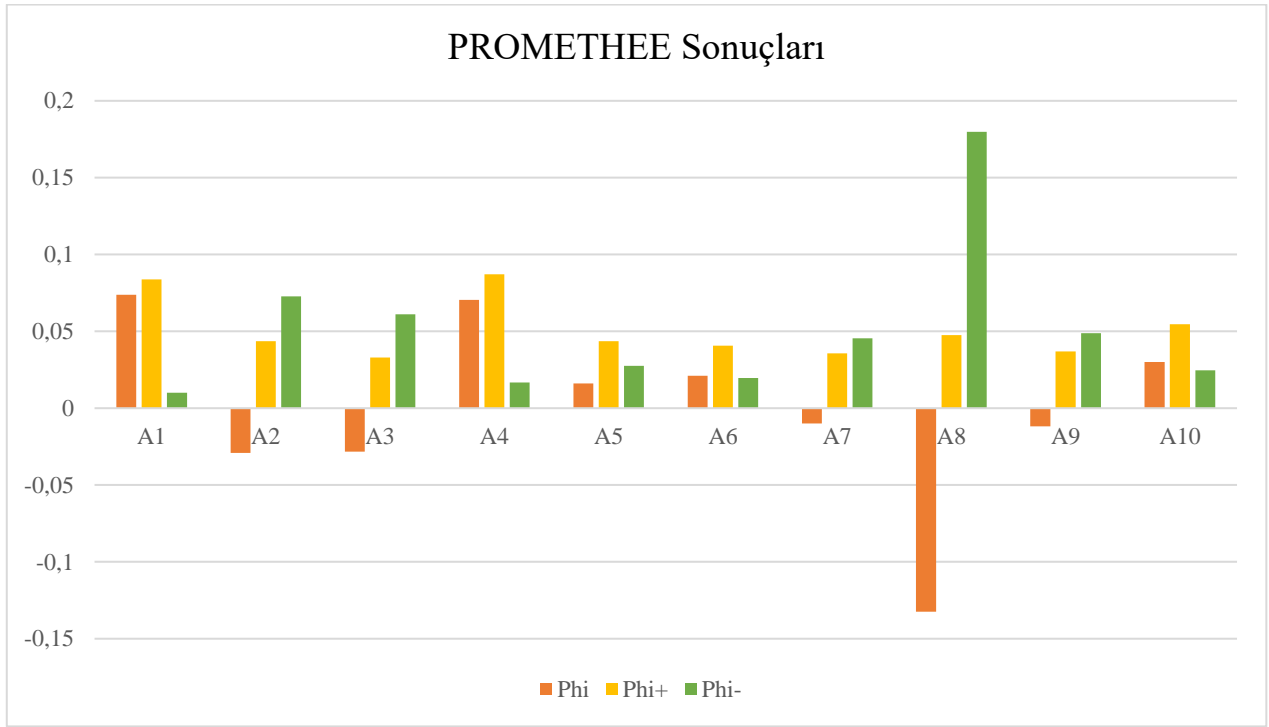
Visual PROMETHEE Academic - 01.09.22.vpg (saved)

File Edit Model Control PROMETHEE-GAIA GDSS GIS Custom Assistants Snapshots Options Help

Scenario1	K11	K12	K13	K14	K21	K22	K23	K31	K32	K33	K41	K42	K43	K44	K45
Unit	unit	unit	unit	y/h		5-point	5-point	unit	unit	unit	impact	unit	unit	unit	unit
Cluster/Group	◆	◆	◆	◆	◆	◆	◆	◆	◆	◆	◆	◆	◆	◆	◆
<b>Preferences</b>															
Min/Max	max	max	max	max	max	max	max	max	max	max	max	max	max	max	max
Weight	0,12	0,04	0,33	0,07	0,04	0,07	0,18	0,01	0,03	0,06	0,01	0,02	0,00	0,02	0,02
Preference Fn.	U-shape	U-shape	U-shape	U-shape	Level	Level	Level	U-shape	U-shape	U-shape	Linear	Linear	Linear	Linear	Linear
Thresholds	absolute	absolute	absolute	absolute	absolute	absolute	absolute	absolute	absolute	absolute	absolute	absolute	absolute	absolute	absolute
- Q: Indifference	1,00	1,00	1,00	1,00	1,00	1,00	1,00	1,00	1,00	1,00	1,00	1,00	1,00	1,00	1,00
- P: Preference	n/a	n/a	n/a	n/a	2,00	2,00	2,00	n/a	n/a	n/a	2,00	2,00	2,00	2,00	2,00
- S: Gaussian	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a
<b>Statistics</b>															
Minimum	0,00	0,00	0,00	0,00	2,00	2,00	2,00	0,00	0,00	0,00	4,00	3,00	4,00	2,00	3,00
Maximum	1,00	1,00	1,00	1,00	4,00	5,00	5,00	1,00	1,00	1,00	8,00	9,00	9,00	9,00	9,00
Average	0,30	0,10	0,10	0,60	3,20	3,90	3,90	0,30	0,70	0,20	6,40	5,40	6,10	5,00	5,40
Standard Dev.	0,46	0,30	0,30	0,49	0,75	0,83	0,83	0,46	0,46	0,40	1,28	1,91	1,45	1,90	1,74
<b>Evaluations</b>															
action1	yes	no	no	yes	good	good	very good	no	yes	yes	7,00	7,00	4,00	6,00	5,00
action2	no	no	no	no	average	bad	good	no	yes	no	6,00	7,00	7,00	4,00	6,00
action3	no	no	no	no	average	good	average	yes	no	no	5,00	9,00	6,00	4,00	3,00
action4	no	no	yes	yes	good	very good	very good	no	yes	yes	8,00	5,00	5,00	7,00	5,00
action5	yes	no	no	yes	average	very good	good	no	yes	no	7,00	3,00	9,00	5,00	4,00
action6	no	no	no	no	average	good	good	yes	no	no	4,00	4,00	6,00	5,00	7,00
action7	no	yes	no	yes	bad	good	good	no	yes	no	7,00	3,00	7,00	3,00	6,00
action8	no	no	no	no	good	good	bad	yes	no	no	5,00	7,00	4,00	9,00	3,00
action9	no	no	no	yes	bad	average	good	no	yes	no	8,00	5,00	7,00	5,00	6,00
action10	yes	no	no	yes	good	good	good	no	yes	no	7,00	4,00	6,00	2,00	9,00

Şekil 8: PROMETHEE veri giriş ekranı

Şekil 8'de verilen bölüm doldurulduktan sonra PROMETHEE sonuçlarına ulaşılmıştır. 10 adayın değerlendirildiği modelden elde edilen sonuçlar Şekil 9'da verilmiştir.



Şekil 9: PROMETHEE yönteminden elde edilen sonuçlar

Sonuçlar incelendiğinde Aday 1'in ilk sırada yer aldığı ve şirketin Aday 1'i işe alması gerektiği sonucuna varılmıştır. Ardından Aday 4 gelmektedir. Aday 4, Aday 1'in değerlerine çok yakın olduğu için güçlü bir rakip olduğu anlaşılmaktadır. Bu çalışmada personelin maaş beklentisi dikkate alınmamıştır. Bu sebepten dolayı şirket Aday 1 ve Aday 4 arasında maaş kriterini de dikkate alarak bir seçim yapabilir.

#### 4. SONUÇ ve ÖNERİLER

Siber güvenlik, globalleşen Dünya'da ilginin giderek arttığı ve öneminin kritik seviyelere ulaştığı bir konudur. Bununla birlikte insan faktörü siber güvenlik süreçlerinin yürütülmesinde en önemli kollardan biridir. Çünkü siber güvenlik uygulamalarında meydana gelen krizlerde konunun uzmanı olan kişilerin kritik müdahaleleri oluşabilecek maddi ve manevi kayıpların önüne geçmektedir. Bu sebepten dolayı siber güvenlik alanında personel seçimi problemi büyük önem arz etmektedir.

Bu çalışmada siber güvenlik hizmeti veren bir şirkete uzman personel alım süreçleri ele alınmıştır. Öncelikle siber güvenlik uzmanı seçimini etkileyen kriterler literatür taraması ve uzman görüşleri doğrultusunda belirlenmiştir. Ardında bu kriterler konunun uzmanı kişilerle AHP yöntemi uygulaması yapılarak değerlendirilmiştir. Çünkü her kriterin personel seçimini etkileme seviyesi farklıdır. Elde edilen kriter ağırlıkları sıralama algoritması olan PROMETHEE yönteminde kullanılarak adaylar değerlendirilmiştir. Yapılan değerlendirme sonucunda Aday 1'in siber güvenlik uzmanı olarak işe alınması gerektiği gözlemlenmiştir. Elde edilen sonuçlar doğrultusunda Aday 4'ün de güçlü bir aday olduğu sonucu dikkat çekmektedir. Yapılan çalışmada maaş ile ilgili kriterler dikkate alınmamıştır. Bu sebepten dolayı siber güvenlik şirketi güçlü iki aday arasında maaş politikasına en uygun olanı tercih edebilir.

Personel seçimi problemi literatürde ele alınan bir konu olmakla birlikte siber güvenlik alanında yapılmış bir uygulama bulunmamaktadır. Bu sebepten dolayı bu çalışmanın yapılan literatür taraması sonucunda ele alınan problem tanımı ve kullanılan kriterler açısından literatüre katkı sağlar nitelikte olduğu tespit edilmiştir. Siber güvenlik alanında personel seçimi ile ilgili literatür zayıftır. Literatürdeki bu eksikliği gidermek için farklı çalışmalar yapılabilir. Bu çalışmada kriterler uzman görüşleri ile belirlenmiş olup ileriki çalışmalarda bu çalışmadaki kriterler de göze alınarak yeni kriterler eklenip kriterler bazında literatüre katkısı analiz edilebilir. Örneğin ele alınan kriterler

arttırılabilir. Kullanılan yöntem kombinasyonları çeşitlendirilerek sonuçlar karşılaştırılabilir. Siber güvenlik alanında personel seçimi problemi için yeni modeller önerilerek modellerin performansı değerlendirilebilir. Siber güvenlik alanı sürekli gelişen ve yeniliklerin hız kesmeden yaşandığı bir alan olduğu için personellerin eğitim planlaması yapılabilir. Son olarak siber güvenlik uzmanlarının performansları çok kriterli karar verme yöntemleri ile ölçülebilir.

## KAYNAKÇA

- Ayough, A., Boshruai, S., & Khorshidvand, B. (2022). A new interactive method based on multi-criteria preference degree functions for solar power plant site selection. *Renewable Energy*, 195, 1165-1173. Elsevier.
- Chuang, Y.-C., Hu, S.-K., Liou, J. J., & Tzeng, G.-H. (2020). A data-driven MADM model for personnel selection and improvement. *Technological and Economic Development of Economy*, 26(4), 751-784.
- Dağdeviren, M. (2008). Decision making in equipment selection: An integrated approach with AHP and PROMETHEE. *Journal of intelligent manufacturing*, 19(4), 397-406. Springer.
- Danişan, T., Özcan, E., & Eren, T. (2022). Personnel Selection with Multi-Criteria Decision Making Methods in the Ready-to-Wear Sector. *Tehnički vjesnik*, 29(4), 1339-1347. Strojariski fakultet u Slavonskom Brodu; Fakultet elektrotehnike, računarstva ....
- Deringöz, A., Danişan, T., & Eren, T. (2021). Covid-19 takibinde giyilebilir sağlık teknolojilerinin ÇKKV yöntemleri ile değerlendirilmesi. *Politeknik Dergisi*, 1-1.
- Dumnić, S., Mostarac, K., Ninović, M., Jovanović, B., & Buhmiler, S. (2022). Application of the Choquet Integral: A Case Study on a Personnel Selection Problem. *Sustainability*, 14(9), 5120. MDPI.
- Erdem, M. B. (2016). A fuzzy analytical hierarchy process application in personnel selection in it companies: A case study in a spin-off company. *Acta Physica Polonica A*, 130(1), 331-334. Institute of Physics, Polish Academy of Science.
- Eren, T., Danişan, T., Deringöz, A., & Aksüt, G. (2022). Comparison and selection of patient follow-up systems for covid-19 pandemic patients. *Fashion and Textiles*, 9(1), 1-13. SpringerOpen.
- Furnell, S., Fischer, P., & Finch, A. (2017). Can't get the staff? The growing need for cyber-security skills. *Computer Fraud & Security*, 2017(2), 5-10. Elsevier.
- Handfield, R., Walton, S. V., Sroufe, R., & Melnyk, S. A. (2002). Applying environmental criteria to supplier assessment: A study in the application of the Analytical Hierarchy Process. *European journal of operational research*, 141(1), 70-87. Elsevier.
- Heidary Dahooie, J., Beheshti Jazan Abadi, E., Vanaki, A. S., & Firoozfar, H. R. (2018). Competency-based IT personnel selection using a hybrid SWARA and ARAS-G methodology. *Human Factors and Ergonomics in Manufacturing & Service Industries*, 28(1), 5-16. Wiley Online Library.
- Karabasevic, D., Zavadskas, E. K., Stanujkic, D., Popovic, G., & Brzakovic, M. (2018). An approach to personnel selection in the IT industry based on the EDAS method. *Transformations in Business & Economics*, 17, 54-65.
- Kose, E., Kabak, M., & Aplak, H. (2013). Grey theory based MCDM procedure for sniper selection problem. *Grey systems: Theory and application*. Emerald Group Publishing Limited.
- Li, P., Xu, Z., Wei, C., Bai, Q., & Liu, J. (2022). A novel PROMETHEE method based on GRADEMA<sup>TEL</sup> for PLTSs and its application in selecting renewable energies. *Information Sciences*, 589, 142-161. Elsevier.
- Maêda, N., Rodrigues, M. V. G., Ângelo, M., Moreira, L., Gomes, C. F. S., & d dos Santos, M. (2021). Bibliometric Studies on Multi-Criteria Decision Analysis (MCDA) Applied in Personnel Selection. *Modern Management Based on Big Data II and Machine Learning and Intelligent Systems III: Proceedings of MMBD 2021 and MLIS 2021*, 341, 119. IOS Press.

- Nong, N.-M. T., & Ha, D.-S. (2021). Application of MCDM methods to Qualified Personnel Selection in Distribution Science: Case of Logistics Companies. *Journal of Distribution Science*, 19(8), 25-35. Korea Distribution Science Association.
- Nursena, O., Yapıcı, S., Yumuşak, R., & Eren, T. (2021). Pandemi sürecinde sürdürülebilir tedarik zinciri yönetimi için ilaç deposu ve aşı dağıtım merkezi yeri seçimi. *Politeknik Dergisi*, 1-1.
- Ozcan, E., Danisan, T., Yumusak, R., & Eren, T. (2020). An Artificial Neural Network Model Supported With Multi Criteria Decision Making Approaches For Maintenance Planning In Hydroelectric Power Plants. *Polish Maintenance Soc.*
- Özcan, E., Yumuşak, R., & Eren, T. (2019). Risk based maintenance in the hydroelectric power plants. *Energies*, 12(8), 1502. Mdpi.
- Özcan, E., Yumuşak, R., & Eren, T. (2021). A novel approach to optimize the maintenance strategies: A case in the hydroelectric power plant. *Eksploatacja i Niezawodność*, 23(2).
- Paçacı, B., Serpil, E., & Çubuk, K. (Basımda). Çok Modlu Taşımacılığa Uygun Lojistik Merkez Yer Seçimi İçin Bir Öneri: Türkiye Uygulaması. *Politeknik Dergisi*, 1-1.
- Raj Mishra, A., Sisodia, G., Raj Pardasani, K., & Sharma, K. (2020). Multi-criteria IT personnel selection on intuitionistic fuzzy information measures and ARAS methodology. *Iranian Journal of Fuzzy Systems*, 17(4), 55-68. University of Sistan and Baluchestan.
- Saaty, T. L. (1980). *The analytical hierarchy process, planning, priority. Resource allocation.* RWS publications, USA.
- Samanlioglu, F., Taskaya, Y. E., Gulen, U. C., & Cokcan, O. (2018). A fuzzy AHP–TOPSIS-based group decision-making approach to IT personnel selection. *International Journal of Fuzzy Systems*, 20(5), 1576-1591. Springer.
- Sang, X., Yu, X., Chang, C.-T., & Liu, X. (2022). Electric bus charging station site selection based on the combined DEMATEL and PROMETHEE-PT framework. *Computers & Industrial Engineering*, 168, 108116. Elsevier.
- Shiva, S., Roy, S., & Dasgupta, D. (2010). Game theory for cyber security. *Proceedings of the Sixth Annual Workshop on Cyber Security and Information Intelligence Research* (ss. 1-4).
- Taşkın, E., Gezik, N., Yumuşak, R., & Tamer, E. (T.Y.). Depo Yönetiminde Endüstri 4.0 Uygulaması: Bir İşletme için RFID Teknoloji Seçimi. *Endüstri Mühendisliği*, 33(1), 194-211.
- Thakkar, J. J. (2021). Preference Ranking Organization Method for Enrichment Evaluations (PROMETHEE). *Multi-Criteria Decision Making* (ss. 119-127). Springer.
- Tong, L. Z., Wang, J., & Pu, Z. (2022). Sustainable supplier selection for SMEs based on an extended PROMETHEE II approach. *Journal of Cleaner Production*, 330, 129830. Elsevier.
- Tuzkaya, G., Gülsün, B., Kahraman, C., & Özgen, D. (2010). An integrated fuzzy multi-criteria decision making methodology for material handling equipment selection problem and an application. *Expert systems with applications*, 37(4), 2853-2863. Elsevier.
- Urosevic, S., Karabasevic, D., Stanujkic, D., & Maksimovic, M. (2017). An Approach To Personnel Selection In The Tourism Industry Based On The Swara And The Waspas Methods. *Economic Computation & Economic Cybernetics Studies & Research*, 51(1).
- Uslu, Y. D., Yılmaz, E., & Yiğit, P. (2021). Developing qualified personnel selection strategies using MCDM approach: A university hospital practice. *Strategic Outlook in Business and Finance Innovation: Multidimensional Policies for Emerging Economies.* Emerald Publishing Limited.
- Yapıcı, S., Oral, N., Yumuşak, R., & Eren, T. (2021). Blokzincir Teknolojisi ile Merkezi ve Dağıtık Veri Tabanının Karşılaştırılması. *Endüstri Mühendisliği*, 32(3), 457-472.
- Yaşar, S., Poyraz, Z., Yumuşak, R., & Eren, T. (2022). ANP ve PROMETHEE Yöntemleri ile Akıllı Şehir Analizi: Ankara’da Bir Uygulama. *Gazi Mühendislik Bilimleri Dergisi*, 8(1), 15-28.

# Sanal Gerçekliği ve Artırılmış Gerçekliği Bir Sonraki Seviyeye Taşımak: Karma Gerçeklik ile Yapay Zeka

\*\*\*

## Taking Virtual Reality and Augmented Reality to The Next Level: Artificial Intelligence with Mixed Reality

DOI:10.33461/uybisbbd.1190125

Ahmet EFE<sup>1</sup> 

### Öz

#### Makale Bilgileri

##### Makale Türü:

Derleme Makalesi

##### Geliş Tarihi:

16.10.2022

##### Kabul Tarihi:

05.12.2022

©2022 UYBİSBBD  
Tüm hakları saklıdır.



*Sanal gerçeklik (VR) ve artırılmış gerçeklik (AR) önemli bir heyecan vermişti, ancak Karma gerçeklik (MR) hayal ettiklerinin ötesinde otantik bir deneyim sunmakta ve bu heyecan verici deneyimi bir sonraki seviyeye taşımaktadır. MR, sanal verileri doğal nesnelere denkleştirerek ve gerçekçi fiziksel ve duygusal etkileşimler oluşturarak, dijital nesnelere gerçek ortamlarda doğal nesnelere gibi konumlandırılmakta ve gerçek nesnelere hizalanarak gerçek zamanlı olarak gerçekleşme sağlanabilmektedir. Metaverse'in ortaya çıkmasıyla, gerçeklik kavramları hem işletmeler hem de akademi tarafından yeniden tanımlanmakta ve vurgulanmaktadır. Metaverse, fiziksel dünyadaki varlığımızın bir kopyasını ve alternatifini oluşturarak var olabileceğimiz sanal ve geçişli bir yapı olarak tanımlandığından, ticari değer ve yüksek teknoloji yönlerinin yanı sıra mistik, manevi ve dini bakış açıları içerebilir. Bu çalışmada multidisipliner analizler ve değerlendirmeler ile sadece fiziksel sanala geçişin değil, aynı zamanda verilerin fiziksel konuma, nesnelere veya sensörlere aktarılması yoluyla sanaldan fiziksele geçişin sağlanabileceği sorgulanmaktadır.*

**Anahtar Kelimeler:** Karma Gerçeklik, Yapay Zeka, Sanal Gerçeklik, Artırılmış Gerçeklik, Metaverse.

### Abstract

#### Article Info

##### Paper Type:

Review Paper

##### Received:

16.10.2022

##### Accepted:

05.12.2022

©2022 UYBİSBBD  
All rights reserved.



*Virtual reality (VR) and augmented reality (AR) have caused significant excitement, but Mixed reality (MR) offers an authentic experience beyond what they have are imagined and it takes this exciting experience to the next level. MR collides virtual data with natural objects, creating realistic physical and emotional interactions so that digital objects are positioned like natural objects in real environments and positioning takes place in real-time by aligning with real things. With the advent of Metaverse, the reality concepts are being redefined and highlighted by both businesses and academia. Apart from business value and high-tech aspects, Metaverse can include mystical, spiritual, and religious perspectives, as it is defined as a virtual and transitive structure in which we can exist by creating a duplicate of our existence in the physical world and its alternative. Through multidisciplinary analyzes and evaluations in this study, the possibility to provide not only the transition from the physical to the virtual, but also the transition from the virtual to the physical through transferring the data to the physical location, objects, or sensors is questioned.*

**Keywords:** Mixed Reality, Artificial Intelligence, Virtual Reality, Augmented Reality, Metaverse.

**Atıf/ to Cite (APA):** Efe, A., (2022). Taking Virtual Reality and Augmented Reality to The Next Level: Artificial Intelligence with Mixed Reality. International Journal of Management Information Systems and Computer Science, 6(2), 78-91

<sup>1</sup> Dr., International Federation of Red Cross and Red Cross, ahmet.efe@ifrc.org.



## 1. INTRODUCTION

The use of MR, VR, AI and AR are rapidly developed with innovative cutting-edge applications in many fields such as design, education, entertainment, military education, and health services. MR also includes auxiliary sectors in its usage areas. Thus, the field of design is dramatically changing, expanding, and evolving. These technologies will provide people receiving instruction in a physical environment with the opportunity to access the needs that the domain does not have virtually. The fact that the objects accessed are close to reality will increase the chance to experience in education. VR, AR, and MR applications aim to overcome the limitations of the physical environment in theory and can provide instant interaction with other physical environments at the speed of light. Unbelievably rapid developments in the innovative dimension of scientific applications do not yet provide definitive evidence on whether this will enable spiritual interaction or whether it will enable communication with angels, demons, and spirits of dead people or result in denial of all past arguments of metaphysical values and elements. For this reason, it is thought that future innovations in AR, VR and MR applications together with super AI should be evaluated. It is assumed that if such an opportunity and functionality is obtained, religious perceptions, faith and worship may be adversely affected. There has not been a study at the literature level on this issue yet. Therefore, this is a research question is set as such: "*Will VR, AR and MR innovative technology lead to interaction with spirit realms, jinn and angels when supported by super AI or result in complete denial of metaphysical values?*"

In this study based on the latest publications in the literature, we have started with setting up research problem and literature information; provided key conceptual definitions of VR, MR and AR; evaluated Security and privacy issues and key risks; briefed different applications in education, military, health, entertainment, and business; discussed the Metaverse with conclusion.

## 2. RESEARCH PROBLEM AND LITERATURE

When it is searched through Scholar database with the keywords "Mixed Reality, Artificial Intelligence, Virtual Reality, augmented Reality, Metaverse and spirituality" 217 publications are found none of which used all terms in the article header. We have found no study assessing possibility of interaction with spiritual realm or denial of all.

In fact, it is all about the conceptualization, theory and positioning of reality related paradigms. The transformation of reality perceived by sense organs into information redefines the concept of reality. The concept of reality determines what exists. The fact that perception can be influenced and misleading in its relationship with memory opens new dimensions in the concept of reality (Ipek, 2020). Many equivalents of MR technology emerged with the blending of VR and AR technology (Yengin and Bayrak, 2017). The key elements of VR had been discussed long before. The HTC Vive Pro Eye is the most equipped example of this technology, developed considerably since the first VR device called "Sensorama Simulator," invented by Morton Heilig in 1962. This headset, the best-known member of MR technologies, includes VR and AR technology, and can follow the user's pupil movements with external eye-tracking technology. This situation dramatically increases the success in the sensitivity of the use of technology. As it is known, all the VR technologies that make it possible to overcome the limits of physical reality involve the integration of the user into the virtual environment through various sensors and computers (Milgram & Kishino, 1994; Grau, 2003; Scoble & Isreal, 2017). After that, the user can navigate the environment positioned with head movements. This function, which is possible due to the communication between the head and the sensors through various methods, is being strengthened using innovative techniques. The newest technique introduced in this context is the eye-tracking technique, which detects the eyeball movements of the user, processes metric and navigation information, and thus increases the integration of the user with the objects in the virtual environment (Bayrak and Yengin, 2020)

With Industry 4.0 applications, the artificial functions of MR have also increased and diversified, as virtual and accurate data have collided in a similar environment in MR. With this

method, while increasing the reality of virtual data, it is ensured that the existing natural environment is transformed into settings where virtual data can be used. Transformations also shape the perception of reality. This pass-through in perceptions of reality causes immersion, distortions in individuals' perception of reality. The next stage of this deterioration is that the individual forgets that he is in a virtual or physical environment. It is among researchers' ideas that the developments in this field can bring along various mental and physical problems. These efforts intersect with those offered by the Internet of Things. This mutual transformation causes many technologies such as digital twins, blockchain, and artificial intelligence to have new use areas. MR technology has been applied in diverse manufacturing and industrial sectors to visualize the digital models supporting teaching and training. Nevertheless, most of these MR projects mainly focus on users' experiences (Tang *et al.*, 2019).

Swalley (2015) aims to determine the effect of Metaverse and augmented reality concepts on online purchasing behavior. Retailing has transformed from traditional to electronic and then to virtual reality in which qualitative and quantitative methods were used together to determine the basic view of technology and consumer interest in retailing. In a study conducted by Ağralı and Aydın (2021), the emotions related to Metaverse were evaluated in the light of the data obtained from social media platforms. The tweets sent one week before and after Mark Zuckerberg's announcement that his company would change its name to Meta were evaluated and tweets sent in English with the hashtag "Metaverse" on Twitter were used as the data set. As a result of the study, it was explained that after Mark's speech, positive tweet rates decreased and both neutral and negative tweet rates increased. Damar (2021) aimed at a bibliometric evaluation of metaverse technology, which has been discussed in the literature since the 1990s. The study was carried out by extracting the data of all documents between the years 1990-2021 from the Web of Science base, and it was deduced that Metaverse will be included in many areas of human life in 15 to 20 years and will shape life by taking advantage of the opportunities of developing technology. Kalkan (2021) aimed I his study to obtain inferences about the development and future of sports and the phenomenon of sports in the Metaverse universe, which is a new concept. While examining the concept of metaverse chronologically, he conceptually discussed the issues of being a coach, athlete, referee, and spectator in the Metaverse era. According to the results obtained from the study, the Metaverse universe was found to be important in terms of sports and its stakeholders in the context of the development of sports and exercise. Kuş (2021) aimed to determine the factors that are effective in the emergence of facts about the concept of Metaverse in his study. Within the scope of the study, the positive and negative factors were listed, and the results of the perception of opportunity and threat were explained. Nalbant and Uyankı (2021) aimed to examine Metaverse technology both in computer vision and general point of view. In addition, they emphasized that these technologies should be more accessible to the society, suitable for people at all economic levels of the organization, and cheap cost products should be developed and made suitable for use by all people. Ozkahveci et al. (2022), on the other hand, predicted that the concept of Metaverse increases every year throughout the world and within the scope of the relevant period, and the world can adapt quickly to this new virtual society.

As it is evident from the above studies in the literature, it is possible for computers to have the capabilities mentioned above and to develop themselves, thanks to the AI technology. As with AR, privacy is a significant concern for VR. The highly personal nature of the data collected, namely biometric data such as iris or retina scans, fingerprints and handprints, face geometry, and voiceprints, is a primary VR privacy concern (Kaspersky, 2021):

- *Eye-tracking*: Data can be valuable to malicious individuals. Knowing precisely what he or she is looking at can help an attacker reveal the user's valuable information, thereby recreating user actions.
- *Finger tracking*: The system records and transmits fingerprint data, indicating fingers typing a PIN code. If the attacker can capture this data, they can regenerate the user's PIN.

Anonymizing VR and AR tracking data are nearly impossible, as people have unique movement patterns. VR and AR tracking data such as zip codes, IP addresses, and voice traces should also be considered potential "Personally Identifiable Information" (PII). These may be considered PII as external parties may use this information to identify or track an individual's credentials. Canadian Chanelle Siggins was recently harassed verbally and physically by another player while using a VR headset. Siggins warned the player about this incident. The actress who harassed Chanelle Siggins, on the other hand, continued the harassment by ignoring the warnings and said, "*I will do what I want*". After a while, the abusive player walked away from Siggins. With Metaverse and VR, this event raises concerns that cyberbullying, which is already a severe problem in the cyber world, will reach dangerous levels. According to the research, violations disturb the players every 7 minutes in VR-based games (Frenkel and Browning, 2021).

The person may be more obese because they can handle everything in the fantasy world. Loneliness and obesity await humanity in the Metaverse. As a result, many new types of diseases may emerge. It is imperative to consider all these with their pros and cons under academic studies. If used for good purposes, technology serves good and evil if used for immoral purposes. Instead of being against technology, we need to think about using it following our ideals. The problems are now in question in the metaverse universe based on AR, VR, and AI. Institutions and companies are in a big race as the concept of Metaverse requires this, it is about how you can work effectively. These are the questions that the concept of Metaverse deals with:

- Which imagination will they bring to the metaverse universe?
- What kind of experiences will other people face?

When we start talking about these questions, the Metaverse will find its place in real terms. We suffer a lot from the game addiction of young people, and there are problems such as school refusal in young people. In such a situation, a new field of gaming addiction emerges. In this virtual world, shopping will be with cryptocurrencies and bitcoins. Challenging experiences in life that need to be overcome, processes that take time and require patience, and the desire to get something quickly attract them the most. In addition to psychological problems such as addiction, there are the main problems to be encountered such as:

- Inability to cope with anxiety, difficulty in emotion management,
- Underdeveloped social communication and problem-solving skills,
- Different physical problems such as obesity, joint and muscle problems due to inactivity, and
- Real-life economic problems due to procrastination, neglect, and not exerting the necessary effort.

### 3. KEY CONCEPTUAL DEFINITIONS

VR is a technology that is still developing, as well as an increasing number of scientists from different fields, are working on it. Different definitions made by Stone (1991) and Oppenheim (1993) also summarize other definitions. According to Stone (1991), VR is multimedia that appeals to human senses, developed to increase communication between humans and machines. According to Oppenheim (1993), VR is a technology that tries to increase human-machine interaction through feeling, not contented with visual and auditory communication (Kurbanoglu, 1996).

Imagination and reality have never been intertwined. For years, VR has been the "next big thing". Still, it has finally emerged to create stunningly realistic images, sounds, and other emotions amid an extraordinary fantasy world. Almost everyone has access to these fascinating technologies, but before pouring hard-earned money into a cutting-edge head-mounted display, let us take a closer look at what is needed for an amazing VR, AR, or MR experience. Technology has caught on, and market researchers have predicted rapid growth for the VR industry. Look-in toys were popular in the 1950s, and indoor flight simulators were first introduced in the 1960s, but the idea of VR goes

even further. As early as the 1940s, inventors, science fiction writers, and fixers dreamed of an environment that could escape reality through art and machinery. VR can be used as a general term to describe similar but different technologies to the actual VR experience. The similarities and differences between VR, AR, and MR should be apparent (Intel, 2021).

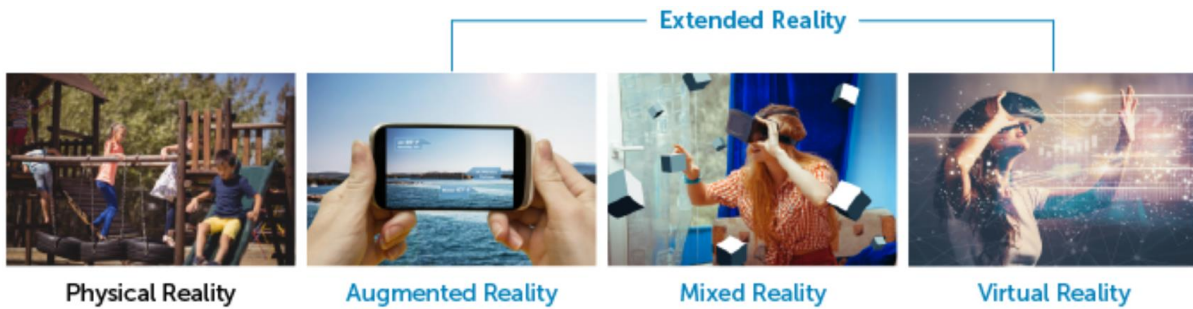


Figure 1. Different pictures for Physical, Augmented, Mixed and Virtual Realities

Source: (Siglin, 2018)

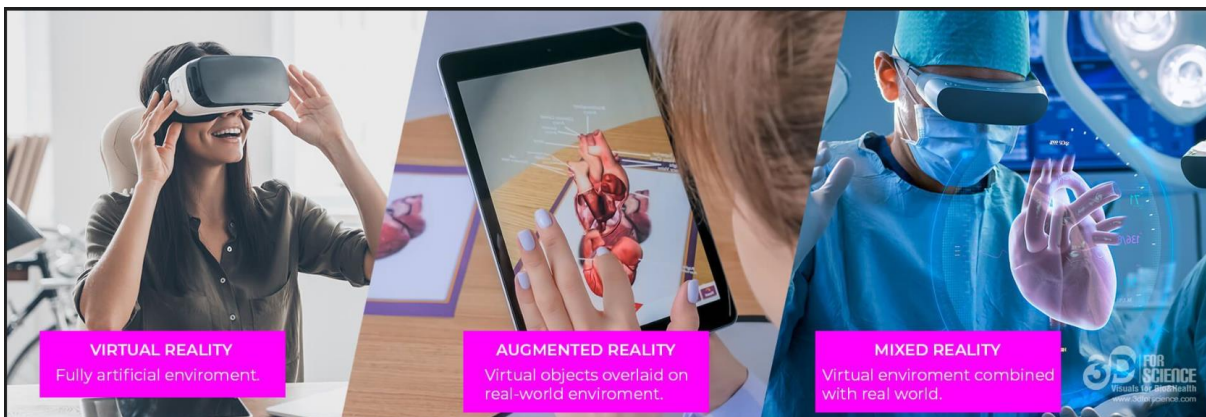


Figure 2. Different pictures for Augmented, Mixed and Virtual Realities

Source: (3Dforscience, 2020)

### *Virtual Reality*

VR is the most widely known technology that is so gripping that it tricks senses into thinking in a different setting or a world apart from the real world. One can move or manipulate objects with touch controls while connected to the console or computer using a head-mounted display (HMD) or headset. One will experience a computer-generated world of video and sound.

### *Augmented Reality*

AR places digital information on top of real-world elements. Pokémon GO\* is one of the best-known examples. AR centers on the natural world but enriches it by layering it with other digital details, a new layer of perception, supporting one's reality or environment (Cem, 2020).

### *Mixed Reality*

Using next generation sensing and imaging technologies in MR, one can interact with and manipulate physical and virtual elements and environments (Sen *et al.*, 2021). MR allows one to see the world he/she lives in and immerse in it, even if someone interacts with a virtual environment by using his hands without ever removing his headset. It breaks down the core concepts between reality and fantasy, delivering an experience that can change the way they play or work today (Holonext, 2019). There are head systems with different features developed by many commercial companies for the use of MR technology. The most known and used among these head systems are Microsoft HoloLens (Bray *et al.*, 2018), Samsung HMD Odyssey (Carbotte, 2018), Lenovo Explorer (Carbotte, 2018; Borandağ & Yücalar, 2020).

VR Hardware VR systems enable users in the virtual world to interact physically with artificial warnings they direct. The interaction can be visual and auditory and through different senses such as tactile, smell, and taste. Specially designed gloves and clothes covering the whole body create these feelings. Above these, some sensors and stimuli allow the system to detect the user's movements and the warnings coming from the system. These stimulating systems are defined as "VR hardware" (Ferhat, 2016). Games to movies and medicine, VR, AR, and MR is becoming more and more common:

- **Games**—Over 2,000 games currently in use, from first-person shooter (FPS) games to strategy games and role-playing adventures
- **Health**—In education, such as in surgical simulations
- **Professional sports**—In training programs like STRIVR that assist professional and amateur athletes<sup>2</sup>
- **Film and TV**—Creating unique experiences in movies and shows
- **Virtual Travel**—Virtual trips to an art museum or another planet that one can take from home

#### 4. SECURITY AND PRIVACY ISSUES AND KEY RISKS

One of the issues that people see as the most significant danger of AR is privacy is that AR collects much more information about who the user is and what they are doing than social media networks or other technologies. User privacy is at risk as AR technologies can see what the user is doing. This raises several concerns and questions (Kaspersky, 2021):

- How do AR companies use the information they collect from users and ensure security?
- If hackers manage to gain access to a device, the risk of loss of privacy is enormous.
- Do AR companies share this data with third parties? So, how do these companies use this data?
- Where do companies store their AR data? Locally on devices or in the cloud? Is data sent to the cloud encrypted?

##### *Untrusted content*

Advanced hackers can replace the user's AR with one of their own, misleading people or giving them false information. This brings reliability issues, as AR is a relatively new field, and authenticated content creation and delivery mechanisms are still being developed. Some cyberthreats can render content unreliable even if the source is genuine. These types of cyber threats include spoofing, eavesdropping, and data manipulation (Kaspersky, 2021).

##### *Social engineering*

Given the possible reliability of the content, AR systems can be an effective tool to deceive users as part of social engineering attacks. For example, hackers can distort users' perceptions of reality through false signs or screens, leading them to take action that will benefit hackers (Kaspersky, 2021).

##### *Malware*

AR hackers can insert malicious content into apps through advertising. Unaware users can compromise AR security by clicking ads that redirect to malicious websites or malware-infected AR servers that serve unreliable images.

---

<sup>2</sup> For more information see: <https://www.strivr.com/>

### *Stealing network credentials*

For merchants using AR and VR shopping apps, hacking can be a cyber threat. Criminals can steal network credentials from Android-powered wearables. Many customers' card information and mobile payment solutions are already registered in their user profiles. Since mobile payment is a straightforward procedure, hackers can access and hijack these accounts.

### *Denial of service*

Another possible AR security attack is a denial of service. In such a case, for example, users who rely on AR applications for their work may suddenly become disconnected from the information flow. For example, a surgeon may suddenly lose access to crucial real-time information from their AR glasses, or a driver whose AR-enabled windshield suddenly turns into a black screen may lose sight of the road (Kaspersky, 2021).

### *Hacking attacks between two ports*

Network attackers can eavesdrop on the AR browser, AR provider, AR channel owners, and third-party servers. These types of attacks are called spying attacks between two ports.

### *Ransomware*

Hackers can access a user's AR device and record their behavior and interactions in the AR environment. This can be embarrassing or unsettling for people who do not want their games and other AR interactions to be publicly known. By adding features designed explicitly to VR platforms, attackers can persuade users to give their personal information. As with AR, it creates a favorable environment for ransomware attacks.

### *Fake IDs or "deepfakes"*

Machine learning technologies make it possible to manipulate audio and video to look like original footage. Since it is possible to access motion tracking data from a VR headset, a hacker could use this data to create a digital copy (also known as a "[deepfake](#)") and aim to circumvent VR security for this purpose. It may be possible to embed it in someone else's VR experience to perform a social engineering attack then. Other problems with VR that critics have identified as negative aspects of VR include (Kaspersky, 2021):

- Health effects include dizziness, nausea, or lack of spatial awareness.
- The potential for excess.
- Loss of connection with people.

## **5. VR APPLICATIONS IN EDUCATION**

Although VR is an expensive technology, it is used in many fields for educational purposes, albeit limited. It is also possible to test factors such as color, lighting, and ergonomics beforehand. Students can do the work they want on cadavers created in virtual environments, or they can make a journey to the human body. Again, it can be used effectively for training students in foreign language education. A student who wants to learn a foreign language can live in that country by traveling to the country where they want to learn it through virtual education (Kayabaşı, 2005).

There is growing evidence that simulation in teaching is necessary to improve learning, skills, and outcomes, efficient skills. In the health sciences, high-fidelity task trainers are ideal for reducing cognitive load and enhancing learning outcomes (Birt *et al.*, 2017). It is expected to be widely used in academic and business education fields. For example, some parts of the training program developed by Microsoft and Japanese Airlines were introduced to the whole world in July 2016. It shows how an airplane jet engine comes in front of the user with hologram images and how the user enlarges, reduces, or rotates this image in the desired directions by using his fingers as a mouse. People using the headgear above can also verbally intervene in the virtual images before them. E.g.,



the MR system reacts very quickly to the command to show the ventilation and cooling system of the jet engine and can instantly bring the desired design. There is no need to go to the hangar where that massive jet engine is located to receive this training. In addition, they do not divide those jet engines, which are very expensive, into parts for educational purposes and present them to the students' examinations. MR technology overcomes all these difficulties. Simulations such as all kinds of laboratory tests and examinations in the battlefields in history lessons will attract the students' attention, reduce the learning time and make learning easier (Özdemir and Öztürk, 2022).



Figure 3. Augmented reality in education Source: (Cem, 2021)

## 6. MEDICAL EDUCATION AND PRACTICE

A versatile MR educational environment is a ubiquitous approach to highly impactful, low cost, high reusability gamified educational approach. Using prolific technologies such as inexpensive AR headsets and a versatile, low maintenance Database (DB) back ends, a natural world environment can be transformed with 3d graphics and audio to various educational spaces of highly impactful content (Antoniou *et al.*, 2016). Indeed, the applications to be used in medicine will significantly increase the quality of education. The failure rate of doctors who can see all the organs of their patients with hologram <sup>3</sup>copies, diagnose them with the help of artificial intelligence, produce organs with new generation printers and transplant them will decrease to almost zero.

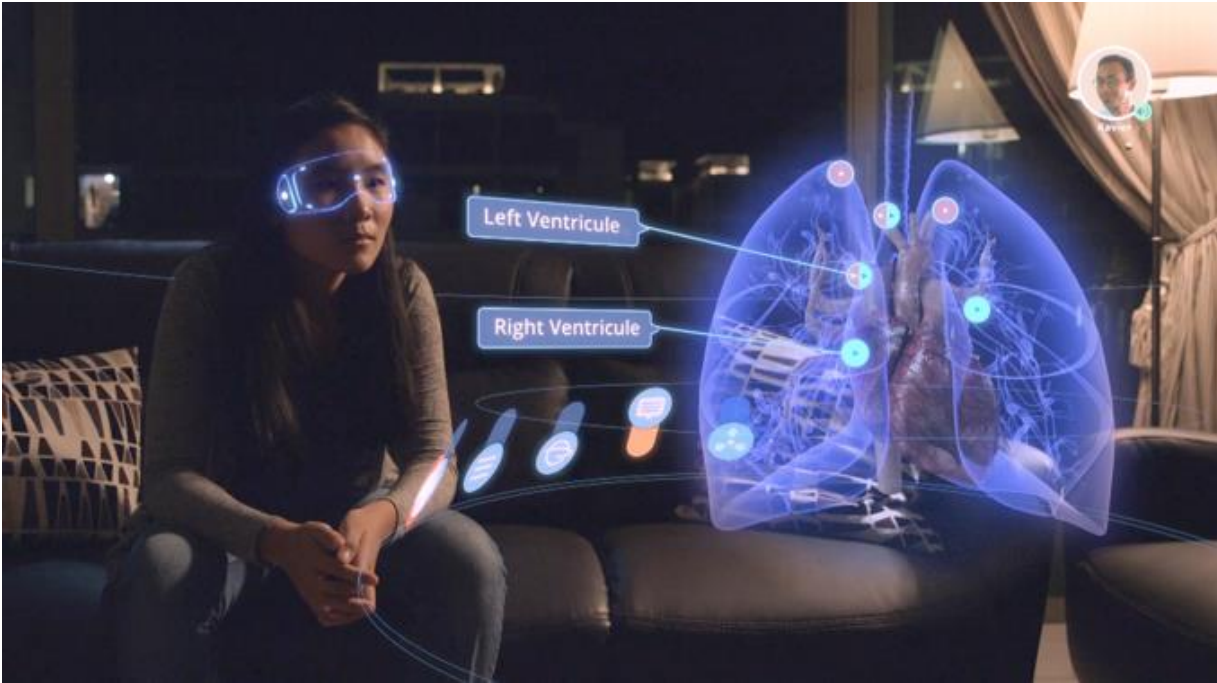
Despite some hardware and technical deficiencies, VR technologies are used to treat many diseases and conditions with programs and interfaces; It is used as an assessment tool to detect heart rate variability and the initial stages of Alzheimer's. At the same time, patients can check their health status at home without going to the hospital by connecting with smartphone applications. The use of VR technologies and other mobile health technologies will increase the functionality of VR technologies and provide important developments in health services (Wiederhold, 2016). As VR technologies decrease production costs and become tools that everyone can use through simple interfaces in terms of use, they will be used more in the education of medicine and other health professions. In this way, students will have an education system to learn the procedures and other vital elements of treating patients in a virtual world or through simulators. As a result, since they can acquire specific skills safely, they will demonstrate their skills when they meet with patients (Reznek,

<sup>3</sup> For more information see: <https://www.microsoft.com/tr-tr/hololens/hardware>

Harter & Krummel, 2002). These technologies, which are used in surgery, treatment, rehabilitation, and education in health services, also benefit from other developments in technology and provide convenience to patients and healthcare professionals in doing their jobs (Demirci, 2018).

## 7. DESIGN PRACTICES

Interaction in a VR environment is the most natural interaction type, and the participants were eager to use both MR and VR environments instead of an emulator (Ergün *et al.*, 2019). The most crucial feature that distinguishes MR applications from AR is the interaction of users with virtual images. Those in different places can see the product under development through their titles and provide and see the changes in real-time with the instructions they will give to the people who manage the project. Alternatively, imagine that the officials of an architectural office evaluate a project drawn with hologram technology with colleagues remotely from different parts of the world. While his Japanese colleague requested the elevator be relocated, the architect in America might want the balconies to be slightly more prominent. Such changes can be made instantly so that everyone can see them in 3D. All kinds of products can be produced error-free and optimal using similar applications during the development phase.



**Figure 4.** Augmented Reality in Medical Trainings *Kaynak: (Cem, 2021)*

## 8. ENTERTAINMENT APPS

VR glasses do not yet have good ergonomics for the physical comfort of the users. Especially VR glasses, in which mobile phones are used, are heavy. However, glasses that use their screens will become much smaller and lighter by time. Studies on this subject continue VR glasses in regular glasses or contact lenses. These glasses or contact lenses can be considered a technological revolution. They make it easier to watch all other digital media products such as television, books, newspapers, movies, and advertisements anytime and anywhere. It will not take long for the movie industry to adapt to VR technology. Thanks to this technology, movies with 360° images to be shot and other factors designed will allow the VR phenomenon to spread to the world quickly. Currently, the scientific world is discussing the development of VR technologies with their positive aspects and some adverse social and psychological side effects. VR will pose a significant danger to some as a mind control method in the future. The adverse effects of the event's moral dimension have divided sociologists, psychologists, and even clergy into two. The products of companies that produce sexual content for adults also cause private life and family concepts to be questioned (Ferhat, 2016).



The quality of AI (i.e., speech recognition and synthesis via machine learning) associated with an augmented object increases MR immersion associated with spatial immersion, MR enjoyment, and consumers' perceptions of novel experiences (Sung *et al.*, 2021). Digital technologies used to entertain people have now gone far beyond gaming applications. The sense of fun and excitement provided by games played with VR or AR headsets/glasses used until recently has been taken to new dimensions. One can watch a concert of a favorite singer, a theater performance, or a sports event from own location in three dimensions with holographic images. Now, instead of going to entertainment venues, entertainment will come to the home.

Digital cinema technology is closely related to the development of computer games. For example, some games such as Tomb Raider, Assassins Creed, Mortal Kombat, which were liked by the masses, were shot as feature films. Likewise, computer games also benefit from cinema in terms of visual effects, framing techniques, and story design. Digital technologies give a new direction to the search for reality in cinema. Amateur video shoots, movies, and movie screenings in which the audience can be included in the process show the need for a new definition of reality. The fact that the platforms, especially Youtube, where amateurs can share their designs, appeal to the audience on a massive scale, shows that movie-watching habits have also changed. Shortly, it is predicted that artificial intelligence will play an active role in all filmmaking processes. We are very close to watching the films in which the scenario will be created by artificial intelligence according to the audience's desire, and the visual and sound effects are presented as options. Even computer games produced by artificial intelligence control the user's experience process. 21st-century cinema has significant advantages in creating and publishing virtual worlds/characters. Many intermediaries of the sector, especially the monopolies of distribution companies, will be adversely affected by this situation. However, new business areas and needed new professions (with high technology skills) will emerge (Barbaros, 2020).

VR sex is a new sex kind. Technological innovations in sex define our future generations and change how we express intimacy and love and how we see and relate to ourselves and others. (Alexandrovich and Gomes, 2020). It allows the user to get tactile sensations or fictional characters from distant participants using computer-controlled sex toys. (Kiiro, 2020). Usually, the user also wears a VR headset to see and interact with the partner (Varina, 2021). So, this type of AI based MR is ostensibly fantastic but, can also severely harm human reproduction, social, psychological and spiritual structures and humanistic values.

## 9. MILITARY TRAINING

Advances in technology paved the way for war simulations in military personnel training. These trainings are not limited to only attack and defense issues. All the training required by military personnel can now be given through these technologies (Rathnayake, 2018). The most popular of these games were war games. Many of these games were developed under the supervision and supervision of military personnel (Cem, 2021).

### *Aviation*

Digital technologies can also be used in the aviation industry, which requires care. AR, VR, and MR technologies can be used in aircraft design to performance evaluation and production line. Today, Airbus uses VR technologies to show aircraft to customers. The European Space Agency (ESA), on the other hand, uses AR technologies in maintenance and plans to use MR and AR technologies in the training of Mars flight personnel.

### *Tourism*

VR technologies provide unique opportunities, especially for holiday lovers. With VR glasses, one can visit virtually any city in the country you want and visit the streets and museums. Today, some tourism companies also allow wandering around the region with VR glasses to explore holiday destinations. On the other hand, AR-based smartphones can guide you by recognizing the objects around you during your vacation (Avzal, *et al.*, 2022).

### *Industry*

In addition to entertainment, these applications also come to the fore with increasing performance and increasing operational efficiency, especially in the industry. Display technologies, which make it easy to access relevant data, enable people and other devices to use resources better. It also reduces the time required, cost, and potential errors. Today, Ford and Audi VR use AR technologies to test different parameters to find the most suitable and efficient products.

### *Partnerships*

AR and MR-based devices also support collaboration. Architects, engineers, and designers can quickly try different parameters with these technologies and see the results. Therefore, reality technologies are expected to break new ground in projects that require teamwork.

### *Installation and Maintenance*

AR is becoming a tool that simplifies the most complex tasks for maintenance and installers. Broken or broken parts can be determined and repaired with the information displayed on the screen for the device held where the operation is performed. It is sufficient for technicians to use head apparatus or tablet for this process. AR guidance can also be used under challenging assemblies in the production area.

### *Marketing*

AR and MR technologies also play an essential role in promoting products to customers. A vehicle dealer or a clothing store can easily use these devices to try different products and colors. As a result, VR, AR, and MR technologies will change lifestyles soon. These technologies will be used in the areas mentioned above and many new areas, so a future you cannot even imagine will be waiting for us. Maybe humanity will completely escape from reality and shelter in the virtual world.

### *Remote work in business*

Collaborative MR technologies enable remote people to work together by sharing communication cues intrinsic to face-to-face conversations, such as eye gaze and hand gestures (Yang *et al.*, (2020). The shared space typically consists of a video capture of the remote workspace displayed on a 2D screen (Huang *et al.*, 2018). Vocational training is frequently held for recruits and existing personnel in medium and large companies. In previous generation technologies, the activity was made available as pre-recorded images, and the user could not ask questions at the time of use. This training can now be done in real-time and interactively between the parties with hologram images. Employees who complete a five-minute training session with such a MR program are said to achieve the same level of success as reading a 50-page training manual. The said technology can be used efficiently in all areas of the business world, from paperwork at the desk to the maintenance and repair of production vehicles, from production planning to logistics and shipping. Of course, the usage area of Mixed Facts is not limited to the examples I have listed above. We will undoubtedly start to use these new digital technologies and applications in all areas of our lives shortly (within a few years) (Cem, 2021).

## **10. DISCUSSION ON METAVERSE**

The definition of Metaverse is "offering a new reality, the world of meaning and collaboration opportunities; providing infrastructure and interaction opportunities for cultural, intellectual and economic production; different advanced technologies are used simultaneously and in an integrated manner; cyber social plane" (Kuş, 2021). Most people describe Metaverse as the next stage of the internet. It entered this virtual universe with virtual and AR devices alongside a personal computer, thus having the opportunity to experience other people or virtual objects directly with the five senses. (Doko, 2021). Although the epidemic process we are experiencing has increased its popularity with Facebook's conversion of the company name to "Meta" and announcing its investments in Metaverse, Metaverse is not a new concept. Despite its popularity, it is not possible to make a clear definition of

what Metaverse is and what it covers. Metaverse seamlessly integrates the real world with the virtual world and allows avatars to carry out rich activities, including creation, display, entertainment, social networking, and trading (Yang et al., 2022). With the "Metaverse" and the Cyber World Order, which is planned to be implemented in 2023, gas, electricity, and power plants and distribution centers, banks, government offices, population, and land registry records may be exposed to cyber-attacks and chaos may occur with "Cyber Terrorism." Everything is being prepared for the "Metaverse," and the products are registered with Blockchain (Aydin, 2021).

Most of the expectations about the Metaverse consist of technologies that we can experience today, but some of them have not yet been massified. In the world of "Second Life", where users create an avatar for themselves and live an alternative virtual life, it is possible to go to school, start a business, earn income, socialize with people, and invest. It is even possible to earn money through Second Life and transfer it to real life.

Metaverse's persistence and interoperability will provide users with more excellent continuity of identity and experience compared to the modern internet. In Metaverse, users will not need separate Twitter profiles, "Fortnite" characters, and Reddit accounts. They will be themselves on all channels. This continuity of identity will be the critical factor behind how users buy and consume content on the Metaverse. There are still quite a few obstacles to a true Metaverse. The biggest hurdles are also hardware limitations. Still, less than 20 percent of Americans are familiar with VR headsets, according to a 2020 report by Thrive Analytics and ARtillery Intelligence. Also, experts predict that devices like VR and AR have a chance to surpass game consoles by 2025 at the earliest. If the hardware and energy technology are sufficient, extensive cultural changes will be necessary to foster an actual Metaverse development.

Many Metaverse projects are still in their infancy. Again, augmented, and virtual reality technologies, or AR, still have a long way to go. In addition to their technological development, they also need to be cheap and mass-produced. It is not new for applications to create their economies in the virtual world. It is possible to watch a concert or a football match while wearing VR glasses and to watch the area where your eyes focus, not the camera, in the 360-degree virtual universe - unlike watching it on a computer and television. Again, online training or meetings can be held in the virtual reality universe. Wearable technology has also come a long way in recent years. With the increasing internet speeds, the internet of things (IoT) is entering our lives more and more every day. With smart homes and innovative city applications, objects can communicate and perform their predetermined tasks without human intervention. With AR applications, it is also possible to buy the shoes you like on phone by first holding the phone's camera on your feet, trying them on your feet, and seeing how they look. Among all these technologies, how Blockchain technology (other than payment systems) will take place in the Metaverse is critical for the future of this virtual universe. Blockchain technology allows database creation without the need for any centralized structure. Will our identities on the Metaverse be kept in a decentralized environment or in the central databases of the technology giants involved in the process, with much more personal data than we produce today, including the data of our moods that we are afraid to admit even to ourselves? In either case, how will nation-states approach this situation? What will be the position of nation-states in these virtual universes? Questions like these remain unclear.

Most of the expectations about the Metaverse consist of technologies that we can experience today, but some of them have not yet been massified. In the world of "Second Life", where users create an avatar for themselves and live an alternative virtual life, it is possible to go to school, start a business, earn income, socialize with people, and invest. It is even possible to earn money through Second Life and transfer it to real life. However, with the current Deep Fake and voice cloning techniques, it is possible to produce a video of someone who does not exist to make someone say things they have never said. This audio-video combination is exceptionally close to reality. Imagine that you encounter this in a 360-degree virtual reality universe, and you can even experience the feeling of touch with the help of a special glove in your hand. Even today, it is getting harder to

understand what is real and what is virtual, and this will be a much more significant challenge. Apart from many start-ups and entrepreneurs working on the Metaverse, the Largest Projects are those:

#### *Decentraland*

Technology giant Samsung has moved its famous Samsung 837 store in New York to Decentraland. One of the largest metaverse universes with 90,601 parcels allows users to build whatever they want. MANA is the cryptocurrency of this universe powered by the Ethereum blockchain. MANA gained 15 percent after Samsung announced that it had entered Decentraland ([www.decentraland.org](http://www.decentraland.org)).

#### *The Sandbox*

The Sandbox (SAND) project, a critical metaverse coin, has recently been a favorite of investors. Sandbox is a gaming metaverse created by Pixowl company in 2011, and the cryptocurrency SAND was built on the Ethereum blockchain. The Sandbox also announced last month that it is collaborating with Adidas. American rapper Snoop Dogg has started building a virtual mansion here ([www.sandboxgame.com](http://www.sandboxgame.com)).

#### *mOVR*

It is a Metaverse platform with Ethereum blockchain-based AR and VR experiences. It is possible to buy plots worldwide through OVR, which combines the real and virtual worlds with AR. With OVR, one can experience the Metaverse, meet new people in OVR rooms, and attend NFT exhibitions in galleries. Two hundred thousand parcels were sold from Turkey, and more than 11 thousand of these parcels are from Istanbul. Many Metaverse projects are still in their infancy. Again, augmented, and virtual reality technologies, or AR (XR), still have a long way to go. In addition to their technological development, they also need to be cheap and mass-produced ([www.ovr.ai](http://www.ovr.ai)).

## **11. CONCLUSION**

Many IT based devices such as televisions, smart mobile phones, ovens, refrigerators work with computer systems. It has become commonplace to benefit from many computers such as the military, architecture, education, and advertising. With the development of technology, computers that used to make only electronic data transfer and complex calculations can solve problems that cannot be solved mathematically, summarize large amounts of data by filtering, comment on events and even learn independently. In other words, computers can both make decisions by collecting information about events and learn the relationships between events by interpreting them.

Although the MR technology, which brings together the virtual and real-world, continues its active development, especially in education and health, it has made itself known with applications in health. Still, it is also used in the military, automotive, marketing, etc. It is a technology suitable for use in many areas. However, it is possible to see different application examples in many fields, from medical education to diagnosis and treatment processes of this technology, which has made a name for itself with its studies in health. Imaging technologies such as MR, which continue to develop and progress rapidly, are based on the ability to operate related processes independently of time, place, and people in many subjects or to execute them remotely. Creating virtual worlds that could only be imagined in the past and overlapping the virtual and real worlds is now possible with augmented, virtual, and MR. In this study, we examined the applications in the field of health, which is an applied field that can make maximum use of the possibilities offered by MR. Especially in medical education, its use for diagnosing and treating diseases will become widespread. HAVELSAN, which has become a world brand in simulation technologies and has original software-oriented solutions and products, acts with the vision of being the national and international leading system integrator in its field. In this context, As HAVELSAN realized, the Hapsim simulator, which is based on MR technology, is one of the new generation technologies since simulation technology is its main field of activity and includes various disciplines. Likewise, it continues to develop different solutions by following the developments in the MR field (Dogan et al., 2021).

Many companies today use MR for production and research. Companies may distribute equipment for MR to their users and enable them to create their digital environments. While screen and media addiction are the subjects of scientific research and congresses, scientists are investigating how children and young people, who are called "netizens" or "digital citizens", will be affected by new technologies. It is thought that the effects of virtual reality on children and young people will be more intense than on adults. It is predicted that new technologies will cause physical discomfort, especially in the head, eye, and neck region of children and young people who are in the age of development physically and cognitively, and even damage their brains. In addition, the infrastructure of controlling the minds of children and young people by hypnotizing them by governments, capital owners, and other power centers seems possible with this new technology. Government regulative bodies, as a public service broadcaster and a major media producer, should take the necessary steps to guide society to the technologies of the future in a healthy way, with the responsibility required by public service broadcasting in the face of these developments in the world and Turkey. It should be a pioneer in the sector by providing R&D support to qualified virtual reality studies in education and entertainment and producing educational content. Together with TRT, RTÜK and other relevant public institutions should prepare the infrastructure for legal regulations regarding this new technology and new content. Establishing an upper advisory board and an ethics committee in this area would also be beneficial. The developing virtual reality world also creates new professions. Universities' software, design, fine arts departments, and educational institutions need to create a specialized training program for these emerging professions.

The Metaverse is an important vision for the new era of the internet. There appears to be a few misperceptions in this area for three reasons:

- The first is that Facebook's Meta launch is supposed to start from today to tomorrow. There is no possibility that it will not materialize at the level of expectations and return to a bubble. However, the glasses technology and processor speeds must reach the desired level; five years at best and ten years.
- Second, it is perceived as a single Metaverse universe and is the subject of plot speculation. However, there is no reason why there should not be thousands of Metaverse universes in the future.
- Third, although it is reduced to it today, it is not just a vision tied to VR glasses. Metaverse is something that we have entered since 2003's Second Life game. Games like Roblox, Fortnite, World of Warcraft always occur in the Metaverse.

While having fun in this virtual world, we will also be with one foot. It is a beautiful version of fantasy and science fiction, but how are we supposed to use it for the benefit of humanity? The purpose of life is not just to live, eat, drink, reproduce. Then we become the same as primitive creatures. The goal of separation is to think about death, reality after death, life. Only worldly thinking cannot make a person happy. Metaverse cannot serve human happiness.

Metaverse inherently transforms into personal, biometric, financial, and even emotional data, raising security, privacy, and intellectual property concerns. Likewise, when avatars are used as a form of identity, personal data becomes vulnerable to being copied, stolen, deleted, or manipulated. Although biometric identification is a solution, it should not be forgotten that identity fraud is always a significant risk, and necessary measures should be started now. At a personal level, there should be some reasonable precautions against VR and AR risks and threats:

- *Reviewing privacy and data confidentiality policies and procedures:* Sometimes, it is easy to bypass lengthy data privacy policies or terms and conditions.
- *Avoiding revealing very personally private information:* Do not disclose personal information that does not need to be disclosed.
- *Using a VPN tunnelling:* One of the ways to keep identity and data privacy on the internet is to use a VPN service.

- *Being active in the real world:* VR is not the reality itself; it will remain virtual.
- *Sticking humanistic, religious and cultural values to protect humanity.*
- *Keeping firmware up to date:* Keep updated for VR headsets and AR wearables. Updates help fix security flaws and add new features and improve existing ones.
- *Using comprehensive antivirus software:* The best way to stay safe online is to use antivirus software to capture viruses, ransomware, malware, phishing, spyware, and other emerging internet security threats.

When we as our research question "*Will VR, AR and MR innovative technology lead to interaction with spirit realms, jinn and angels when supported by super AI?*", with the limits of literature knowledge it is found no answers yet. VR, AR and MR applications aim to overcome the limitations of the physical environment in theory and can provide instant interaction with other physical environments at the speed of light. Unbelievably rapid developments in the innovative dimension of scientific applications do not yet provide definitive evidence on whether this will enable spiritual interaction or whether it will enable communication with angels, demons and spirits of dead people though it is still on the enhancement of physical interaction and made-up virtual artefacts. It is assumed that if such an opportunity and functionality is obtained, religious perceptions, faith and worship may be adversely affected. But this area should be studied with a multidisciplinary approach by theologians, psychologists, spiritualists, and philosophy branches.

## REFERENCES

- 3dforscience (2020) Differences between VR, AR and MR, <https://3dforscience.com/differences-between-vr-ar-and-mr/>
- Ağralı, Ö. ve Aydın, Ö. (2021). "Tweet Classification and Sentiment Analysis on Metaverse Related Messages", *Journal of Metaverse*, 1(1), s.25-30.
- Aleksandrovich, A. & Gomes, L. (2020). Shared multisensory sexual arousal in virtual reality (VR) environments. *Paladyn, Journal of Behavioral Robotics*, 11(1), 379-389. <https://doi.org/10.1515/pjbr-2020-0018>
- Antoniou P. E., Daflı E., Arfaras G. and Bamidis P. D., (2016) "Versatile Mixed Reality Educational Spaces - A Medical Education Implementation Case," 15th International Conference on Ubiquitous Computing and Communications and 2016 International Symposium on Cyberspace and Security (IUCC-CSS), 2016, pp. 132-137, DOI: 10.1109/IUCC-CSS.2016.026.
- Avzal, A., Özdemir, D., & Erarslan, K. (2022). Aizanoi Antik Kentinin Artırılmış ve Sanal Gerçeklik Teknolojileri Kullanılarak Turistik Amaçlı Tanıtım Uygulamala Örneklerinin Tasarımı. *Eskişehir Türk Dünyası Uygulama ve Araştırma Merkezi Bilişim Dergisi*, 3(2), 66-73.
- Aydın, İ. H. (2021) Metaverse ve Sonrası, *Luminosophy*, [https://www.acapublishing.com/dosyalar/baski/LUMINOSOPHY\\_2021\\_566.pdf](https://www.acapublishing.com/dosyalar/baski/LUMINOSOPHY_2021_566.pdf)
- Barbaros, I. (2020) Sinema ve Sanal Gerçekliğin Ortak Geleceği Üzerine Bir Değerlendirme, Yaşar Üniversitesi Master Tezi
- Birt, J., Moore, E., & Cowling, M. (2017). Improving paramedic distance education through mobile mixed reality simulation. *Australasian Journal of Educational Technology*, 33(6). <https://doi.org/10.14742/ajet.3596>
- Borandag, E. & Yücalar, F. (2020). Arttırılmış gerçeklik ile scrum task board uygulaması. *Uluslararası Yönetim Bilişim Sistemleri ve Bilgisayar Bilimleri Dergisi*, 4 (1), 1-12. DOI: 10.33461/uybisbbd.652366
- Bray B., Mcculloch J., Schonning N., Zeller M. (2018). What is mixed reality?. <https://docs.microsoft.com/en-us/windows/mixed-reality/mixed-reality>

- Burdea, G. ve Coiffet P.(1994). *Virtual Reality Technology*. Paris: Wiley-Interscience, Hermes.
- Carbotte K. (2018). *Lenovo Explorer Windows Mixed Reality Headset Review: A Low-Cost VR OnRamp*. <https://www.tomshardware.com/reviews/lenovo-explorer-windows-mixed-realityvr,5468.html>.
- Cem H. (2021) Karma Gerçeklik, Haber24, <https://t24.com.tr/yazarlar/hayri-cem-haftalik/karma-gerceklik,29959>
- Cem Hayri, (2020) Augmented Reality, T24, June 30, URL: <https://t24.com.tr/yazarlar/hayri-cem-haftalik/artirilmis-gerceklik,27213>
- Damar, M. (2021). "Metaverse Shape of Your Life for Future: A Bibliometric Snapshot", *Journal of Metaverse*, 1(1), s.1-8.
- Demirci Ş. (2018) Sağlık Hizmetlerinde Sanal Gerçeklik Teknolojileri, nönü Üniversitesi Sağlık Hizmetleri Meslek Yüksekokulu Dergisi, ISSN: 2147-7892, Cilt 6, Sayı 1(2018) 35-46
- Doko, E. (2021). Alternatif bir dünya arayışı: Metaverse. *Lacivert*, 85, 79-81.
- Ergün O., Akın Ş., Dino İ. G. and Surer E., (2019) "Architectural Design in Virtual Reality and Mixed Reality Environments: A Comparative Analysis," *IEEE Conference on Virtual Reality and 3D User Interfaces (VR)*, 2019, pp. 914-915, doi: 10.1109/VR.2019.8798180.
- Ferhat S. (2016) Dijital Dünyanın Gerçekliği, Gerçek Dünyanın Sanallığı Bir Dijital Medya Ürünü Olarak Sanal Gerçeklik, TRT Akademi ISSN 2149-9446 Cilt 01 Sayı 02 Temmuz Dijital Medya Sayısı
- Frenkel S., Browning K., (2021) The Metaverse's Dark Side: Here Come Harassment and Assaults, *NYTIMES*, <https://www.nytimes.com/2021/12/30/technology/metaverse-harassment-assaults.html>
- Grau, O. (2003). *Virtual Art: From Illusion to Immersion*. Cambridge: MIT Press.
- Holonext, Admin, (2019) What is Mixed Reality? 02.01. 2019, URL: <https://holonext.com/karma-gercelik-nedir/>
- Huang, W., Alem, L., Tecchia, F. et al. Augmented 3D hands: a gesture-based mixed reality system for distributed collaboration. *J Multimodal User Interfaces* 12, 77–89 (2018). <https://doi.org/10.1007/s12193-017-0250-2>
- Intel (2021) Comparison between VR, AR and MR: What do you need to know?, <https://www.intel.es/content/www/es/es/tech-tips-and-tricks/virtual-reality-vs-augmented-reality.html>
- İpek, A.R. (2020) Naming And Identification Problems At Augmented Reality, Virtual Reality And Mixed Reality, *IDIL*, s. 1061–1072. <https://doi.org/10.7816/idil-09-71-0>
- Kalkan, N. (2021). "Metaverse Evreninde Sporun Bugünü ve Geleceğine Yönelik Bir Derleme", *Ulusal Spor Bilimleri Dergisi*, 5(2), s.163-174
- Kaspersky (2021) What are the Security and Privacy Risks of VR and AR, <https://me-en.kaspersky.com/resource-center/threats/security-and-privacy-risks-of-ar-and-vr>
- Kayabaşı, Y. (2005) Sanal Gerçeklik Ve Eğitim Amaçlı Kullanılması, *The Turkish Online Journal of Educational Technology – TOJET* July 2005 ISSN: 1303-6521 volume 4 Issue 3
- Kiiro (2020) Realistic, Interactive VR Sex Is Finally Here, and It's Affordable, *Futurism*, <https://futurism.com/vr-sex-kiiroo-titan-headset-vibrating-stroker>
- Kurbanoglu S. (1996) Sanal Gerçeklik: Gerçek Mi, Değil Mi?, *Türk Kütüphaneciliği* 10,1, 21-31, <http://www.openaccess.hacettepe.edu.tr:8080/xmlui/bitstream/handle/11655/10496/1461-2921-1-PB.pdf?sequence=1&isAllowed=y>

- Kuş, O. (2021). Metaverse: ‘Dijital Büyük Patlamada’ Fırsatlar ve Endişelere Yönelik Algılar .  
Intermedia International E-journal, 8 (15) , 245-266 . DOI: 10.21645/intermedia.2021.109
- Milgram, P. & Kishino, F. (1994). A Taxonomy of Mixed Reality Visual Displays. IEICE  
Transactions on Information Systems, Vol E77-,D, No.12, 1-15.
- Nalbant, K. G. & Uyanık, Ş. (2021). "Computer Vision in The Metaverse", Journal of Metaverse,  
1(1), s.9-12.
- Oppenheim, Charles. (1993). "Virtual reality and the virtual library", Information Services and Use  
(13):215-227.
- Özdemir, D., & Öztürk, F. (2022). The Investigation of Mobile Virtual Reality Application  
Instructional Content in Geography Education: Academic Achievement, Presence, and Student  
Interaction. International Journal of Human–Computer Interaction, 38 (16), 1487-1503, DOI:  
10.1080/10447318.2022.2045070
- Özkahveci, E.; Civek, F. & Ulusoy, G. (2022). “Endüstri 5.0 Döneminde Metaverse (Kurgusal  
Evren)’ÜN Yeri”, Journal of Social, Humanities and Administrative Sciences, 8(50):398-409.
- Rathnayake W. G. R. M. P. S., (2018) "Usage of Mixed Reality for Military Simulations,"  
International Conference on Current Trends towards Converging Technologies (ICCTCT),  
2018, pp. 1-5, doi: 10.1109/ICCTCT.2018.8550993.
- Reznek M, Harter P, Krummel T. (2002) Virtual Reality and Simulation: Training the Future  
Emergency Physician. Academic Emergency Medicine, 9(1): 78-87.
- Scoble, R. & Israel, S. (2017). The Fourth Transformation: How Augmented Reality and Artificial  
Intelligence Change Everything. USA: Patrick Brewster Press.
- Shen G., Dai J., Moustafa H. and Zhai L., (2021) "5G and Edge Computing Enabling Experience  
Delivery Network (XDN) for Immersive Media," IEEE 22nd International Conference on High  
Performance Switching and Routing (HPSR), 2021, pp. 1-7, doi:  
10.1109/HPSR52026.2021.9481809.
- Siglin T. (2018) AR, MR, XR and VR Streaming: Understanding the 4 R’s, WOWZA,  
<https://www.wowza.com/blog/ar-mr-xr-and-vr-streaming-understanding-the-4-rs>
- Stone, Robert J. (1991). "Virtual reality and cyberspace: From science fiction to science fact",  
Information Services and Use (11): 283-300.
- Sung Eunyong (Christine), Sujin Bae, Dai-In Danny Han, Ohbyung Kwon, (2021). Consumer  
engagement via interactive artificial intelligence and mixed reality, International Journal of  
Information Management, Volume 60, 102382, ISSN 0268-40 12,  
<https://doi.org/10.1016/j.ijinfomgt.2021.102382>
- Tang YM, Au KM, Yohana Leung, (2018) Comprehending products with mixed reality: Geometric  
relationships and creativity, International Journal of Engineering Business Management,  
<https://doi.org/10.1177/1847979018809599>
- Varina R. (2021) I Cheated on My Husband by Fucking a Virtual Reality Porn Star, and I Want to  
Do It Again—IRL, Cosmopolitan, <https://www.cosmopolitan.com/sex-love/a35843984/virtual-reality-sex-review/>
- Wiederhold BK. Lessons Learned as We Begin the Third Decade of Virtual Reality.  
Cyberpsychology, Behavior, and Social Networking, 2016; 19(10): 577-578.
- Yang Qinglin, Zhao Yetong, Huang Huawei, Zheng Zibin (2022) Fusing Blockchain and AI with  
Metaverse: A Survey, Journal of Computer Science Computers and Society,  
<https://arxiv.org/abs/2201.03201>



Yang, J., Sasikumar, P., Bai, H. et al. (2020). The effects of spatial auditory and visual cues on mixed reality remote collaboration. *J Multimodal User Interfaces* 14, 337–352  
<https://doi.org/10.1007/s12193-020-00331-1>

Yengin, D. ve Bayrak, T. (2017). *Sanal Gerçeklik - VR*. İstanbul: Der publications.

# Yazılım Sektöründe Kişilik Tiplerinin Bir Uzman Sistem Kullanılarak Belirlenmesi: Pilot Çalışma<sup>1</sup>

\*\*\*

## Determination Of Personality Types in Software Industry by Using an Expert System: A Pilot Study

Emine KURUOĞLU<sup>2</sup> 

Çağla OZEN<sup>3</sup> 

Mustafa Ersel KAMASAK<sup>4</sup> 

DOI:10.33461/uybisbbd.1214494

### Öz

#### Makale Bilgileri

##### Makale Türü:

Araştırma Makalesi

##### Geliş Tarihi:

05.12.2022

##### Kabul Tarihi:

26.12.2022

©2022 UYBİSBBD  
Tüm hakları saklıdır.



Bir yazılım geliştirme projesi üzerinde çalışan grup üyelerinin kişilikleri, ekip üyelerinin rollerine uygun olmasına bağlı olarak, çalışma sonucunun başarısını etkiler. Bu nedenle, ekip üyelerinin yazılım projelerindeki rollerini kişilik farklılıklarına göre belirlemek önemlidir. Bu çalışma, yazılım projelerinde yazılım mühendislerinin uygun rollerini belirlemek ve yazılım sektöründeki yetenek yönetimi için, yenilikçi ve yaratıcı bir uzman sistem önermektedir. Önerilen sistemin en belirgin yönü, teste girenler için kişilik analizi hissini uyandırmayan sorulardan oluşmasıdır. Sorular, yazılımcıların projelerindeki görev ve doğal davranışlarıyla alakalı olacak şekilde tasarlanmıştır. Yazılım mühendisleri genellikle psikometrik testlere inanmazlar ve bunları manipüle etmeye çalışırlar. Ek olarak, bu testlerdeki soru sayısı oldukça fazla olduğundan sıkılabilmektedirler. Önerilen test, Jung Tabanlı Yazılım Yetenek Yönetimi (JBSTM) ve Keirsey Mizaç Sıralayıcı testinin sonuçları ile karşılaştırılmıştır. Bu iki test, 20 yazılım sektörü çalışanına rastgele bir sıra ile tek bir oturumda uygulanmıştır. Testlerden sonra da bir karşılaştırma anketi doldurtulmuştur. E / I, T / F ve J / P'de JBSTM ve Keirsey testi arasında orta derecede bir korelasyon ortaya çıkmıştır. JBSTM ve Keirsey test sonuçları test tamamlanma süreleri açısından da karşılaştırılmıştır. Karşılaştırma anketinin sonuçları kullanılarak JBSTM ve Keirsey testi kolaylık, uygunluk, rahatlık ve kullanılabilirlik açısından da karşılaştırılmış ve JBSTM'in dört açıdan da Keirsey testine kıyasla daha yüksek puanlar aldığı görülmüştür.

**Anahtar Kelimeler:** MBTI, Keirsey testi, kişilik testleri, yazılım mühendisleri, yetenek yönetimi.

#### Article Info

##### Paper Type:

Research Paper

##### Received:

05.12.2022

##### Accepted:

26.12.2022

©2022 UYBİSBBD  
All rights reserved.



The personalities of the group members working on a software development project affect the success of the work result, depending on the suitability of the team members for their roles. Therefore, it is important to determine the roles of team members in software projects according to their personality differences. This study proposes an innovative and creative expert system for determining the appropriate roles of software engineers in software projects and talent management in the software industry. The most obvious aspect of the proposed system is that it consists of questions that do not evoke the feeling of personality analysis for those who take the test. The questions are designed to be relevant to the tasks and natural behaviors of the software engineers in their projects. Software engineers often do not believe in psychometric tests and try to manipulate them. Furthermore, the number of questions in these tests is quite large, so they can get bored. The proposed test was compared with the results of the Jung-Based Software Talent Management (JBSTM) and the Keirsey Temperament Sorter test. These two tests were applied to 20 software industry employees in a single session in a random order. A comparison questionnaire was completed after the tests. A moderate correlation between JBSTM and the Keirsey test appeared in E / I, T / F and J / P. The JBSTM and Keirsey test results were also compared in terms of test completion times. Using the results of the comparison questionnaire, the JBSTM and Keirsey test were also compared in terms of easiness, suitability, comfort and usefulness, and it was found that the JBSTM received higher scores compared to the Keirsey test in all four respects.

**Keywords:** MBTI, Keirsey test, personality tests, software engineers, talent management.

**Atıf / to Cite (APA):** Kuruoglu, E., Ozen, Ç., Kamasak M., (2022). Determination Of Personality Types in Software Industry by Using an Expert System: A Pilot Study. International Journal of Management Information Systems and Computer Science, 6(2), 149-161

<sup>1</sup> This work is related with an ongoing doctorate thesis that is titled as “Developing an Expert System for Talent Management in the Software Industry”

<sup>2</sup> Istanbul Technical University, Turkey kuruoglu17@itu.edu.tr.

<sup>3</sup> Asst. Prof. Dr. Information Systems and Technologies, Yeditepe University, Turkey cagla.ozen@yeditepe.edu.tr

<sup>4</sup> Prof. Dr. Bilgisayar Mühendisliği, Bilgisayar ve Bilişim, Istanbul Technical University, Turkey kamasak@itu.edu.tr

## 1. INTRODUCTION

Various personality tests are adopted throughout the software development industry in order to hire the right person for the right position or enabling that person to pursue the correct career path after recruitment. Yet, software engineers typically have high technical and analytic intelligence, and they do not believe in these tests. Hence, they might manipulate the results in a way that they do not reflect the truth. They may even choose the necessary answers that will enable them to have valid results that will help them get the desired results. As there is no invalidity or contradictoriness in the results report, an unexperienced person might not see the manipulation and may interpret the results incorrectly. Besides, as the number of questions in almost all the psychometric tests are too many, software engineers find it boring to answer all of them.

Considering all these observations, we believe a novel test is required specifically for software engineers. This test should have questions that are related to the tasks in software projects.

With this approach, participating software engineers would respond naturally as they do not feel like they are taking a personality test and the most accurate, realistic and beneficial results can be obtained.

## 2. BACKGROUND

For the determination of personality types within software engineering field, there have been many studies conducted. Methods that are used widely can be listed as; Myers-Briggs Type Indicator (MBTI), Keirsey Temperament SorTer test (Keirsey test) (based on MBTI), Five Factor Model (FFM), NEO Personality Inventory – Revised (NEO PI-R) (based on FFM) etc. In the following subsection, these tests are compared. Next, Jung's 16 personality types which both MBTI and Keirsey tests are based are explained.

### 2.1 Comparison of MBTI, Keirsey Test, FFM and NEO PI-R

(White, 1984) and (Cruz et al., 2015) argues that MBTI is widely used within software engineering studies. In software projects, heterogeneous team members are more productive and efficient in unstructured tasks, while homogeneous team members working together in structured jobs provide the optimum (White, 1984). (Cruz et al., 2015) systematically reviewed 90 articles that empirically explored personalities in software engineering. These articles were selected from among 19,000 articles published between 1970 and 2010. 72% of these studies were published after 2002. Based on these 90 articles, we concluded that the most used test to investigate personality traits in software engineering is the MBTI.

According to (Kaluzniacky, 2004), in studies where MBTI is employed, it was shown that emotional intelligence is important. Besides, (Cruz, 2015) pointed out that there are specific situations that FFM is employed. (McCrae, 1992) and (Feldt et al., 2010) studied traits such as Openness (O), Conscientiousness (C), Expressiveness (E), Appropriateness (A) and Neuroticism (N) in FFM applications. Yet, specifically within the software engineering industry, research of traits such as Neuroticism (irritability, anxiety, bad temper or disappointment) have disturbed the participants. Due to this, many software companies refrain from using FFM-focused personality inventories. Besides, in Turkey, some of the human resources experts from several software engineering companies have reported that the participants have been unwilling to answer NEO PI-R since it has several questions regarding private issues.

(Kosti et al., 2014) provided the compared results between MBTI-based personality inventories and FFM. (Rutherford, 2001) applied Keirsey test on computer engineering students to provide good project teams that have been selected based on the results. According to these results, ESTJ type groups generate a lot of ideas and they are following a traditional path. Contrary to this, ISTJ type are very quiet and behave on their own. When groups consisted of a wide variety of types, they were successful in having the ability to listen and discuss and thus find what was "best" for the group.

Therefore, in software engineering, heterogeneous groups were good at solving problems and generating more innovative and productive ideas. On the other hand, in homogeneous groups where everyone is of the same type, strengths and weaknesses can be distorted because there is no counterbalance to meet them.

(Capretz and Ahmed , 2010a) and (Capretz and Ahmed , 2010b) have published the connections between the software engineering job descriptions and both MBTI and Keirsey test results. In this publication, the authors mentioned the difficulty of researching personality types associated with software engineering since these studies usually include theoretical research on psychological type connections between human-oriented disciplines and software engineering. In this work, the authors contributed towards linking software engineering concepts and the psychology of software developers. The authors investigated the mapping of human roles and MBTI personality types at different stages of the software engineering lifecycle. For this, they first determined the roles in software engineering under five groups as system analyst, software designer, programmer, tester and maintainer. The authors then defined the job descriptions for these five software engineering roles across 12 soft skills. Then a map that matches them with MBTI personality types was proposed. In this manner, a map that presents soft skills needed for the job descriptions of these five roles in software projects was formed. Therefore, a link between soft skills and the MBTI's personality types (E/I, S/N, T/F, and J/N) was created with this map. Next, which role in software engineering would match which personality type of MBTI was deduced. As a result of this effort, they concluded that the best systems analysts would be E-F, the best designers would be N-T, the best programmers would be I-S-T, the best testers would be S-J, and the best maintainers would be S-P.

(Capretz et al., 2015) applied MBTI to 100 Cuban software engineers. According to these results, ESTJ type was the most common MBTI personality type (25%). It was followed by ESTP (15%) and ISTJ (10%) types. The least common personality types were INFJ and INFP types with 1%. ISFP, ENTP and ESFJ types constituted 2% of the total study participants. In this study, participants were also asked which software development roles they preferred. As a result, system analyst, system designer and programmer were the most preferred roles, while tester and maintenance roles emerged as less preferred roles

MBTI and Keirsey are widely used in the software engineering industry for recruitment, careers and talent management. Because MBTI and Keirsey results give very accurate results in matching employees with the right roles and positions. It is also widely and reliably used to identify promising “high potential” employees for the organization and to create talent pools (Cruz et al., 2015), (White, 1984).

## 2.2 Jung’s Personality Types

Myers-Briggs personality types (Myers, 2010) and the personality types that are used in Keirsey test (Keirsey and Bates, 1984) were created upon (Jung, 1921) personality types. According to (Yilmaz et al., 2017) MBTI and Keirsey test can be used both in education and in business life as a determiner of personality traits. JBTSM, the test which is explained in terms of its design, application and results within the framework of this article, was also founded upon Jung’s personality types.

Jung’s two main classifications for personality types are Introversion (**I**) and Extraversion (**E**) types (Fordham, 2011). By adding Thinking (**T**), Feeling (**F**), Sensing (**S**), INtuition (**N**), Jung defines eight types of Introversion and Extraversion types.

- 1) Thinking (**T**) - Introversion (**I**): This way of thinking focuses on subjective ideas.
- 2) Thinking (**T**) - Extraversion (**E**): This way of thinking focuses on the physical and outside world.

- 3) Feeling (F) - Introversion (I): Feelings are kept inside and feelings are not usually exposed.
- 4) Feeling (F)- Extraversion (E): Feelings are affected by the outside world and are also exposed.
- 5) Sensing (S) - Introversion (I): Reactions are based on the internal subjectivity.
- 6) Sensing (S) - Extraversion (E): Progresses on perceptible experiences and does not get stuck on anything for long.
- 7) INtuition (N) - Introversion (I): Knowledge is based on intuitionism.
- 8) INtuition (N) - Extraversion (E): Connections are established with the outer world through intuitions.

Finally, with the addition of Judging (J) and Perceiving (P) to these eight types, a total of 16 different personalities have been generated (see Table 1).

**Extraversion (E) - Introversion (I):** Introversion types (I) gain their energy from within themselves and they are mostly self-motivating. Extraversion (E) types, on the other hand, gather their energy externally and thrive in a group environment. They are the ones to initiate a conversation in a public atmosphere.

**Sensing (S) – Intuition (N):** INtuition (N) types prefer abstract concepts and innovative thinking. They prefer solving new problems and refrain from tasks that do not provide added value. Sensing (S) types, on the other hand, prefer knowledge through experience and have an eye for details. They solve problems through experiencing.

**Thinking (T) - Feeling (F):** Thinking (T) types mostly behave objectively and rationally when making a decision and try to be impartial. Feeling (F) types, on the other hands, prefer social and subjective evaluations and decide in accordance with their feelings. They also create strong personal connections.

**Judging (J) - Perceiving (P):** Judging (J) types prefer being planned, organized and systematic. Perceiving (P) types, on the other hand, do not like planning and are flexible. They are not bothered by disorganized environments and might leave their jobs to the very last minute.

Table 1. Personality types of MBTI and Keirsey test

ISTJ	ISFJ	INFJ	INTJ
ISTP	ISFP	INFP	INTP
ESTP	ESFP	ENFP	ENTP
ESTJ	ESFJ	ENFJ	ENTJ

(Baron,1988) mentioned that these 16 different personalities and characters are specialized in different fields of competence, have interests in different jobs and they are successful in relevant jobs to their personality type. Furthermore, (Keirsey and Bates, 1984) provide an emphasis on important details regarding the process of choosing a job in the business life.

(Bush and Schkade,1985), mention that there is an abundance of use of MBTI or Keirsey test within the industry of software engineering. In addition to this, they argued that ISTJ type of personality is the most prevalent one to this industry. (Turley and Bieman, 1995) also observed that I and T factors are the most common ones among software engineers. (Gorla and Lam, 2004) concluded that people with S and J factors are more inclined to programming work and project team duties. As stated by (Ahmed et al. , 2012), ESTJ types are more effective in software projects due to their extraversion characteristic, so that they can be defined as important and distinctive figure in these projects. Specifically in projects that are custom-made for a client, with their highly-developed social skills, ESTJ types are commonly sought after. (Martinez et al., 2010) stated that according to

MBTI, NT types tend to be more creative than ST types. NT types see the possibility beyond the reality, seek patterns and make connections between them. They have more capability on determining principles through data. NT types can provide a theoretical framework and, by going beyond the details, they make new principles visible. Research and development of software systems by using latest technologies is more interesting for ST types. On the other hand, (Martinez et al., 2010) also points out that many NF and SF types are happy to work with clients directly as engineers focusing on less technical parts and parts that need more human interaction. Furthermore, (Martinez et al., 2010) states that projects are more successful when different behavioral competencies as well as technical competencies of system analysts, designers, programmers, testers and maintainers involved in software lifecycle stages are taken into account. (Capretz and Ahmed, 2010) also discuss that the studies focusing on which personality type is more suitable for what kind of software development effort have a great importance. As there is a need for client interaction, people with high E factors are able to empathize with the client. Despite their high technical knowledge, if there is a lack of E factor, many software engineers might fail at being a system analyst. E and F types are able to connect and empathize with clients much better than I and T types. F types make the client feel better while T types might be unable to feel the needs of a client. Due to this, ENFJ, ENFP and ESFJ types are generally chosen for customer interface sides of a project.

People working on the design side have a perspective that is holistic and wide, which enables them to see all the components of a system. They are creative and innovative. They think abstractly, and by having a conceptual look and modeling, they bring abstract definitions to uncertainties. Hence, people with N factor are great designers. They conduct various tasks such as prototyping, developing processing functions and definition of inputs and outputs. As the first phase of the design phase requires team discussions and client interaction, designers might require similar abilities with the analysts. That is why (Capretz and Ahmed, 2010a) claim that people with these duties require a T factor in order to think and find the best logical solution. Programmers convert the designed version into a program. By processing a programming language in detail, they turn a designed version into a working one. That is the reason programmers have the ability to pay attention to details, logical and analytical competence. Due to this, (Capretz and Ahmed, 2010a) determined that programmers with a T factor are better than the others. Besides, they also concluded that F type people might have difficulties meeting the expectations in programming due to the fact that they are emotional, not logical. S type people pay attention to details and I types might stay long hours in front of a computer without feeling the need to socialize. Hence (Capretz and Ahmed, 2010a) concluded that ISTJ and ISTP types are better programmers.

Testers mainly find software problems. While unit tests are generally conducted by the programmers who run the code for the first time, functionality and feature tests, interface tests, integration and systems tests and validation tests which confirm that the software is working as a whole are run by testers. Test strategies and steps are planned and methodical. Hence, while J types are better at testing with their planned and orderly approach, S types are better in applications. (Capretz and Ahmed, 2010a) also concluded that in order to provide ease of access for the client in interface tests and for "user-friendly interfaces", E and F types are more suitable, while I types are more suitable for working with databases or complicated tests. Generally, for user interface and user functionality, ESFJ types are much more suitable, while ISFJ types are better in detailed system tests. In addition to this, as system analysts are mainly ESFJ types, it is only natural that they create the test scenarios. While ISTJ types are better with more complex tests that require more analytical thinking, they find the role of a "programmer" more suitable for their jobs and personalities. Just like the role of maintenance, the role of documentation also requires patience, attention and putting similar and iterative topics on paper time and again. F types are generally preferred here in order to understand the user needs. I types with their strong written communication skills, are more successful in documentation. As N factor improves the imagination, design and visual presentation skills, INFJ and INFP types are generally more successful in documenting jobs. Generally, software engineers do not think that this is an appropriate job for their skillset, for a person who graduated as a software engineer

who has an interest in editorship and writing, this role might be perfectly suitable. As they can manage the main components of the software project management lifecycle, scope, time and cost, with their extrovert qualities, ESTJ types show strong operational management qualities with their knowledge of teamwork and their ability to keep the team together with their strong communication skills, hence they are the best suitable type for the role of managing a software project. While projects including research, projects that include the latest developments and innovations mainly fall into the interest areas of N type, people who take the role of protecting and developing software systems are mainly S types with their practical, realistic and observant qualities. Generally, S types prefer having a job that has been proven to be successful before. Yet, on the contrary, an N type person generally prefers to conduct a business in a way that has never been tried and tested before since they enjoy working in a field that has never been tried and developed and requires a high level of innovative thinking and creativity. Hence, N types are generally bored with the optimizations and small corrections that are required in the software maintenance side, due to their preference of new projects. S types, however, prefer jobs that require the use of established knowledge instead of developing new ideas. Also, they are great observers and have a tendency to focus on the details. P types are mostly calm, they enjoy searching for the possibilities and take their time deciding, yet J types are mostly determined and impatient. Hence, for maintenance projects that require patience, sometimes monotony and recurring “patching”, P types are more suitable. Because they are more open to changes in adaptations and they are more sympathetic towards the constant flow of changes asked by the client. Problem-solving ability and the applied approach of SP types are two good qualities of the maintenance projects, as these people enjoy solving practical and defined problems. So, ESTP types provide great maintenance software engineers. Furthermore, (Omar et al.,2015) mentions a similar study conducted in Malaysia and in that study, INTJ, ENTJ and ISTJ types were the most prevalent ones while no ENFP types were encountered.

### **3. METHODOLOGY**

#### **3.1 Participants**

20 software developers (system analysts, programmers, testers, and team leaders) who have an advanced level of English and work in the software department of a technology firm were participated in the study. There were 16 male and four female participants with a mean age of 30.5 years. The participants’ work experience in software industry was 8.5 years on average. Before conducting the study, they were informed that their responses will not affect their career lives, will not be shared with the officials in the company, and will only be used as statistical data in the article without specifying the company and the subject name. The participants are also informed about other details and they are free to leave anytime. Their consent is taken in a written form.

#### **3.2 Materials**

##### **3.2.1 Proposed Test**

Proposed test consists of 20 questions with two answer choices (see Appendix). These questions are generated by considering project management steps of software engineering processes, agile software development methodology, people’s activities, tasks, roles, behaviors and attitudes within each step of a software project. Permission was taken from Istanbul Technical University Ethics Committee (ITU\_INAREK) for the whole study and proposed questionnaire.

##### **3.2.2 Keirseey Test**

Keirseey test was discussed in detail in the previous section (see Background section).

### 3.2.3 Comparison Questionnaire

Comparison questionnaire was developed with four questions to compare proposed questionnaire and the Keirsey test in terms of easiness, suitability, comfort (in terms of privacy) and usefulness with ten-point scales (1 is “Not at all likely”, 10 is “Extremely likely”) (see Appendix).

### 3.3 Procedure

Firstly, the participants were informed about the study. Next, proposed and Keirset tests were conducted in random order finally comparison questionnaire was applied. Test completion time of each participant has been measured for the first two questionnaires.

### 3.4 Data Preparation and Analysis - Jung Based Software Talent Management (JBSTM) Expert System

Jung Based Software Talent Management (JBSTM) is an expert system for effective talent management in software industry. The system identifies each participant’s personality type by using proposed questionnaire based on Jung’s 16 personality types. The 20 questions were developed by creating five questions for each of the following four group pairs; E/I, S/N, T/F and J/P. The ultimate goal of this system is to determine people’s most suitable roles in a software project by taking their psychological type into consideration.

As shown in the following figure (see Figure 1), JBSTM has three inputs; proposed questionnaire, Keirsey test and the comparison questionnaire. The expert system has four processes; conducting tests and questionnaire that are inputs of the system, checking validity and reliability, determining people’s psychological types, and matching determined psychological type with a role in a software project. The system produces three outputs; JBSTM expert system report and results of the Keirsey test and the comparison questionnaire.

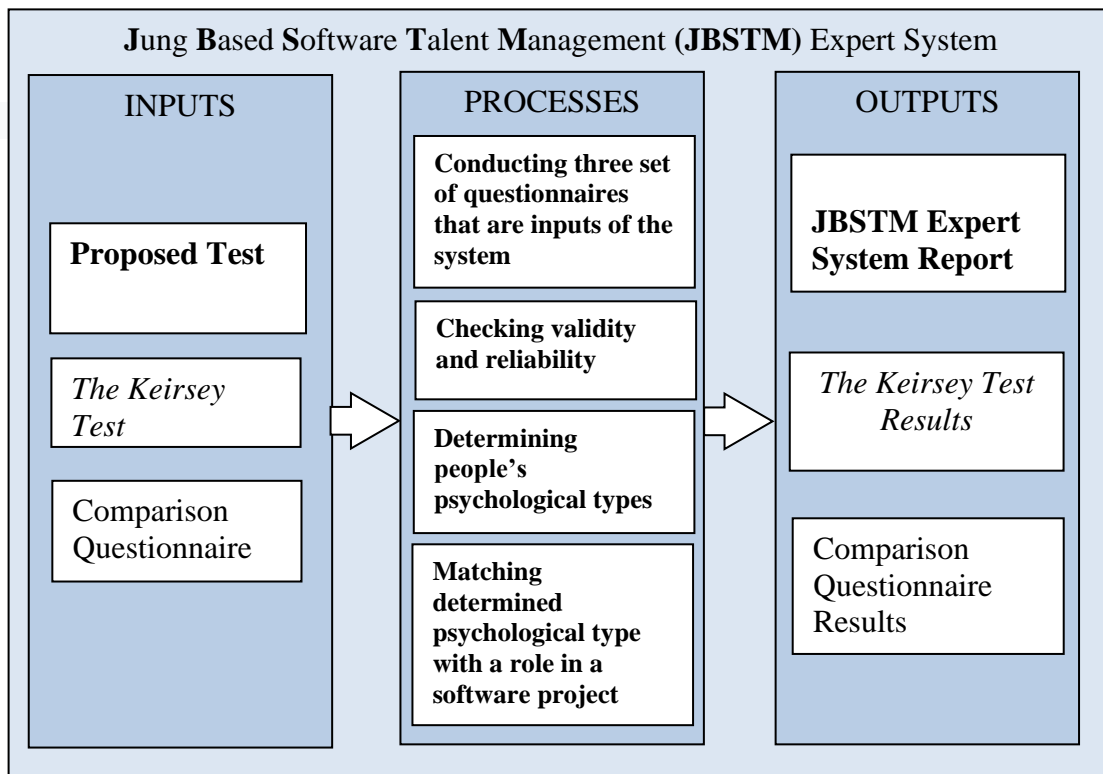


Figure 1. Jung Based Software Talent Management (JBSTM) Expert System for Talent Management in Software Industry



To investigate the validity and reliability of the system, the system uses compliance validity (Alpar, 2018:502) that is comparing the results of the referenced test with the results of the proposed test. Therefore, the system checks the outputs of the Keirsey test and the JBSTM by taking (Alpar, 2018) Pearson Correlation Coefficient (R) values into consideration (see Table 2).

Table 2. R values and definitions (Alpar, 2018)

R values	Definition
0.00-0.19	No correlation
0.20-0.39	Low degree
0.40-0.69	Moderate degree
0.70-0.89	High degree
0.90-1.00	Perfect

#### 4. RESULTS AND DISCUSSION

The Keirsey test and JBSTM in E/I, T/F and J/P revealed a moderate degree correlation based on (Alpar, 2018) R value classification (R = 0.42 for E/I dimension, R = 0.51 for T/F dimension, R = 0.47 for J/P dimension). Only for the S/N dimension's R is low (R = 0.1).

JBSTM and Keirsey test results were compared in terms of test completion times. On average, 20 participants responded JBSTM in seven minutes while they replied Keirsey test in 16 minutes. This result indicates that the time participants spent answering the Keirsey test was almost twice the time they spent answering the JBSTM.

JBSTM and Keirsey test results were also compared in terms of easiness, suitability, comfort and usefulness (see Table 3). JBSTM acquired higher scores compared to Keirsey test in all four aspects.

Table 3. Comparison Questionnaire Results

	JBSTM Score	Keirsey Score
Easiness	157	132
Suitability	171	154
Comfort	183	167
Usefulness	165	129

The results showed that the participants found JBSTM easier and more convenient. Furthermore, participants stated that they feel more comfortable while they were answering JBSTM. Besides, they think that JBSTM is much more useful for the software industry than the Keirsey test.

This work is based on the observations and experiences which were derived from the software engineering industry. Questions in the test are based on the project processes, business manners and behaviors in software engineering and these questions were directed to the participants in the pilot scheme.

Based on Jung, which is widely used for the validity of the test, when looking at its correlation with Keirsey test, validity correlation was obtained in E/I, T/F and J/P types. Intentionally, or unintentionally, the crisis period at the time of the test and the participants' concerns about their careers caused N weighted markings in the test. It was also concluded from the feedbacks given by the participants that they had anxiety that would affect their career life and that they marked the way they wanted to be seen, not as they were. The post-application feedbacks of the participants also support this view. Therefore, it would be better to improve the questions in a way that the subjects will not be able to feel if they are S or N.

It is also found that the time participants spent answering the Keirse test was almost twice the time they spent answering the JBTSM. This result can be considered as a validation for our claim. Since psychometric tests contain plenty of questions and they require very long time to provide responses, software engineers do not prefer these tests.

The study results showed that the participants found JBTSM easier and more convenient. Furthermore, participants stated that they feel more comfortable while they were answering JBTSM. Besides, they think that JBTSM is much more useful for the software industry than the Keirse test. These are promising and significant results that reinforce the purpose of the development of JBTSM.

On the other hand, in practice, some difficulties have occurred within the companies which the pilot scheme is already performed or intended to be performed. The companies are not enthusiastic to take part in this practice on the ground of the law on the protection of personal data, as they have encountered such cases that the subjects brought a suit against the companies following their participation. In one of the companies that the survey was conducted, since another survey which was going to be applied to the employees was rejected by the company because of the law, our practice has also drawn negative reaction from some employees. The study was requested to be done remotely from a company by using electronic devices. However, such a method is not compatible with our pilot scheme rules. In another company, when a voluntary invite for the participation was sent to the employees, they were not eager to attend the practice, since they had some worries about accessibility of the results, specifically by the HR department, and the management. It can be inferred that the resistance against the voluntary study derives from the employees' career concerns and trust issues with their companies.

## 5. CONCLUSION

Most of the software engineers feel disturbed by the psychologically-based personality tests. Also as previously mentioned, these disadvantages are not present in an innovative and creative approach, yet, on the contrary, by considering whether it is possible to maintain a positive approach and trust by the software engineers, a method has been developed that is naturally adaptive to the environment of the software engineering projects and questions were prepared in this manner.

Most of the software engineers who are working in the software industry have very strong analytical thinking skills. Therefore, it is observed that they can manipulate the personality tests to provide the most common and desired results. Because, they believe that the psychometric tests are not aimed at what they do in their own business world. Moreover, they claim that the tests are providing false results and these tests may affect their careers in a negative way. Also, as there are plenty of questions in psychometric tests, they find them rather boring.

In conclusion, JBTSM, which has been developed due to the manipulations of introverted and analytical software engineers who can get bored easily because of the question numbers in the psychometric tests, has provided considerably good results in terms of validity and reliability. Also, as expected, it has produced notably positive outcomes in terms of being easy, beneficial, and comfortable. Within this respect, it can be inferred that JBTSM is worth studying with a high motivation.

## REFERENCES

- Ahmed, F., Capretz, L. F., & Campbell, P. (2012). Evaluating the demand for soft skills in software development. *It Professional*, 14(1), 44-49.
- Alpar, C. (2018). Spor Sağlık Ve Eğitim Bilimlerinden Örneklerle Uygulamalı İstatistik Ve Geçerlik Güvenirlik. 5.baskı, Detay Yayıncılık.
- Baron, R. (1998). *What type am I?: Discover who you really are*. Penguin

- Bush, C. M., & Schkade, L. L. (1985). In search of the perfect programmer. *Datamation*, 31(6), 128-132.
- Capretz, L. F. (2012, February). Software development and personality traits. In *Tenth International Conference on Computer Applications* (pp. 1-8).
- Capretz, L. F., & Ahmed, F. (2010). Making sense of software development and personality types. *IT professional*, 12(1), 6-13.
- Capretz, L. F., Ahmed, F. (2010a). Making sense of software development and personality types. *IT professional*, 12(1), 6-13.
- Capretz, L. F., Ahmed, F. (2010b). Why do we need personality diversity in software engineering?. *ACM SIGSOFT Software Engineering Notes*, 35(2), 1-11.
- Capretz, L. F., Varona, D., & Raza, A. (2015). Influence of personality types in software tasks choices. *Computers in Human behavior*, 52, 373-378.
- Cruz, S., da Silva, F. Q., & Capretz, L. F. (2015). Forty years of research on personality in software engineering: A mapping study. *Computers in Human Behavior*, 46, 94-113.
- Feldt, R., Angelis, L., Torkar, R., & Samuelsson, M. (2010). Links between the personalities, views and attitudes of software engineers. *Information and Software Technology*, 52(6), 611-624.
- Fordham, F., & Yalçiner, A. (2001). Jung psikolojisinin ana hatları. Say.
- Gorla, N., & Lam, Y. W. (2004). Who should work with whom? Building effective software project teams. *Communications of the ACM*, 47(6), 79-82.
- Jung, C. G. (1921). 1971. Psychological types. *Collected Works of CG Jung*, 6.
- Kaluzniacky, E. (2004). Managing psychological factors in information systems work: An orientation to emotional intelligence. IGI Global.
- Keirsey, D., & Bates, M. M. (1984). Please understand me: Character & temperament types (p. 210). Del Mar, CA: Prometheus Nemesis Book Company.
- Kosti, M. V., Feldt, R., & Angelis, L. (2014). Personality, emotional intelligence and work preferences in software engineering: An empirical study. *Information and Software Technology*, 56(8), 973-990.
- Martínez, L. G., Licea, G., Rodríguez-Díaz, A., & Castro, J. R. (2010, June). Experiences in software engineering courses using psychometrics with RAMSET. In *Proceedings of the fifteenth annual conference on Innovation and technology in computer science education* (pp. 244-248).
- McCrae, R. R., & John, O. P. (1992). An introduction to the five-factor model and its applications. *Journal of personality*, 60(2), 175-215.
- Myers, I. B., & Myers, P. B. (2010). Gifts differing: Understanding personality type. Nicholas Brealey.
- Omar, M., Katuk, N., Syed Abdullah, S. L., Hashim, N. L., & Romli, R. (2015). Assessing personality types preferences amongst software developers: a case of Malaysia. *ARPN Journal of Engineering and Applied Sciences*, 10(3), 1499-1504.
- Rutherford, R. H. (2001). Using personality inventories to help form teams for software engineering class projects. *ACM Sigcse Bulletin*, 33(3), 73-76.
- Turley, R. T., & Bieman, J. M. (1995). Competencies of exceptional and nonexceptional software engineers. *Journal of Systems and Software*, 28(1), 19-38.
- White, K. B. (1984). A preliminary investigation of information systems team structures. *Information & Management*, 7(6), 331-335.

Yilmaz, M., O'Connor, R. V., Colomo-Palacios, R., & Clarke, P. (2017). An examination of personality traits and how they impact on software development teams. *Information and Software Technology*, 86, 101-122.

## APPENDIX

### Proposed Test

1. Some software projects are time-independent research projects. These are not always expected to be finalized. Some projects are development projects where it is important to reach the target and finalize it on time. Which one would you like to work in?  
 Research projects                       Development projects
2. When designing a product's database, it requires speculative thinking and considering many possibilities in the future, such as expanding the feature set and changing technology. Coding is the transformation of designed requirements into concrete functionality through coding. Based on these definitions, in which role would you like to work?  
 Database design                       Coding
3. The Product Owner, who manages the needs of the customers by meeting with the customers frequently, and the Product Manager on the technical side of the project have equal authority and rights with the projects using Agile methodology. What role would you like to work in?  
 Product owner                       Product manager
4. Which of the following describes your physical working environment in your workplace?  
 I don't mind the disorder of my desk. I may appear disorderly to others, but in my mind, it has its own order.                       If my desk is messy and untidy, I can't focus. After getting it in order, I get to work.
5. Some software development projects are customer-specific and require frequent contact with the customer. Some are machine-based (M2M, robotics, etc.), customer-independent software development projects. Which project would you prefer to work on?  
 Machine-based and customer-independent                       Customer-based
6. There are two open positions in the software project. In the job description of the first position, the design in which the components are constructed and associated is required, and in the job description of the second position, the coding of the designed structure is required. Which position do you prefer?  
 Design                       Coding
7. You need to make a decision to end or continue a project in progress in the unit you manage. How do you take this decision?  
 I decide by considering how the people involved in the project will be affected by the situation.                       I make decisions based on data, thinking realistically and logically.
8. Which of the following would suit you best when communicating with software project teams?  
 I prefer e-mail or convey written information.                       I choose to convey information by speaking face to face.
9. Which of the following would you like to deal with during the design phase of a software development project?

- Database design that can meet the features and functions that can be added to the product in the future, and that requires fictional and relational thinking.
- Functional design in the process of converting current needs of customers or users into software.
10. In addition to the tasks, they perform themselves in software development projects, mentoring teams are formed for the future. Which team would you like to be on?  
 Technical and professional orientation       Project teams coordination
11. In which of the following tasks would you be more productive in a software project?  
 Working on a single project with clear inputs and goals       Working on a project where the goal can change according to new inputs
12. Assuming you have the same powers, fees, and rights, would you prefer to advance your career in technical positions (API design, setting coding standards, etc.) or administrative positions (sales, marketing, management, etc.)?  
 I prefer to advance in technical positions       I prefer to advance in administrative positions
13. It is important to analyze user behavior especially in mobile and web applications. Thus, it is ensured that effective and easy designs are made that users will like and prefer. What is your level of interest in these subjects?  
 I am very interested, I would like to understand how users feel.       I am not very much interested, I prefer to focus more on technical issues.
14. In customer-specific software development projects, firstly, by connecting with people in different positions on the customer side, their needs are analyzed. After the analysis, the focus is on coding independently of the customer. In which of these processes would you prefer to be involved?  
 Requirements analysis       Coding
15. The following two different software pools are created in a software company. Which one would you like to take part in?  
 Pool of software developers focusing on design and modeling       Pool of software developers focused on optimization and quality
16. There are two different types of projects in a software company:  
A) Long-running, extensive but routine software projects, subject to specifications, where requirements and schedule do not change.  
B) Dynamic projects where client's requirements and schedule can change.  
Which project would you like to take part in?  
 Projects subject to specification, with wide-ranging requirements and schedule unchanged       Dynamic projects where client requirements and schedule can change
17. Which of the following best describes you?  
 I have many different Software Development Project ideas and suggestions.       I am more concerned with the implementation of software projects rather than generating ideas.
18. API (Application Programming Interface) developers often work alone or with a small team. Analyzing customer requirements requires working with different people on the customer side. Which one is suitable for you?  
 Customer requirements analysis       API development

19. Which of the following situations describes you in software development projects?  
 I am happy when I deal with multiple projects at the same time or when I switch to new projects.       I am happy when I work on a single project and complete my work.
20. Which of the following two tasks in the software project is suitable for you?  
 By meeting with the customer, feeling the urgency of the features s/he needs and ensuring that the scope of the sprint is determined accordingly.       To take part in the development of the scope and software decided at the sprint meeting.

## COMPARISON QUESTIONNAIRE

Your comments about the inventory:

(A) with 20 questions

(B) with 70 questions

Can you rate the following expressions by giving values from 1 to 10?

**Ease** of making decisions when ticking choices (10: Very easy, 1: Not easy)

A (.....)       B (.....)

**Suitability** of questions (10: Very suitable, 1: Not suitable)

A (.....)       B (.....)

The extent to which the questions made you feel **comfortable** in terms of privacy (10: I felt very comfortable, 1: I did not feel comfortable)

A (.....)       B (.....)

The degree of **usefulness** of the questions to the software industry career life: (10: Very useful: 1 Not useful)

A (.....)       B (.....)

# Covid-19 Pandemisi Döneminde Öğrencilerin Uzaktan Eğitimde Yaşadıkları Sorunların Kelime Bulutu ve Gizli Anlamsal Dizinleme Yöntemleriyle Tespit Edilmesi \*

\*\*\*

## Determination of the Problems Experienced by Students in Distance Education During the Covid-19 Pandemic Period with Word Cloud and Latent Semantic Indexing Methods

Halit IRMAK<sup>1</sup> 

Zümrüt ECEVİT SATI<sup>2</sup> 

DOI:10.33461/uybisbbd.1214562

### Makale Bilgileri

**Makale Türü:**  
Araştırma Makalesi

**Geliş Tarihi:**  
05.12.2022

**Kabul Tarihi:**  
28.12.2022

©2022 UYBİSBBD  
Tüm hakları saklıdır.



### Öz

COVID-19'un tüm dünyada hızla yayılması, birçok ülkenin eğitim kurumlarını geçici olarak kapatma ve uzaktan eğitime geçme kararı almasına neden olmuştur. Zaman ve mekan kısıtlaması olmaksızın uygulanabilen uzaktan eğitim modeli bu süreçte eğitimin sürdürülebilirliği açısından faydalı olmuştur. Ancak, herhangi bir ön hazırlık yapılmadan uzaktan eğitime ani geçiş yapılması beraberinde birtakım sorunları da ortaya çıkarmıştır. Bu bağlamda, uzaktan eğitim süreçlerine katkı sağlaması açısından bu süreçte yaşanan sorunların tespit edilmesi önem arz etmektedir. Bu çalışmanın amacı, COVID-19 pandemisi döneminde öğrencilerin uzaktan eğitimde yaşadıkları sorunları güncel metin madenciliği yöntemleri ile tespit etmektir. Bu amaçla öğrencilerin uzaktan eğitim sürecindeki destek kayıtlarına ait metinler toplanarak bir veri havuzu oluşturulmuş ve bu veriler metin madenciliği yöntemlerinden olan Kelime Bulutu ve Gizli Anlamsal Dizinleme (GAD) teknikleri ile analiz edilmiştir. Analiz sonucunda, uzaktan eğitim sistem altyapısından kaynaklandığı anlaşılan sorunlar (derslerin sistemde görülmemesi, açılmaması, ses sorunları), internet, sınav ve ödev ile ilgili sorunlar, en çok öne çıkan sorunlar olarak tespit edilmiştir.

**Anahtar Kelimeler:** Metin Madenciliği, LSI, Kelime Bulutu, Uzaktan Eğitim, COVID-19

### Abstract

The rapid spread of COVID-19 all over the world has caused many countries to decide to temporarily close their educational institutions and switch to distance education. The distance education model, which can be applied without time and space restrictions, has been beneficial for the sustainability of education in this process. However, the sudden transition to distance education without any preliminary preparation has also revealed a number of problems. In this context, it is important to identify the problems experienced in this process in order to contribute to distance education processes. The aim of this study is to identify the problems experienced by students in distance education during the COVID-19 pandemic with current text mining methods. For this purpose, a data pool was created by collecting the texts of students' support records in the distance education process and these data were analyzed with Word Cloud and Latent Semantic Indexing (LSI) techniques, which are text mining methods. As a result of the analysis, the problems that were understood to be caused by the distance education system infrastructure (courses not being seen or opened in the system, audio problems), internet, exam and homework problems were identified as the most prominent problems.

**Keywords:** Text Mining, LSI, Word Cloud, Distance Education, COVID-19.

### Article Info

**Paper Type:**  
Research Paper

**Received:**  
05.12.2022

**Accepted:**  
28.12.2022

©2022 UYBİSBBD  
All rights reserved.



**Atıf/ to Cite (APA):** Irmak, H., Ecevit Satı, Z., (2022). Covid-19 Pandemisi Döneminde Öğrencilerin Uzaktan Eğitimde Yaşadıkları Sorunların Kelime Bulutu ve Gizli Anlamsal Dizinleme Yöntemleriyle Tespit Edilmesi. Uluslararası Yönetim Bilişim Sistemleri ve Bilgisayar Bilimleri Dergisi, 6(2), 162-171

\* Bu çalışma, Doç. Dr. Zümrüt Ecevit Satı danışmanlığında Halit Irmak'a ait Doktora Tez çalışması kapsamında hazırlanmıştır.

<sup>1</sup> Doktora Öğrencisi Enformatik Bölümü, Fen Bilimleri Enstitüsü, İstanbul Üniversitesi, hirmak@ogr.iu.edu.tr

<sup>2</sup> Doç. Dr. İşletme Bölümü, Siyasal Bilgiler Fakültesi, İstanbul Üniversitesi, zsati@istanbul.edu.tr

## 1. GİRİŞ

İnsanın içinde bulunduğu çevreye uyum sağlaması ve farklı çevrelerdeki gereksinimleri gidermesi yaşamı boyunca edinmiş olduğu öğrenmelere bağlıdır. Bu açıdan bakıldığında insanların yaşamları boyunca öğrenmesi bireysel ve toplumsal bir gereklilik olarak görülmektedir. Bu gereklilik nedeniyle eğitimlerde önemli bir görev üstlenerek gerekli öğrenim ortamını sağlamaya çalışmaktadırlar (Kaya, 2002). Eğitimi toplumun her kesimine ulaştırmak ve böylece hem bireylerin hem de toplumun eğitim seviyesini yükseltmek amacıyla uzaktan eğitim sisteminin gerekli olduğu görülmektedir (Arıcı ve Yekta, 2005). Uzaktan eğitim zaman ve mekandan bağımsız olarak isteyen herkese yaşam boyu öğrenme fırsatı sunan bir eğitim sistemi modelidir. Günümüzde bilgi ve iletişim teknolojilerinin yoğun olarak kullanıldığı uzaktan eğitimde öğrenenler daha çok kendi kendine öğrenme sağlamaktadır (İşman, 2011; Kaya, Erden, Çakır ve Bağırşakçı, 2004).

Uzaktan eğitim, bilgi ve iletişim teknolojilerindeki gelişmelerin yarattığı fırsatlar sayesinde birçok önemli gelişme göstermiştir. Bu gelişmelere örnek olarak, zamandan ve mekândan bağımsız eğitim fırsatının sağlanması, kendi hızında ilerleme, dijital ortamlarda depolanan tüm eğitsel içeriklere hızlı ve kolay erişimin olması gösterilebilir. Uzaktan eğitim, herkesin öğrenme imkanlarını arttırmasının yanı sıra, daha verimli bir eğitim ortamı oluşturmak için öğrencilerin öğrenme deneyimlerini geliştirmelerine de yardımcı olmaktadır (Deveci, 2019). Ayrıca Öğrencilere ulaşılabilir ve kaliteli bir eğitim ortamı sağlamak için uzaktan eğitim maliyetleri düşürürken kaliteli bir eğitim ortamı sağlayarak öğrencilerin ve öğretmenlerin her an erişebilecekleri çevrimiçi ortamları kullanarak öğrenmeyi kolaylaştırmaktadır. Bu sayede öğrenciler istedikleri kadar ve istedikleri zaman öğrenme etkinliklerine katılabilmekte ve kendi öğrenme hızlarında ilerleyebilmektedirler (Özbay, 2015).

Dünyada yaşanan COVID-19 pandemisi sonrası birçok ülkede yüz yüze yapılan eğitimler durdurulmuş ve uzaktan eğitime geçiş yapılarak eğitime devam edilmiştir. Çin'in Hubei eyaleti Wuhan şehrinde ortaya çıktığı ifade edilen COVID-19 salgını, Çinli uzmanlar Dünya Sağlık Örgütü Çin Ofisine 2019 aralık ayında bildirmiştir (Dikmen ve Bahçeci, 2020; T.C. Sağlık Bakanlığı, 2020). Yaşanan bu gelişme sonrası Türkiye'de 2020'nin başlarında yüz yüze eğitimin durdurulması kararı alınmış ve uzaktan eğitim sistemine tamamiyle geçiş yapılmıştır. Uzaktan eğitimin sağladığı avantajlar sayesinde hemen hemen her düzeyde eğitim herhangi bir aksama olmadan sürdürülmüştür. Ancak uzaktan eğitimin avantajlı yönlerine rağmen bazı dezavantajları da bulunmaktadır. Bunların başında uygulama derslerinin işleme zorluğu, ilk yatırım maliyeti, sistem hataları, öğrencilerde görülebilen motivasyon düşüklüğü ve teknoloji kullanımı konusunda yeterli bilgi ve deneyime sahip olmayan kişilerin teknoloji tabanlı uzaktan eğitim sistemlerinden yeterince verim alamaması gibi farklı sorunlar görülebilmektedir (Altunsoy, 2019; Ünlükahraman, 2011).

Uzaktan eğitimde karşılaşılan sorunların tespit edilmesinde veri madenciliği yöntemleri kullanılabilir. Veri madenciliği ham veriden anlam çıkarma, daha önceden bilinmeyen geçerli ve uygulanabilir bilgilerin elde edilmesi amacıyla istatistiksel teknikleri ve makine öğrenmesi algoritmalarını kullanan bir yöntemdir (Çelik, 2009). Veri madenciliği büyük veri yığınlarından anlamlı veriler elde etme, öngörülebilir bulunma ve daha önce hiç akla gelmemiş düşünülmemiş sonuçları ortaya koymasıyla diğer yöntemlerden farklılık göstermektedir (Silahtaroglu, 2016). Uzaktan eğitim sistemlerinde yaşanan sorunlara yönelik öğrenci ve akademisyenlerin sistem üzerinde oluşturdukları destek kayıtları birer metin verisi olarak sistemde saklanmaktadır. Metinsel veriler, veri madenciliğinin bir alt alanı olan metin madenciliği yöntemi ile analiz edilebilmektedir. Metin madenciliği, büyük metin verilerini kullanarak anlam çıkarmaya çalışan bir veri madenciliği yöntemidir (Hossain, Karimuzzaman, Hossain ve Rahman, 2021).

Günümüzde yaygın uygulama alanı bulan metin madenciliği yöntemlerinden biri olan konu modelleme yöntemleri sayesinde bir metindeki gizli veya açık konular tespit edilebilmekte ve bu sayede araştırmacılar kısa sürede konu ile ilgili fikir sahibi olabilmektedir (Kaya ve Gülbandılar, 2022). Uzaktan eğitimde öğrencilerin yaşamış oldukları sorunların tespiti ve bu konuda bilgi sahibi olunması, uzaktan eğitim süreçlerinin iyileştirilmesi bakımından önemlidir. Bu çalışmada, COVID-



19 pandemisi döneminde öğrencilerin uzaktan eğitimde yaşadıkları sorunların metin madenciliği yöntemleri ile tespit edilmesi amaçlanmıştır. Bu amaç doğrultusunda metin madenciliği yöntemlerinden olan Kelime Bulutu ve Gizli Anlamsal Dizinleme (GAD) teknikleri çalışmada kullanılarak metinsel veriler analiz edilmiştir. Güncel metin madenciliği yöntemlerinin kullanılarak uzaktan eğitimde öğrencilerin yaşamış oldukları sorunların analiz edilmesi çalışmanın özgün yönünü oluşturmaktadır.

## 2. YÖNTEM

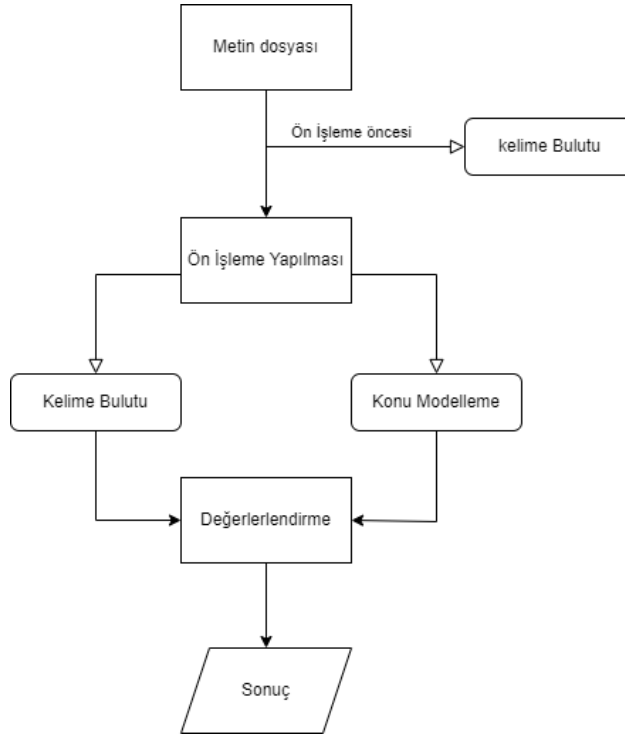
Metin madenciliği, e-ticaret, turizm, bankacılık ve siyaset gibi verinin olduğu birçok farklı alanda kullanılmaktadır (Yılmaz ve Aslan, 2020). Eğitim alanında da metin madenciliği çalışmaları popüler bir konu olmakla birlikte hem ülkemizde (Onan, 2021) hem de yurtdışında bu alanda çalışmalar yapıldığı görülmektedir (Meneses III, Cantilang, Dala ve Madeja, 2021). Uzaktan eğitim öğrenci destek verilerinin kullanıldığı bu çalışmada veri madenciliği yöntemlerinden biri olan metin madenciliği yöntemi tercih edilmiştir. Veriler Türkiye’de bulunan bir devlet üniversitesinden temin edilmiş ve çalışmada 2020-2021 bahar dönemine ait öğrenci destek kayıtları kullanılmıştır.

## 3. VERİ ÖN İŞLEME

Elde edilen metinler üzerinde analizlerin yapılabilmesi için metin ön işleme yapılmıştır. Tüm metinlerin küçük harflere dönüştürülmesi, şapkalı özel karakterlerin düzeltilmesi, url, tarih ve sayıların kaldırılması, tüm noktalama işaretleri ve tek başına bir anlam ifade etmeyen bağlaç, edat ve zamirlerin kaldırılması veri ön işleme basamağında yapılması gereken önemli bir basamaktır (Bostancı ve Albayrak, 2021). Ön işleme ham veri üzerinde işlenen, bilginin toplanması ve basitleştirilmesinde kritik öneme sahip bir adımdır (Hossain ve diğerleri, 2021).

## 4. VERİ ANALİZİ

Çalışmada, metin madenciliği yöntemlerinden olan Kelime Bulutu ve Konu Modelleme yöntemleri kullanılmıştır. Kelime bulutu, metinde geçen kelimelerin frekans değerlerine göre bir görsel oluşturmaktadır. Bu yöntem sayesinde metnin görsel bir özeti oluşturulmakta ve veri görselleştirilebilmektedir. Veri görselleştirme, herhangi bir verideki soyut kavramların belirli frekanslara göre düzenlenerek somut bir forma dönüştürülmesi ve böylelikle görsel bir forma dönüşen metinlerin kolay anlaşılır ve yorumlanabilir hale getirilmesidir (Küçükkartal, 2020). Veri görselleştirme amacıyla kelime bulutları sıklıkla başvurulan bir yöntemdir (Temizhan ve Mendes, 2021). Metin madenciliğinin alt alanlarından biri olan konu modellemede amaç metinlerde geçen gizli veya açık konuları tespit etmektir (Kaya ve Gülbandırlar, 2022). Konu modelleme tekniği olarak Gizli Anlamsal Dizinleme (Latent Semantic Indexing) tekniği kullanılmıştır. Gizli Anlamsal Dizinleme’de (GAD) anlam bakımından birbirine yakın kelimelerin benzer metin parçalarında geçeceği varsayılmaktadır (Casillano, 2019). Bu teknik sayesinde bir metinde yer alan gizli tema ve konular ortaya çıkartılarak metnin genel fikri ortaya çıkarılabilmektedir (Meneses III, Cantilang, Dala ve Madeja, 2021). Metin madenciliği aşamasında izlenen adımlar Şekil 1’de sunulmuştur.



Şekil 1 Metin Analizi Adımları

## 5.BULGULAR

Elde edilen veriler ön işleme çalışmalarına tabi tutulmuş, sonrasında kurulan model yardımıyla analiz edilmiştir. Ön işleme öncesine ait kelime bulutu Şekil 2’de gösterilmiştir.



Şekil 2 Ön İşleme Öncesi Oluşturulan Kelime Bulutu

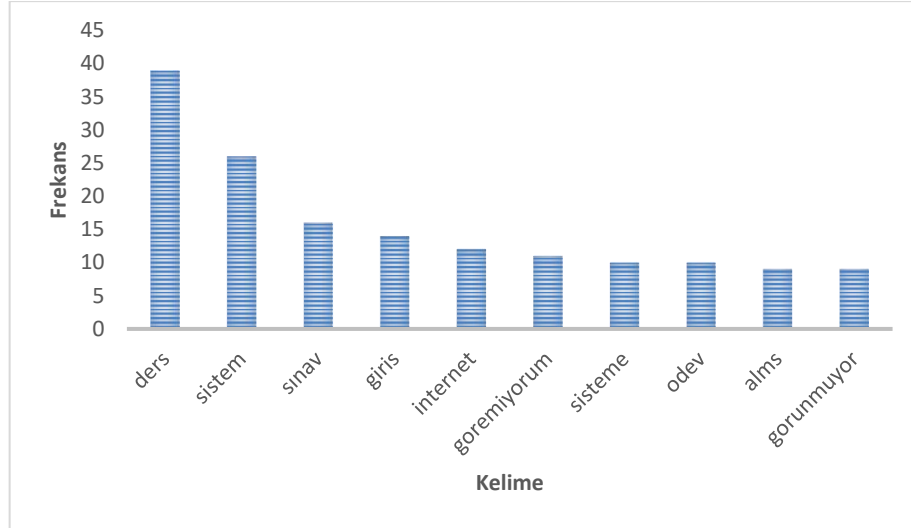
Kelime bulutlarında oluşturulan kelimeler metin içerisindeki kullanım sıklıklarına (frekans) göre bulutun merkezinde ve büyük puntolarla gösterilmektedir (Karadal ve Saygın, 2016). Kelime

bulutları, metindeki kelimelerin frekansına dayalı bir veri görselleştirme yöntemidir (Temizhan ve Mendeş, 2021). Ön işleme öncesi oluşturulan kelime bulutuna bakıldığında bağlaç, noktalama işaretleri, soru ekleri gibi kelimelerin metin içerisindeki frekanslarının fazla olmasından dolayı kelime bulutunun merkezinde yer aldığı görülmektedir. Bu haliyle bir değerlendirmenin yapılması bir anlam ifade etmeyeceği ve ön işlemenin gerekliliği böylece bir kez daha görülmektedir. Kelime bulutu yöntemi sayesinde metnin kolay anlaşılır bir görsel özeti oluşturulurken aynı zamanda metin ile ilgili genel bir düşünce, anlam çıkarılabilmektedir (Hossain ve diğerleri, 2021). Bu yüzden kelime bulutu oluşturulurken ön işlemenin yapılarak kelime bulutlarının oluşturulması genel fikri bulmak ve anlam çıkarmak için çok önemli bir basamaktır. Ön işleme çalışmaları sonrası yeniden bir kelime bulutu oluşturulmuştur (Şekil 3).



Şekil 3 Ön İşleme Sonrası Oluşturulan Kelime Bulutu

Ön işleme sonrası oluşturulan kelime bulutuna ait frekans değerleri (ilk 10 kelime) Şekil 4’te gösterilmiştir. Kelime bulutuna ait şekil 3 ve şekil 4’e bakıldığında öğrencilerin en çok derslerle ilgili sorun yaşadıkları görülürken, devamında lms (Learning Management System) sistemi, sınav uygulaması, internet bağlantısı ve ödev konularında sorun yaşadıkları görülmektedir.



Şekil 4 Kelime Bulutunda Geçen (ilk 10) Kelimelere Ait Frekans Değerleri

Kelime bulutu çalışmaları sonrası metinde geçen gizli veya açık konuların ortaya çıkarılması amacıyla GAD tekniği uygulanmıştır. GAD analizinde beş konu (tema) oluşacak şekilde ayarlama yapılmıştır. GAD analizi sonrası oluşturulan konulara ilişkin anahtar kavramlar Tablo 1’de sunulmuştur.

Tablo 1 GAD Tekniği Kullanılarak Oluşturulan Anahtar Kavramlar

Konu	Anahtar Kavramlar
1	derse, ag, ag_JJ dan, dan, sorun, ders, giris, sorunu, alms, girdigim
2	ders, kaydını, ders_NNS kaydını, olamadım, kaydını_VBP olamadım, derse, olur, sistemde, görünmüyor, sistemde_NN görünmüyor
3	sınav, sınava, sistem, doldu, hakkın, hakkın_NN doldu, yazısı, yenileme, hakkınız, sayfa
4	sorun, yapabilirsiniz, yukleme, giris, ders, sisteme, üzerinden, pdf, sisteme_NN giris, odevi
5	sınava, giris, hakkın, doldu, yazısı, hakkın_NN doldu, sistem, internet, ogrenci, hata

GAD tekniği sonrası beş konu ve on anahtar kavram otomatik olarak oluşturulmuştur. Analiz sonrası ortaya çıkan anahtar kavramlardan yola çıkılarak temalar oluşturulmuş ve bu temalar Tablo 2’de sunulmuştur.

Tablo 2 Konu Anahtar Kavramlarından yola Çıkarak Oluşturulan Temalar

Konu	Anahtar Kelimeler	Tema
1	derse, ag, ag_JJ dan, dan, sorun, ders, giris, sorunu, alms, girdigim	Ağ (internet) sorunundan kaynaklı öğrencilerin uzaktan eğitim sistemine, derslere girememesi sorunu
2	ders, kaydını, ders_NNS kaydını, olamadım, kaydını_VBP olamadım, sistem, derse, olur, sekilde, olur_NN musunuz	Ders kaydı sorunu veya senkronizasyon sorunundan dolayı derslere katılamama
3	sınav, sınava, sistem, doldu, hakkın, hakkın_NN doldu, yazısı, yenileme, hakkınız, sayfa	Çevrimiçi sınav uygulamasında yaşanan sorunlar.
4	sorun, odev yukle, yukleme, alms, ders, sisteme, üzerinden, pdf, sisteme_NN giris, odevi	Uzaktan eğitim sistemine ödev doküman yükleme ile ilgili sorunlar

5	sınava, giris, hakkın, doldu, yazısı, hakkın_NN doldu, sistem, internet, ogrenci, hata	Sınav uygulaması ile ilgili internet kaynaklı sorunlar
---	--	--

## 6. TARTIŞMA VE SONUÇ

Bu çalışmada, öğrencilerin uzaktan eğitim sürecinde yaşadıkları sorunların metin madenciliği yöntemleri ile tespit edilmesi amaçlanmıştır. Bu amaç doğrultusunda metin madenciliği yöntemlerinden olan kelime bulutu ve GAD teknikleri çalışmada kullanılarak metinsel veriler analiz edilmiştir. Veri görselleştirme amacıyla kelime bulutunun oluşturulması bulguların daha kolay anlaşılmasını sağlamıştır. Veriler GAD tekniği ile analiz edilerek anahtar kavramlar oluşturulmuş ve sonrasında anahtar kavramlardan yola çıkılarak temalar oluşturulmuştur.

Pandemi salgını nedeniyle dünyada birçok ülkede uzaktan eğitim sistemine geçilmiş ve yüz yüze eğitime ara verildiği görülmüştür. Ancak yaşanan bu hızlı geçişle birlikte uzaktan eğitimde çeşitli sorunlar ortaya çıkmıştır (Alhazmi, 2022; Caliskan ve diğerleri, 2020). Öğrencilerin bu süreçte yaşadıkları sorunların tespiti üzerine yaptığımız bu çalışmada, metin verilerinden oluşturulan kelime bulutu Şekil 3'te ve kelime bulutunda en çok frekansa sahip ilk on kelime Şekil 4'te sunulmuştur. GAD tekniği sonrası ortaya çıkan anahtar kavramlara ilişkin bulgular Tablo 1'de ve bu anahtar kavramlardan yola çıkılarak oluşturulan temalar Tablo 2'de sunulmuştur.

Öğrencilerin pandemi döneminde uzaktan eğitimde yaşadıkları sorunlara baktığımızda: LMS'de yer alan derslerin açılmaması, derslerin izlenmemesi, ses sorunları, ders kaydı ve ders kaydını yaptıkları derslerin LMS üzerinde görünmemesi sorunlarından dolayı öğrencilerin derslere katılmadığı ve bu konuda çok başvuru yaptıkları anlaşılmaktadır. Alanyazında yapılan benzer çalışmalarda bu sonuçlar desteklenmektedir (Ayyıldız ve Zahal, 2022; Baykal ve Tutuncu, 2022). Ayyıldız ve Zahal'in (2022) pandemi döneminde üniversitelerde uzaktan eğitim ile verilen gitar eğitiminin, gitar eğitimcilerinin görüşlerinden yola çıkarak inceledikleri çalışmada ses sorunları, internet/bağlantı sorunlarını tespit ettikleri görülmüştür. Benzer şekilde Baykal ve Tutuncu'nun (2022) sağlık bölümü öğrencilerinin online eğitime yönelik duygu, düşünce ve deneyimlerinin betimlendiği çalışmalarında, öğrencilerin kendilerinden, sistemden ve az da olsa öğretmenlerden kaynaklı sorunlar yaşadıkları tespit edilmiştir.

İnternet bağlantısından kaynaklanan sorunlar öğrencilerin yaşadıkları bir diğer önemli sorun olarak karşımıza çıkmaktadır. Alanyazın incelendiğinde benzer tespitlerin yapıldığı çalışmalara rastlanmaktadır (Salomao, Ricardo ve Ortiz, 2021; Tarrayo, Paz ve Gepila, 2021). Salomao, Ricardo ve Ortiz'in (2021) COVID-19 sürecinde uzaktan eğitimin analizine yönelik 1082 öğrenci ile yaptıkları çalışmada öğrencilerin 43%'ünün maddi kaynak yetersizliği ve internet erişim sorunu yaşadıklarını tespit etmişlerdir. Tarrayo, Paz ve Gepila (2021) pandemi ortamında esnek öğrenmenin avantaj, dezavantaj, sorunları ve iyileştirilmesi gereken noktalarının neler olduğuna yönelik 38 Filipinli İngilizce bölümü akademisyeniyle anket ve odak grup görüşmesi şeklinde yaptıkları çalışmada katılımcılar belirtmiş oldukları diğer sorunlarla birlikte yeterli internet kaynaklarının sağlanması gerekliliğini vurgulamışlardır.

Öğrencilerin çevrimiçi sınavlar ile ilgili sorun yaşadıkları tespit edilmiştir. İlgili alanyazına bakıldığında (Romaniuk ve Lukasiewicz-Wieleba, 2021; Valizadeh, 2022) çevrimiçi sınavlarla ilgili benzer tespitlerin yapıldığı görülmektedir. Romaniuk ve Lukasiewicz-Wieleba (2021), Maria Grzegorzewska Üniversitesi'nde pandemi döneminde birinci ve ikinci sınavlar sonrası öğrenci ve akademisyenlerle tekrarlı anket çalışması yapmışlardır. Araştırmada öğrencilerin çevrimiçi sınavlarla ilgili katı zaman kısıtlamaları ve dikkat dağıtıcı unsurları olumsuz durumlar olarak ifade ettikleri, her iki grubun da sınav sırasında yaşanan teknik sorunları ciddi sorunlar olarak ifade ettikleri belirlenmiştir. Valizadeh'in (2022) Türk yükseköğretim öğrencilerinin çevrimiçi öğrenme programlarında kopya çekmeye ilişkin, öğrencilerin algılarını ve çözümlerini vurgulamayı

amaçladığı ve 163 öğrenci ile yaptığı çalışmada bazı öğrencilerin yaşanan teknik sorunlardan dolayı bu konuda bir şey yapılamayacağını açıkça belirttikleri tespit edilmiştir.

Öğrencilerin karşılaşmış olduğu bir diğer sorunun ise ödevlerle ilgili sorunlar olduğu sonucuna varılmıştır. Alanyazın incelendiğinde, ödevlerle ilgili benzer tespitlerin yapıldığı çalışmalara rastlanılmıştır (Elçiçek, 2022; Kürtüncü ve Kurt, 2020). Elçiçek'in (2022) COVID-19 pandemisi döneminde, uzaktan eğitim faaliyetlerinde karşılaşılan engellere yönelik yapmış olduğu meta-sentez çalışmasında, artan ödev yükleri nedeniyle öğrencilerin zorlandığı ve ödevlerin öğrencilerde stres oluşturduğu tespitinde bulunmuştur. Benzer şekilde Kürtün ve Kurt (2020) tarafından yapılan çalışmada, internet ve bilgisayar konusunda yaşanan sıkıntılar ve bu konudaki kısıtlı imkanlardan dolayı derslerin takibi ve ödev konularında öğrencilerin sıkıntı yaşadığı tespit edilmiştir. Bu sonuçlar çalışma sonuçlarımızı destekler niteliktedir.

Sonuç olarak, pandemi döneminde öğrencilerin uzaktan eğitimde ağırlıklı olarak sistem altyapısından kaynaklı sorunlar yaşadıkları görülürken, internet erişimi, sınavlar ve ödevlerle ilgili de sorun yaşadıkları görülmüştür. Uzaktan eğitim faaliyetlerinin sorunsuz veya en az sorunla sürdürülebilmesi için sistem altyapısının güçlendirilmesi ve ders dönemi başlamadan önce ses kontrolleri, canlı ders ve arşiv ders testleri başta olmak üzere tüm sistem testlerinin yapılmasının yararlı olacağı düşünülmektedir. Ders dönemi başlamadan önce yapılması gereken bir diğer önemli husus ise öğrenci bilgi sistemi otomasyonunun da bulunan öğrenci, akademisyen ve derslerin LMS sistemine aktarılması ve bu iki sistem arasındaki senkronizasyonun sürekliliğinin sağlanmasıdır. Sınav ve ödev uygulamalarında öğrencilere yeterli sürenin tanınması ve sınav uygulamasında karşılaşılabilecek sorunlara yönelik bir alternatif planın yapılması öğrenci motivasyonunun sağlanması, ölçme ve değerlendirme süreçlerinin doğru bir şekilde işlemesi açısından önemlidir. Bununla birlikte hem okullarda hem de ülke genelinde internet altyapısının iyileştirilmesi ve karar vericilerin bu yönde politikalar gerçekleştirmesi uzaktan eğitimde internet kaynaklı sorunların giderilmesi gerekliliğini ortaya koymaktadır. Çalışmada elde edilen bir diğer sonuç uzaktan eğitim süreçlerinde uzaktan eğitim birimi, üniversite yönetimi, bilgi işlem ve öğrenci işleri birimlerinin koordineli olarak birlikte çalışması ve hareket etmesinin sürecin etkinliğini artmada önemli bir rol oynadığıdır.

## KAYNAKÇA

- Alhazmi, H. (2022). Detection of Students' Problems in Distance Education Using Topic Modeling and Machine Learning. *Future Internet*, 14(6). doi:10.3390/fi14060170
- Altunsoy, F. (2019). Uzaktan Eğitim Öğrencilerinin Başarılarının Yapay Zeka Teknikleri İle Tahmini. Süleyman Demirel Üniversitesi.
- Arıcı, N. ve Yekta, M. (2005). Mesleki ve Teknik Eğitimde Çoklu Ortam Araçları Kullanılmış Web Tabanlı Öğretimin Öğrenci Başarısına Etkisi. *Ticaret ve Turizm Eğitim Fakültesi Dergisi*, 1, 144–153.
- Ayyıldız, E. B. ve Zahal, O. (2022). Instructor experiences with online guitar lessons during the Covid-19 pandemic in Turkey. *International Journal of Music Education*. doi:10.1177/02557614221123078
- Baykal, D. ve Tutuncu, S. K. (2022). Online Education Experiences Of The Students Studying In Health Care Departments During The Covid-19 Pandemic. *Turkish Online Journal Of Distance Education*, 23(1), 131–143.
- Bostanci, B. ve Albayrak, A. (2021). Duygu Analizi İle Kişiyeye Özel İçerik Önerme. *Veri Bilimi Dergisi*, 4(1), 53–60. www.dergipark.gov.tr/veri adresinden erişildi.
- Caliskan, S., Kurbanov, R. A., Platonova, R. I., Ishmuradova, A. M., Vasbieva, D. G. ve Merenkova, I. V. (2020). Lecturers Views of Online Instructors about Distance Education and Adobe

- Connect. International Journal Of Emerging Technologies In Learning, 15(23), 145–157. doi:10.3991/ijet.v15i23.18807
- Casillano, N. F. (2019). Unraveling Views of Students Towards Computer Programming A Sentiment Analysis and Latent Semantic Indexing Application. International Journal of Recent Technology and Engineering (IJRTE), 8(1).
- Çelik, M. (2009). Veri Madenciliğinde Kullanılan Sınıflandırma Yöntemleri ve Bir Uygulama. İstanbul Üniversitesi.
- Deveci, M. (2019). Türkiye’de Yeni İletişim Teknolojilerinin Yükseköğrenim Kurumlarında Uzaktan Eğitim Alanında Kullanımı. Marmara Üniversitesi.
- Dikmen, S. ve Bahçeci, F. (2020). Covid-19 Pandemisi Sürecinde Yükseköğretim Kurumlarının Uzaktan Eğitime Yönelik Stratejileri: Fırat Üniversitesi Örneği. Turkish Journal of Educational Studies (C. 7).
- Elçiçek, M. (2022). Uzaktan Eğitim Engelleri: Bir Meta-Sentez Çalışması. Üniversite Araştırmaları Dergisi, 5(1), 35–45. doi:10.32329/uad.937271
- Hossain, A., Karimuzzaman, M., Hossain, M. M. ve Rahman, A. (2021). Text Mining and Sentiment Analysis of Newspaper Headlines. Information, 12(10), 414. doi:10.3390/info12100414
- İşman, A. (2011). Uzaktan Eğitim. Pegem Akademi (4. Baskı.). Ankara.
- Karadal, H. ve Saygın, M. (2016). İnovasyon Yetenekleri İle Girişimci Kişilik Özellikleri Arasındaki İlişki Üzerine Ampirik Bir Araştırma. Eurasian Business & Economics Journal, S1, 282–294.
- Kaya, A. ve Gülbandılar, E. (2022). Konu Modelleme Yöntemlerinin Karşılaştırılması. Eskişehir Türk Dünyası Uygulama ve Araştırma Merkezi Bilişim Dergisi, 3(2), 46–53. doi:10.53608/estudambilisim.1097978
- Kaya, Z. (2002). Uzaktan Eğitim (1. Baskı.). Ankara: Pegem Akademi Yayıncılık.
- Kaya, Z., Erden, O., Çakır, H. ve Bağırşakçı, N. B. (2004). Uzaktan Eğitimin Temelleri Dersindeki Uzaktan Eğitim İhtiyacı Ünitesinin Web Tabanlı Sunumunun Hazırlanması. The Turkish Online Journal of Educational Technology, 3(3), 165–175.
- Küçük kartal, H. K. (2020). Twitter ’daki Verilere Metin Madenciliği Yöntemlerinin Uygulanması. ESTUDAM Bilişim Dergisi, 1(2), 10–13.
- Kürtüncü, M. ve Kurt, A. (2020). Covid-19 Pandemisi Döneminde Hemşirelik Öğrencilerinin Uzaktan Eğitim Konusunda Yaşadıkları Sorunlar. Avrasya Sosyal ve Ekonomi Araştırmaları Dergisi, Covid-19 Ö(1), 66–77.
- Meneses III, J. M., Cantilang, K. W., Dala, A. D. ve Madeja, J. B. (2021). Decoding Views and Sentiments of Program Heads Towards the Supervision of Instruction During the COVID-19 Pandemic. EPRA International Journal of Multidisciplinary Research (IJMR), (2), 323–329. doi:10.36713/epra8863
- Onan, A. (2021). Sentiment analysis on massive open online course evaluations: A text mining and deep learning approach. Computer Applications in Engineering Education, 29(3), 572–589. doi:10.1002/cae.22253
- Özbay, Ö. (2015). Dünyada Ve Türkiye’de Uzaktan Eğitimin Güncel Durumu. Uluslararası Eğitim Bilimleri Dergisi, 2(5), 376–394. <https://dergipark.org.tr/inesj/issue/40015/475774> adresinden erişildi.
- Romaniuk, M. W. ve Lukasiewicz-Wieleba, J. (2021). Challenges of administering university examinations remotely during the COVID-19 pandemic. E-MENTOR, (3), 22–31.

- Salomao, J., Ricardo, A. ve Ortiz, P. (2021). Distance learning during Covid-19: experience of the Higher Polytechnic Institute of Benguela, Angola. *Revista Edapeci-Educacao A Distancia E Praticas Educativas Comunicacionais E Interculturais*, 21(3), 53–67. doi:10.29276/redapeci.2021.21.315866.53-67
- Silahtaroglu, G. (2016). *Veri Madenciliği Kavram ve Algoritmaları (3. bs.)*. İstanbul: Papatya Yayıncılık Eğitim.
- T.C. Sağlık Bakanlığı. (2020). *COVID-19 (SARS-CoV-2 Enfeksiyonu) Rehberi Bilim Kurulu Çalışması*. Ankara.
- Tarrayo, V. N., Paz, R. M. O. ve Gepila, E. C. (2021). The shift to flexible learning amidst the pandemic: the case of English language teachers in a Philippine state university. *Innovation in Language Learning and Teaching*, 1–14. doi:10.1080/17501229.2021.1944163
- Temizhan, E. ve Mendeş, M. (2021). COVID-19 Pandemisi ile İlgili Twitter Mesajlarının Metin Madenciliği Tekniği İle Değerlendirilmesi. *Türkiye Klinikleri J Biostat*, 13(2), 185–200. doi:10.5336/biostatic.2020-79992
- Ünlükahraman, O. (2011). *Web Tabanlı Eğitimde Web Madenciliği Uygulaması İle Öğrenci Davranışlarının Analizi*. Fırat Üniversitesi.
- Valizadeh, M. (2022). Cheating In Online Learning Programs: Learners' Perceptions And Solutions. *Turkish Online Journal Of Distance Education*, 23(1), 195–209.
- Yılmaz, A. U. ve Aslan, Z. (2020). Uzaktan Eğitim Forum Sistemlerinde Bir Metin Madenciliği Uygulaması. *AURUM Mühendislik Sistemleri ve Mimarlık Dergisi*, 4(1), 9–21.