ISSN: 1307-9697 | e-ISSN: 2147-0715

JOURNAL OF INFORMATION TECHNOLOGIES

BİLİŞİM TEKNOLOJİLERİ DERGİSİ

Volume / Cilt 18 Number / Sayı 2 Year / Yıl 2025 Month / Ay April / Nisan

Gazi Üniversitesi | Bilişim Enstitüsü

GAZİ ÜNİVERSİTESİ *(GAZİ UNIVERSITY)*
BİLİŞİM ENSTİTÜSÜ *(INSTITUTE OF INFORMATICS)*

BİLİŞİM TEKNOLOJİLERİ DERGİSİ

(JOURNAL OF INFORMATION TECHNOLOGIES)
ISSN: 1307-9697 **e-ISSN:** 2147-0715

Cilt *(Volume)*: 18 **Sayı *(Issue)*: 2** ***Nisan(April)* 2025**

Bilişim Teknolojileri Dergisi uluslararası hakemli bir dergidir.
Journal of Information Technologies is an international refereed journal.

Yazışma Adresi *(Contact Address)*
Tunus Cad. No: 35 Kavaklıdere Çankaya/ANKARA
Telefon / Telephone: 0312 202 38 01
Faks / Fax: 0312 212 79 29

Çevrimiçi Değerlendirme Sistemi *(Online Evaluation System)*
http://dergipark.gov.tr/gazibtd
E-posta *(e-mail)***:** btd@gazi.edu.tr

Bilişim Teknolojileri Dergisi 3 ayda bir (Ocak, Nisan, Temmuz, Ekim) yayınlanmaktadır.
Journal of Information Technologies is published every 3 months (January, April, July, October).

Dijital Olgunluk Modelleri Üzerine Bibliyometrik Bir İnceleme

Literatür Makalesi/Review Article

Esra GÜNEŞ [1]*, Yenal ARSLAN [2], Hakan VELİOĞLU [3], Çetin KARAHAN [4]

[1] Denetim ve Risk Yönetimi Programı, Ankara Sosyal Bilimler Üniversitesi, Ankara, Türkiye
[2] Yazılım Mühendisliği Bölümü, Ankara Yıldırım Beyazıt Üniversitesi, Ankara, Türkiye
[3] İç Denetim Başkanlığı, T.C. Tarım ve Orman Bakanlığı, Ankara, Türkiye
[4] Strateji Geliştirme Daire Başkanlığı, T.C. Savunma Sanayii Başkanlığı, Ankara, Türkiye
esra.gunes2@student.asbu.edu.tr, yenalarslan@aybu.edu.tr, hakanveli@gmail.com, cetin.karahan@gmail.com
(Geliş/Received:20.09.2024; Kabul/Accepted:26.12.2024)
DOI: 10.17671/gazibtd.1553658

***Özet*—**Dijital olgunluk modelleri, kuruluşların dijital dönüşüm yeteneklerini değerlendirmesi ve geliştirmesi için temel araçlardır. Bu modeller, teknoloji, organizasyon, kültür, müşteri gibi çeşitli boyutlar açısından bir kuruluşun dijital dönüşüm durumunu değerlendirmek amacıyla yapılandırılmış ve bütünsel bir yaklaşım sunmaktadır. Bu çalışmanın amacı, dijital olgunluk modelleri üzerine yapılan araştırmalara ait detaylı bir analiz yapmaktır. Bu doğrultuda, Scopus ve Web of Science veri tabanlarında yer alan akademik çalışmalar bibliyometrik analiz yöntemiyle incelenmiştir. Scopus ve Web of Science veri tabanlarında yer alan İngilizce dilindeki ve makale türündeki yayınlar, dijital olgunluk modellerinin olgunluk seviyeleri ve boyutları açısından sistematik bir şekilde incelenmiştir. Bulgular, 2020 yılından itibaren dijital olgunluk modeli konusundaki yayın sayısında artış olduğunu, Almanya ve Türkiye'nin bu alanda en fazla yayın üreten ülkeler arasında yer aldığını göstermiştir. İncelenen modellerde ise en çok teknoloji, strateji ve organizasyon boyutlarının ele alındığı tespit edilmiştir. Çalışmanın, bu alanda araştırma yapmak isteyen araştırmacılara kapsamlı bir genel bakış sunarak, gelecekteki çalışmalar için odaklanılabilecek potansiyel alanları belirlemeleri konusunda yardımcı olacağı düşünülmektedir.

***Anahtar Kelimeler*—** dijital olgunluk, dijital olgunluk modeli, dijital dönüşüm, bibliyometrik analiz

A Bibliometric Analysis on Digital Maturity Models

***Abstract*—** Digital maturity models are essential tools for organisations to assess and improve their digital transformation capabilities. These models provide a structured and holistic approach to assessing an organisation's digital transformation status in terms of various dimensions such as technology, organisation, culture, customers, etc. The purpose of this study is to conduct a detailed analysis of the research on digital maturity models. In this direction, academic studies in Scopus and Web of Science databases were analysed by bibliometric analysis method. English-language article-type publications in Scopus and Web of Science databases were systematically analysed in terms of maturity levels and dimensions of digital maturity models. The findings show that there has been an increase in the number of publications on digital maturity models since 2020 and that Germany and Turkey are among the countries that produce the most publications in this field. In the models examined, it was found that technology, strategy and organisation dimensions were mostly addressed. It is thought that the study will help researchers who want to conduct research in this field to identify potential areas to focus on for future studies by providing a comprehensive overview.

***Keywords*—** digital maturity, digital maturity model, digital transformation, bibliometric analysis

1. GİRİŞ (INTRODUCTION)

Dijital teknolojilerin hızla gelişmesi ve organizasyonel süreçlere entegrasyonun zorunlu hale gelmesi, kuruluşların müşterilerle etkileşim kurma ve sürekli değişen pazar dinamiklerine uyum sağlama yeteneklerini şekillendirmektedir. Çoğunlukla dijital dönüşüm olarak adlandırılan bu süreç, sadece teknolojik gelişmelerin benimsenmesini kapsamaz aynı zamanda iş stratejilerinin, kültürünün ve müşteri deneyimlerinin bütünsel olarak yeniden tasarlanmasını da içerir. Dijital dönüşüm kavramı, dijital teknolojilerin ortaya çıkması ve insanların bu yöndeki taleplerinin artmasıyla birlikte önem kazanmıştır. Dijital dönüşüm; kuruluşların dijital teknolojiyi kullanarak hizmetlerini sürdürebilmesi ve varlıklarını muhafaza edebilmeleri için stratejik ve teknolojik bir yenilenme süreci olarak görülmektedir.

Dijital dönüşüm, mobil, yapay zeka, bulut, blok zinciri gibi yeni dijital teknolojilerin, kurumsal müşteri deneyimini artırmak, mevcut operasyonları kolaylaştırmak veya yeni iş modelleri oluşturmak için büyük iş iyileştirmeleri sağlamak amacıyla kullanılan bir süreç olarak tanımlanabilir [1]. Dijital dönüşüm yalnızca yeni bilgi teknolojilerinin sağladığı yeni stratejilerin ve iş modellerinin benimsenmesiyle ilgili değildir, aynı zamanda dijital teknolojileri yaygın bir şekilde entegre eden ve hizmetleri ihtiyaçlara göre yeniden düzenleyen kültürel bir değişimi de içermektedir [2].

Dijital dönüşüm süreci birçok kuruluş için karmaşık ve zor bir süreçtir. Kuruluşlar bu süreci etkili ve verimli bir şekilde yürütmek için kendilerine rehberlik edecek, yaptıkları çalışmaların geçerliliğini ölçebilecek çerçeve veya modellere ihtiyaç duyarlar. Dijital olgunluk modelleri son yıllarda dijital dönüşüm faaliyetlerinin artmasıyla önemli bir konu başlığı haline gelmiştir ve akademik yayınların yanı sıra uluslararası şirketler, yazılım ve danışmanlık firmaları yeni olgunluk modelleri ortaya koymuşlardır [3].

Dijital olgunluk modelleri, dijital dönüşüm yeteneklerini değerlendirmek ve geliştirmek, dijital dönüşümün karmaşık sürecinde yol almak ve dijital teknolojilerin faydalarından yararlanmak isteyen kuruluşlar için temel araçlardır [4]. Dijital olgunluk modeli, bir kuruluşun dijital dönüşüm ve dijital teknolojileri benimseme düzeyini değerlendirmek için kullanılan bir çerçeve olarak da ifade edilebilir [5]. Dijital dönüşüm sürecini yönetmek ve değerlendirmek isteyen bir kuruluşa, dijital olgunluk modelleri, teknolojik, organizasyonel ve stratejik yönler gibi çeşitli boyutları kapsayan sistematik bir yaklaşım sunar [6]. Bir başka tanımla dijital olgunluk modelleri, dijital yetenekleri sistematik bir şekilde değerlendirmek, yönlendirmek ve geliştirmek için yapılandırılmış bir araç sunarak dijital teknolojilerin benimsenmesinin bütünsel olmasını, insan, süreç ve teknoloji arasındaki karşılıklı bağımlılıkların dikkate alınmasını sağlar[7]. Kuruluşlar, mevcut dijital altyapılarını, süreç ve yeterliliklerini dijital olgunluk modelleri sayesinde belirlenmiş kriterlere göre değerlendirebilirler. Dijital olgunluk değerlendirmelerinin, farklı sektörlerdeki kuruluşların dijitalleşme düzeylerini, küçük ölçekli işletmeler de dahil olmak üzere, kapsamlı bir şekilde ölçme potansiyeline sahip olduğu belirtilmektedir[8-9]. Bu değerlendirmeler, kuruluşların operasyonel verimliliğe katkıda bulunacak dijital girişimler için hazır olup olmadıklarına ilişkin değerli iç görüler sunarak, dijital dönüşüm stratejilerinin şekillendirilmesine önemli bir destek sağlayabilir. [8-9-10].

Kuruluşlar, dijital dönüşüm süreçlerindeki ilerleme düzeyini değerlendirmek ve bu süreçleri daha etkili ve düzenli bir şekilde yönetmek istemektedir ve bu durum dijital olgunluk model veya araçlarına olan talebi arttırmaktadır [11-12]. Akademisyenler, araştırmacılar, danışmanlık firmaları, hatta devletler tarafından geliştirilen dijital olgunluk modellerinin sayısı her geçen gün artmaktadır [13-14-15-16-17-18]. Dijital olgunluk modellerinin farklı sektörler için farklı faktörler üzerinden değerlendirilmesi gerektiğini vurgulayan ve boyutların sektörün ihtiyaçlarını yansıtacak şekilde geliştirilmesi gerektiğini belirten çalışmalar mevcuttur (14-19).

Dijital olgunluk kavramının giderek artan önemi, dijital olgunluk modellerinin gelişim süreci ve bu modellerin literatürde ne ölçüde ele alındığı konuları, araştırma ilgimizin odak noktalarından biri haline gelmiştir. Bu bağlamda, söz konusu modellerin boyutları, seviyeleri ve hangi sektörlere yönelik uygulandığı gibi unsurların daha iyi anlaşılması, bu çalışmayı gerçekleştirmemizin temel motivasyonunu oluşturmuştur Bu çalışmada, dijital olgunluk modellerine ilişkin Scopus ve Web of Science (WoS) veri tabanlarında yer alan akademik çalışmalar üzerinde güncel bir bibliyometrik analiz gerçekleştirilmiştir. Araştırma kapsamında ulaşılan İngilizce dilinde ve makale türündeki yayınlarda ele alınan dijital olgunluk modellerine sistematik bir analiz uygulanmış ve dijital olgunluk modelinde yer alan boyut ve seviyeler belirlenmiştir.

Araştırmacıların, araştırma konularını belirlemelerine ve çeşitli alanlar arasındaki ilişkiyi daha iyi anlamalarına yardımcı olmak için birçok farklı alanda yaygın olarak bibliyometrik analiz uygulanmaktadır [20]. Literatürdeki çalışmaların değerlendirmesi ve gelecekteki çalışmalara yön vermesi konusunda bibliyometrik ve sistematik analiz çalışmalarından faydalanılmaktadır [21]. Scopus ve WoS veri tabanlarının atıf analizi ve arama özellikleri açısından farklıları vardır. Kapsamlı ve güvenilir sonuçlar elde etmek için her iki veri tabanındaki akademik çalışmalara yönelik analiz yapılmıştır [22-23].

2. METODOLOJİ (METHODOLOGY)

Bu çalışmada, Scopus ve WoS veri tabanlarından doküman incelemesi yöntemiyle elde edilen ikincil veriler bibliyometrik analiz yöntemi ile incelenmiştir. Araştırmanın ilk aşamasında, mevcut literatürün durumunu ortaya koymak üzere yapılacak inceleme için anahtar kelimeler belirlenmiştir. Belirlenen anahtar kelimeler için

Scopus ve WoS veri tabanlarında arama yapılmış ve elde edilen sonuçlara Tablo-1'de yer verilmiştir.

Tablo 1. Belirlenen anahtar kelimeler ve arama sonuçları[1]
(Identified keywords and search results)

Anahtar Kelimeler	Scopus		Web of Science (WoS)	
	Tüm Diller	*İngilizce*	*Tüm Diller*	*İngilizce*
"Digital Maturity"	733	689	385	357
"Digital Maturity Model"	110	106	42	40
"Digital Transformation"	25,854	24,177	13,233	12,351
"Digital Transformation Maturity Model"	19	17	9	8
"Digital Maturity Assessment"	55	51	27	26
"Digital Maturity Assessment Tool"	5	5	3	3
"Digital Transformation" + "Digital Maturity"	440	410	200	184
"Digital Maturity Model" + "Digital Maturity Assessment"	9	8	8	7

İkinci aşamada, Scopus ve WoS veri tabanlarında sadece "Digital Maturity Model" anahtar kelimesi için yapılan arama sonuçları ayrıntılı olarak incelenmiş ve yayın sayısı, ülke, atıf sayısı, anahtar kelime sıklığı gibi bulgulara yer verilmiştir. Elde edilen veriler Microsoft Office Excel ve VOSviewer paket programıyla derlenmiştir.

Literatür taraması, Scopus veri tabanında "Article Title" veya "Abstract" veya "Keywords"; Web of Science veri tabanında "Title" veya "Abstract" veya "Author Keywords" alanları içinde yapılmıştır. Yapılan tarama sonucunda, İngilizce dilinde ve makale türünde, Scopus veri tabanında 45, WoS veri tabanında ise 25 çalışma tespit edilmiştir. Bu çalışmalar arasında, bazı yazarların eserlerinin her iki veri tabanında listelendiği belirlenmiştir. Her iki veri tabanında ortak olarak yer alan 22 makale, her iki veri tabanında da dikkate alınarak analiz edilmiştir. Toplamda 48 makale detaylı bir şekilde incelenmiştir. Yapılan aramalarda, herhangi bir zaman kısıtlaması uygulanmamıştır.

3. BULGULAR (FINDINGS)

Scopus ve Web of Science veri tabanlarından, "Digital maturity model" anahtar kelimesi için elde edilen sonuçlar detaylı olarak incelenmiştir. Scopus veri tabanında arama, "Article Title", "Abstract" ve "Keywords" alanları içinde; sorgu olarak "TITLE-ABS-KEY ("digital maturity model")" kullanılarak yapılmıştır. Arama sonucunda toplam 110 yayına ulaşılmıştır. Yayın türlerine göre yayınların sayıları ve yüzdeleri Tablo 2'de verilmiştir.

Tablo 2. Yayın türüne göre arama sonuçları, Kaynak: Scopus
(Search results by publication type, Source: Scopus)

Yayın Türü	**Sayısı**	**Yüzdesi**
Makale	49	44.5 %
Konferans Bildirisi	46	41.8 %
Kitap Bölümü	7	6.4 %
Konferans İncelemesi	5	4.5 %
İnceleme	2	1.8 %
Kitap	1	0.9 %
Toplam	110	100.0 %

WoS veri tabanında, "Digital maturity model" anahtar kelimesi için sonuçları detaylı olarak incelenmek amacıyla "Title", "Abstract" ve "Author Keywords" arama alanları ve arama sorgusu olarak "((TI=("digital maturity model")) OR AB=("digital maturity model")) OR AK=("digital maturity model")" kullanılmıştır. Arama sonucunda toplam 45 yayına ulaşılmıştır. Yayın türlerine göre yayınların sayıları ve yüzdeleri Tablo 3'te verilmiştir. Scopus ve WoS veri tabanlarındaki yayınların, yayınlandıkları dillere göre sayıları Tablo 4'te verilmiştir.

Tablo 3. Yayın türüne göre arama sonuçları, Kaynak: WoS
(Search results by publication type, Source: WoS)

Yayın Türü	**Sayısı**	**Yüzdesi**
Makale	27	60.0 %
Bildiri Kitabı	13	28.9 %
Erken Erişim	2	4.4 %
Makale İncelemesi	2	4.4 %
Editoryal Materyal	1	2.2 %
Toplam Yayın	45	100.0 %

Tablo 4. Dillere göre arama sonuçları, Kaynak: Scopus, WoS
(Search results by language, Source: Scopus, WoS)

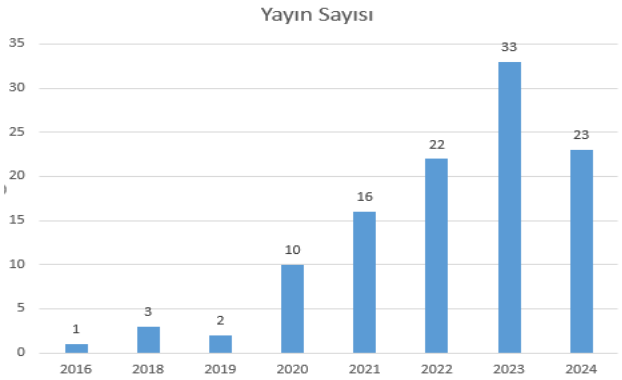
Yayın Dili	**Scopus Yayın Sayısı**	**WoS Yayın Sayısı**
İngilizce	106	40
Rusça	2	1
İspanyolca	1	1
Fransızca	1	-
Tüm Dillerde	110	42

3.1. Yıllara göre yapılan yayın sayılarına ilişkin bulgular
(Findings according to the number of publications by years)

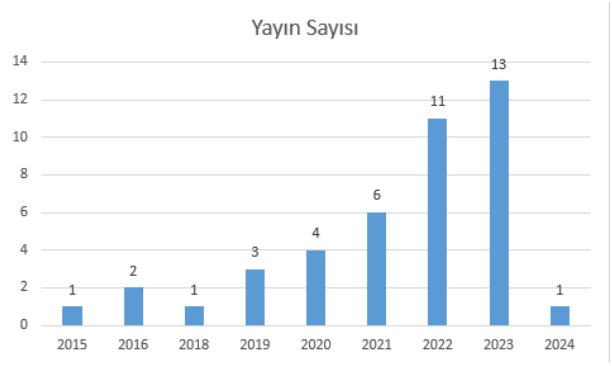
Herhangi bir zaman kısıtlaması uygulamadan, yıllara göre yapılan yayın sayıları incelendiğinde (Şekil 1, Şekil 2), Scopus veri tabanında ilk yayının 2016 yılında listelendiği ve 2019 yılından sonra bir artış eğiliminin olduğu görülmektedir. WoS veri tabanında ilk yayının 2015

[1] Scopus ve WoS veri tabanlarında arama 21.06.2024 tarihinde yapılmıştır.

yılında listelendiği ve her iki veri tabanında 2017 yılında hiç yayın listelenmediği görülmüştür.



Şekil 1. Yıllara göre yayın sayısı, Kaynak: Scopus
(Number of publications by year, Source: Scopus)

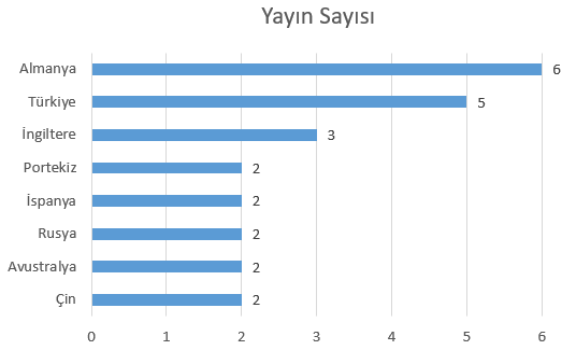


Şekil 2. Yıllara göre yayın sayısı, Kaynak: WoS
(Number of publications by year, Source: WoS)

3.2. Ülkelere göre yapılan yayın sayılarına ilişkin bulgular
(Findings according to the number of publicaitons by countries)

Yayınların ülkelere göre dağılımı incelendiğinde (Şekil 3, Şekil 4), Almanya'nın Scopus veri tabanında 20, WoS veri tabanında 6 yayınla en çok yayını olan ülke olduğu; Türkiye'nin Scopus veri tabanında 8, WoS veri tabanında 5 yayınla en çok yayın üreten ikinci ülke olduğu görülmektedir.

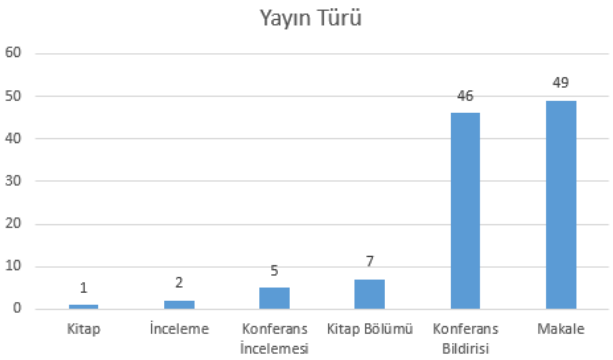
Şekil 3. Ülkelere göre yayın sayısı, Kaynak: Scopus
(Number of publications according to countries, Source:Scopus)



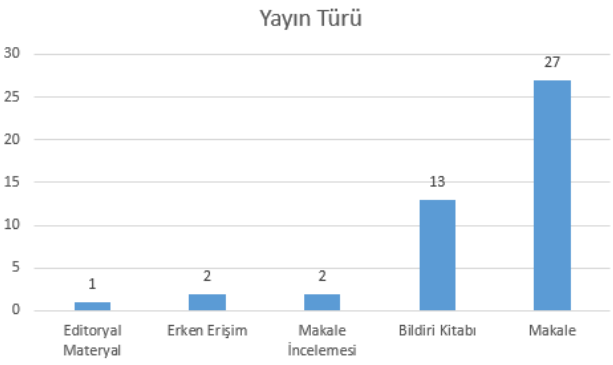
Şekil 4: Ülkelere göre yayın sayısı, Kaynak: WoS
(Number of publications according to countries, Source: WoS)

3.3. Yayın türüne ilişkin bulgular (Findings according to the publication type)

"Digital maturity model" anahtar kelimesi için yapılan arama sonucunda, zaman kısıtlaması olmadan Scopus veri tabanında toplam 110; WoS veri tabanında toplam 42 yayın bulunmuştur. Yayın türü olarak her iki veri tabanında en çok makale türü yer almaktadır (Şekil 5, Şekil 6). WoS veri tabanında 2 yayın hem makale hem de erken erişim; 1 yayın hem makale hem de bildiri kitabı olarak sınıflandırılmıştır.



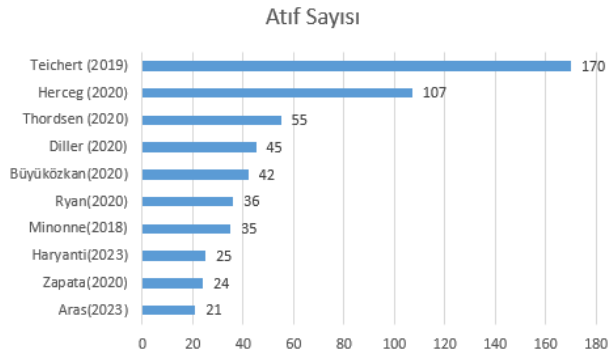
Şekil 5. Yayın türüne göre yayın sayısı, Kaynak: Scopus
(Number of publications according to publication type, Source: Scopus)



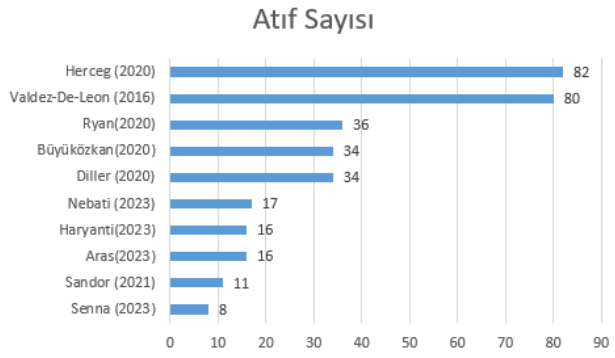
Şekil 6. Yayın türüne göre yayın sayısı, Kaynak: WoS
(Number of publications according to publication type, Source: WoS)

3.4. Atıf sayılarına göre yayınlara ilişkin bulgular (Findings according to the number of citations)

"Digital maturity model" anahtar kelimesi için yapılan arama sonucunda, en çok atıf alan ilk 10 yayın, Şekil 7 ve Şekil 8'de gösterilmiştir.



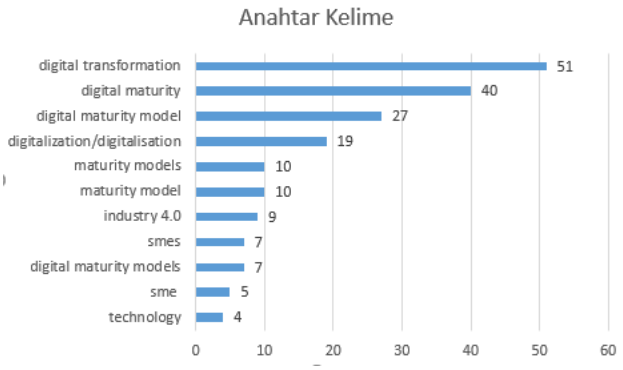
Şekil 7. Yayınlara göre atıf sayısı, Kaynak: Scopus,
(Number of citations by publications, Source: Scopus)



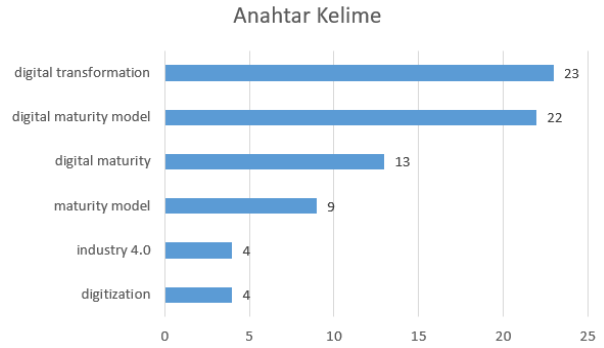
Şekil 8. Yayınlara göre atıf sayısı, Kaynak: WoS
(Number of citations by publications, Source: WoS)

3.5.Anahtar kelime incelemesine ilişkin bulgular(Findings according to keyword analysis)

Anahtar kelimeler için yapılan incelemede en az 4 defa görülen anahtar kelimelerin listesi Şekil 9 ve Şekil 10'daki gibidir. Scopus veri tabanında en çok "digital transformation" ve "digital maturity" , WoS veri tabanında en çok "digital transformation" ve "digital maturity model" kelimelerinin anahtar kelime olarak kullanıldıkları görülmektedir.

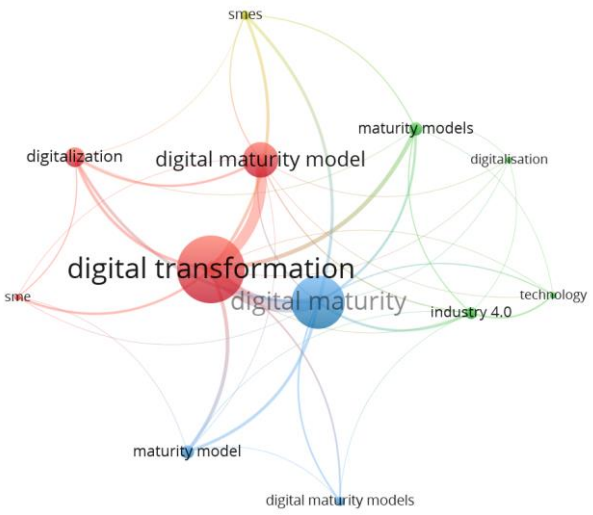


Şekil 9. En çok kullanılan anahtar kelimeler, Kaynak: Scopus
(Most used keywords, Source: Scopus)



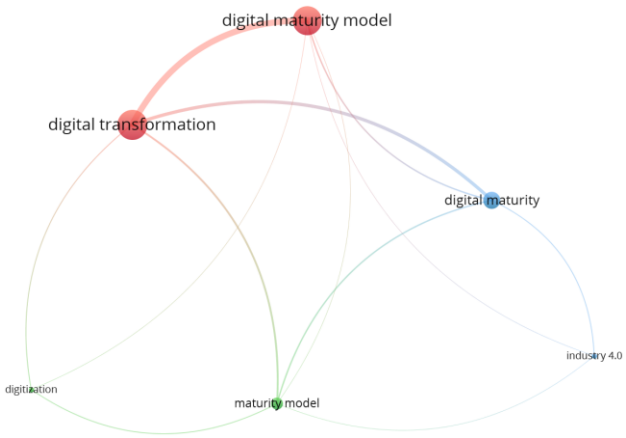
Şekil 10. En çok kullanılan anahtar kelimeler, Kaynak: WoS
(Most used keywords, Source: WoS)

Anahtar kelimeler için yapılan incelemede en az 4 defa görülen anahtar kelimelerin birlikte kullanımlarını gösteren haritalar Şekil 11 ve Şekil 12'de sunulmuştur. Her iki veri tabanında en çok "digital transformation" ve "digital maturity model" kelimelerinin anahtar kelime olarak birlikte kullanıldıkları görülmektedir.



Şekil 11. Anahtar kelime dağılımı, Kaynak: Scopus, Materyal: VOSviewer

(Keyword distribution, Kaynak: Scopus, Materyal: VOSviewer)



Şekil 12. Anahtar kelime dağılımı, Kaynak: WoS, Materyal VOSviewer
(Keyword distribution, Kaynak: WoS, Materyal: VOSviewer r)

3.6. Dijital olgunluk modellerinin incelenmesi (Analysing digital maturity models)

İncelenen sistematik literatür tarama sonuçlarına dayanarak, dijital olgunluk modellerinin yaygın olarak araştırılan bir konu olduğu söylenebilir [24]. Scopus ve WoS veri tabanları incelendiğinde birçok farklı sektör için geliştirilen dijital olgunluk modeli olduğu görülmüştür. Örneğin, eğitim kurumları [25], havalimanları [26], savunma sanayi sektörü [27], telekomünikasyon servis sağlayıcıları [28] için dijital olgunluk modelleri önerildiği görülmüştür.

Bu çalışma kapsamında yapılan makale incelemesi, dijital olgunluk modellerinin genellikle boyutlar temelinde ele alındığını ortaya koymuştur. Dijital olgunluk modeli boyutları, bir kuruluşun dijital kabiliyetlerini; teknoloji benimseme, kurum kültürü, liderlik, strateji ve müşteri yönetimi gibi çeşitli yönlerden temsil etmektedir [29]. Dijital olgunluk modellerine boyutların dâhil edilmesini destekleyen akademik çalışmalar mevcuttur, zira bir olgunluk modelinin, bir kuruluşun mevcut dijital durumunu farklı boyutlar üzerinden değerlendirmesine olanak tanımak için kullanılabileceğini belirten çalışmalar mevcuttur[29]. Dijital olgunluk modellerindeki aşama veya seviyeler, bir kuruluşun her bir boyuta ilişkin ilerleme veya olgunluk seviyesini göstermektedir. Dijital olgunluk modelindeki aşama veya seviyeler, kuruluşların zaman içinde dijital kabiliyetlerini geliştirirken izleyecekleri bir yol haritası sunmaktadır.

Bu çalışma kapsamında, bibliyometrik analiz sonucunda elde edilen sonuçlar doğrultusunda, Scopus veri tabanında listelenen 45 adet, WoS veri tabanında listelenen 25 adet İngilizce dilinde ve makale türündeki çalışmada ele alınan dijital olgunluk modelleri boyut ve seviyeleri açısından değerlendirilmiştir (Ek.1). 22 adet makalenin hem Scopus hem de WoS veri tabanında listelendiği görülmüştür. Scopus veri tabanında, yer alan 13 makalede dijital olgunluk modelinin hem boyut hem de seviye; 11 makalede sadece boyut; 9 makalede sadece seviye bakımından ele alındığı görülmüştür. WoS veri tabanında yer alan 8 makalede dijital olgunluk modelinin hem boyut hem de seviye; 8 makalede sadece boyut; 5 makalede sadece seviye bakımından ele alındığı görülmüştür.

Ek.1'de sunulan makale incelemesi sonucuna göre dijital olgunluk modellerinin çoğunlukla, küçük ve orta büyüklükteki işletmeler, eğitim, sağlık ve imalat sektörleri için önerildiği görülmüştür. Dijital olgunluk boyutu açısından ele alınan konular incelendiğinde, en çok teknoloji (15), strateji (13), organizasyon (12), kültür (11), yönetim/liderlik (11), müşteri (9) boyutlarının ele alındığı belirlenmiştir. Elde edilen bu sonuçlar, 45 olgunluk modelinin incelendiği, toplamda 100'den fazla boyutun tanımlandığı ve bu boyutlar arasında karşılaştırmalı analiz yapıldıktan sonra organizasyon, teknoloji, strateji, müşteri, çalışan, kültür ve süreç dönüşü/iş süreci boyutlarının temel boyutlar olarak belirledikleri çalışma sonucuyla benzerlik göstermektedir [10]. Yapılan başka bir çalışmada benzer şekilde, dijital olgunluk modellerinde en yaygın olarak ele alınan konular; teknoloji, strateji, insan, organizasyon, ürün, kültür, yönetim/liderlik, iş süreçleri ve müşteriler olarak belirlenmiştir [30].

Dijital dönüşüm sürecinin tetikleyicisi ve göz ardı edilemez bir parçası olan teknolojinin en çok ele alınan konu olması şaşırtıcı değildir. Çeşitli çalışmalarda, teknoloji boyutunun, dijital iş süreçlerini desteklemek amacıyla teknoloji planlamasını ve entegrasyonunu sağlayan yetenekler, teknolojik altyapı, veri kullanımı, bilgi sistemleri ve yönetimi açısından ele alındığı görülmüştür [1-10-31].

Son yıllarda yapılan çalışmalarda, yapay zeka, bulut bilişim, büyük veri analitiği gibi dijital teknolojilerin kaynak yönetimini optimize ederek ve israfı azaltarak çevresel sürdürülebilirliği önemli ölçüde artırabileceği ortaya konulmuştur [32-33]. Dijital teknolojilerin benimsenmesi üretkenliği artırabileceği gibi doğru yönetilmediği takdirde daha yüksek enerji tüketimine de neden olabilir [34]. Başka bir ifadeyle dijital dönüşüm süreci artan enerji tüketimi ve elektronik atık gibi zorlukları da beraberinde getirmektedir [35-35-37]. Geliştirilen dijital olgunluk modellerinde, sürdürülebilirlik ve çevresel faktörlerin sınırlı bir şekilde ele alındığı, dijital dönüşümün sürdürülebilirliğinin çevresel etkiler açısından değerlendirilmesine nadiren odaklanıldığı görülmüştür [31].

4. SONUÇ (RESULT)

Dijital dönüşüm yolculuğu birçok organizasyon için oldukça karmaşıktır. Bu yolculuğu planlı bir şekilde yönetmek ve yürütmek için dijital dönüşüm modellerinden faydalanmak süreci kolaylaştıracaktır. Dijital dönüşümün her aşamasında, organizasyonun mevcut durumunu

değerlendirebileceği referans bir rehberin veya çerçevenin olması, yapılacak çalışmaların daha veremli bir şekilde planlanmasına yardımcı olacaktır. Organizasyonların nereden başlayacağını bilmesi, anahtar/kilit faaliyetlere odaklanması sürecin etkili bir şekilde yürütülmesine fayda sağlayacaktır. Dijital olgunluk modelleri, kuruluşun güçlü olduğu alan ve yetkinliklerini tanımlamakla kalmaz, aynı zamanda dijital dönüşümü kolaylaştırmak için ele alınması gereken düşük yetkinliğe sahip yönleri de ortaya çıkarır. Böylece olgunluk modeli, dijital dönüşüm süreci boyunca başvurulacak bir rehber veya araç olarak sürecin etkili bir parçası olur.

Bu çalışma, dijital olgunluk modellerinin araştırılması, dijital dönüşüm süreçlerinin daha iyi anlaşılması, organizasyonel yeteneklerin değerlendirilmesi, stratejik yönlendirme sağlanması ve sektörel uygulamaların geliştirilmesi gibi çeşitli alanlarda literatüre katkılar sağlamaktadır. Bibliyometrik analiz bulguları, bu alanda araştırma yapmak isteyen araştırmacılara kapsamlı bir genel bakış sunarak, gelecekteki çalışmalar için odaklanılabilecek potansiyel alanları belirlemektedir. Analiz bulgularına göre Scopus ve WoS veri tabanlarında dijital olgunluk modeli konusunda 2020 yılı itibariyle yayın sayısında belirgin bir artışın olduğu ve ülke incelemesine göre Almanya ve Türkiye'nin yayın sayısı olarak dikkat çeken ülkeler olduğu görülmüştür.

Ayrıca, dijital olgunluk modellerinin boyut ve seviyelerine yönelik yapılan sistematik inceleme, literatürdeki mevcut kavramsal çerçevelerde en çok teknoloji, strateji ve organizasyon boyutlarının ele alındığını ve sonucun yapılan önceki çalışmalarla tutarlı olduğu sonucunu ortaya koymaktadır[24-30]. Dijital dönüşüm, sürdürülebilirliği artırmak için birçok konuda fırsat sunarken aynı zamanda artan enerji tüketimi ve elektronik atık gibi zorlukları da beraberinde getirmektedir. Dijital dönüşümün sürdürülebilirliği açısından çevresel etkileri ele alan dijital olgunluk modellerinin geliştirildiği, ancak bu konunun çoğu modelde yeterince ele alınmadığı görülmüştür.

KAYNAKLAR(REFERENCES)

[1] K. Warner, M. Wäger, "Building dynamic capabilities for digital transformation: an ongoing process of strategic renewal", *Long Range Planning*, 52(3), pp. 326-349, 2019.

[2] I. Iyamu, A. Xu, O. Gómez-Ramírez, A. Ablona, H. Chang, G. McKee, M. Gilbert, "Defining digital public health and the role of digitization, digitalization, and digital transformation: scoping review", *Jmir Public Health and Surveillance*, 7(11), e30399, 2021.

[3] J. Poeppelbuss, B. Niehaves, A. Simons, & J. Becker, "Maturity models in information systems research: literature search and analysis", *Communications of the Association for Information Systems*, 29(1), 27, 2021.

[4] A. K. Pani, H. S. Pramanik, "Digital transformation of organizations–defining an emergent construct", *In Re-imagining Diffusion and Adoption of Information Technology and Systems: A Continuing Conversation: IFIP WG 8.6 International Conference on Transfer and Diffusion of IT, TDIT 2020, Proceedings*, Part II pp. 511-523, 2020.

[5] M. Kupriyanova, E. Evdokimova, I. Solovyova, I. Simikova, "Methods of developing digital maturity models for manufacturing companies", *In E3S Web of Conferences*, Vol. 224, p. 02034, 2020.

[6] R. Teichert, "Digital transformation maturity: a systematic review of literature" *Acta Universitatis Agriculturae Et Silviculturae Mendelianae Brunensis*, 67(6), pp. 1673-1687, 2019.

[7] P. Jafari, "Investigating a roadmap for digital work practices based on a maturity model approach". *Ghent University. Faculty of Economics and Business Administration, Ghent, Belgium*, 2024.

[8] V. Okrepilov, "Digital maturity level assessment as an element of digitalization of russian economy", *European Proceedings of Social and Behavioural Sciences*, 2021.

[9] D. V. Kuzin, "Problems of digital maturity in modern business", *The World of New Economy*, 13(3), pp.89-99, 2019.

[10] T. Haryanti, N. A. Rakhmawati, A. P. Subriadi, "The extended digital maturity model", *Big Data and Cognitive Computing*, , 7(1), 17, 2023.

[11] A. A. Tubis, "Digital maturity assessment model for the organizational and process dimensions", *Sustainability*, 15(20), 15122, 2023.

[12] T. Thordsen, M. Murawski, M. Bick, "How to Measure Digitalization? A Critical Evaluation of Digital Maturity Models", **In Responsible Design, Implementation and Use of Information and Communication Technology**., 19th IFIP WG 6.11 Conference on e-Business, e-Services, and e-Society, I3E 2020, Skukuza, South Africa, Proceedings, Part I 19 pp. 358-369, Springer International Publishing April 6–8, 2020.

[13] S. Perera, X. Jin, P. Das, K. Gunasekara, M. Samaratunga, "A Strategic Framework for Digital Maturity of Design and Construction through a Systematic Review and Application" J. *Ind. Inf. Integr*., 31, 100413, 2023.

[14] C.van Tonder, B. Bossink, C. Schachtebeck, & C. Nieuwenhuizen, "Key dimensions that measure the digital maturity levels of small and medium-sized enterprises (smes)", *Journal of Technology Management &Amp; Innovation*, 19(1), pp.110-130, 2024.

[15] E.Gökalp, V. Martinez, "Digital transformation capability maturity model enabling the assessment of industrial manufacturers", *Computer Ind*. 132, 103522, 2021.

[16] G. Remane, A. Hanelt, F. Wiesboeck, L. Kolbe, "Digital Maturity in Traditional Industries—An Exploratory Analysis", **In Proceedings of the 25th European Conference on Information Systems (ECIS)**, Guimarães, Portugal, 5–10 June 2017

[17] South Australian Government, Department of the Premier and Cabinet, "The Digital Strategy Toolkit", https://www.dpc.sa.gov.au/__data/assets/pdf_file/0008/46565/Digital_Transformation_Toolkit_Guide.pdf , Erişim tarihi: 18.09.2024

[18] Türkiye Cumhuriyeti Cumhurbaşkanlığı, Strateji ve Bütçe Başkanlığı, "On İkinci Kalkınma Planı (2024-2028)", p.234 (964.4), 2023 https://www.sbb.gov.tr/wp-content/uploads/2023/12/On-Ikinci-Kalkinma-Plani_2024-2028_11122023.pdf, Erişim tarihi: 10.07.2024

[19] F. Blatz, R. Bulander & M. Dietel, "Maturity model of digitization for SMEs", **In 2018 IEEE International Conference on Engineering, Technology and Innovation** (ICE/ITMC), pp.1-9, June, 2018.

[20] G. Erboz, H. Abbas, S. Nosratabadi, "Investigating supply chain research trends amid Covid-19: a bibliometric analysis". Management Research Review, 46(3), pp. 413-436, 2023. [21] M. A. Khan, D. Pattnaik, R. Ashraf, I. Ali, S. Kumar, N. Donthu, "Value of special issues in the journal of business research: A bibliometric analysis", *Journal of business research*, 125, pp.295-313. 2021.

[23] N. Donthu, S. Kumar, D. Mukherjee, N. Pandey, L. Marc, "How to conduct a bibliometric analysis: an overview and guidelines", *Journal of Business Research*, 133, pp.285-296, 2021.

[23] W. Hassan, A. E. Duarte, "Bibliometric analysis: A few suggestions", Current Problems in Cardiology, Volume 49, Issue 8, 2024.

[24] L. Ochoa-Urrego, J. Peña-Reyes, "Digital maturity models: a systematic literature review", Management for Professionals, pp.71-85, 2021.

[25] M. Al-Ali, A. Marks, "A digital maturity model for the education enterprise", Perspectives: Policy and Practice in Higher Education, 26(2), pp.47-58, 2022.

[26] N. Halpern, T. Budd, P. Suau-Sanchez, S. Bråthen, D. Mwesiumo, "Conceptualising airport digital maturity and dimensions of technological and organisational transformation", *Journal of Airport Management*, 15(2), pp. 182-203, 2021.

[27] E. E. Nebati, B. Ayvaz, A. O. Kusakci, "Digital transformation in the defense industry: A maturity model combining SF-AHP and SF-TODIM approaches", Applied Soft Computing, 132, 109896, 2023.

[28] O. Valdez-de-Leon, "A digital maturity model for telecommunications service providers", Technology innovation management review, 6(8), 2016.

[29] R. Duncan, R. Eden, L. Woods, I. Wong, C. Sullivan, "Synthesizing dimensions of digital maturity in hospitals: systematic review", *Journal of Medical Internet Research*, 24(3), e32994, 2022.

[30] N. Kaszás, I. Ernszt, B. Jakab, "The Emergence of Organizaional and Human Factors in Digital Maturity Models", *Management: Journal of Contemporary Management Issues*, 28(1), pp. 123-135, 2023.

[31] D. Merdin, F. Ersöz, H. Taşkın, "Digital Transformation: Digital Maturity Model for Turkish Businesses", *Gazi University Journal of Science*, 36(1), pp. 263-282, 2022.

[32] G. Vial, "Understanding digital transformation: A review and a research agenda". *J. Strateg. Inf. Syst.* 28, pp.118–144, 2019.

[33] A. K Feroz, H. Zo, A. Chiravuri, "Digital transformation and environmental sustainability: a review and research agenda", *Sustainability*, 13(3), 1530, 2021.

[34] J. Yan, J. Li, X. Li, & Y. Liu, "Digital transition and the clean renewable energy adoption in rural family: evidence from broadband china", *Frontiers in Ecology and Evolution*, 11, 2023.

[35] L. Zhao, Y. Zhang, & H. Zhang, "Research on the impact of digital literacy on farmer households' green cooking energy consumption: evidence from rural china", *International Journal of Environmental Research and Public Health*, 19(20), 13464, 2022.

[36] H. Wen, C. Lee, & Z. Song, "Digitalization and environment: how does ict affect enterprise environmental performance?", Environmental Science and Pollution Research, 28(39), pp.54826-54841, 2021.

[37] J. Lan, H. Wen, "Industrial digitalization and energy intensity: evidence from china's manufacturing sector". Energy Research Letters, 2(2), 2021.

[38] M. A. Abdullah, "Digital maturity of the Egyptian universities: goal-oriented project planning model", Studies in Higher Education, pp.1-23, 2023

[39] A. Altundag, M. Wynn, "Advanced Analytics and Data Management in the Procurement Function: An Aviation Industry Case Study", Electronics, 13(8), pp.1554, 2024.

[40] A. Aras, G. Büyüközkan, "Digital Transformation Journey Guidance: A Holistic Digital Maturity Model Based on a Systematic Literature Review", Systems, 11(4), pp.213, 2023.

[41] A. Aristovnik, D. Ravšelj, E. Murko, "Decoding the Digital Landscape: An Empirically Validated Model for Assessing Digitalisation across Public Administration Levels", Administrative Sciences, 14(3), pp. 41, 2024.

[42] A. S. Barry, S. Assoul, N. Souissi, "Strengths and Weaknesses of Digital Maturity Models", *Journal of Computer Science*, 19 (6), pp. 727.738, 2023.

[43] T. Benazzouz, K. Auhmani, "Digital maturity assessment model for pharmaceutical supply chain: a patient and hospital-centred development", *International Journal of Healthcare Management*, 17(2), pp. 261-284, 2024.

[44] Bouncken, R., & Schmitt, F. SME family firms and strategic digital transformation: inverting dualisms related to overconfidence and centralization. *Journal of Small Business Strategy*, 32(3), 1-17, (2022).

[45] S. Bozkus Kahyaoglu, S. Durst, E. Coskun, "To digitalize or not? COVID effect in medium sized companies based on three cases", EDPACS, 68(4), pp.1-25, 2023.

[46] G. Büyüközkan, M. Güler, "Analysis of companies' digital maturity by hesitant fuzzy linguistic MCDM methods", *Journal of Intelligent & Fuzzy Systems*, 38(1), pp.1119-1132, 2020.

[47 T. Červinka, "Digital transformation of strategic management of SMEs in the Czech Republic", *Journal of Information and Organizational Sciences*, 47(2), pp. 387-400, 2023.

[48] B. Cognet, J. P. Pernot, L. Rivest, C. Danjou, "Digital maturity models: comparing manual and semi-automatic similarity assessment frameworks", *International Journal of Product Lifecycle Management*, 13(4), pp.291-316, 2021.

[49] B. Cognet, J. P. Pernot, L. Rivest, C. Danjou, "Systematic comparison of digital maturity assessment models", Journal of Industrial and Production Engineering, 40(7), pp. 519-537, 2023.

[50] M. Ćurak, M. Duvnjak, M. Pervan, "The relationship between the digital maturity and efficiency of Croatian non-life insurers: Exploratory research", *Journal of Economics and Management*, 46(1), pp.55-78, 2024.

[51] M. Diller, M. Asen, T. Späth, "The effects of personality traits on digital transformation: Evidence from German tax consulting", *International Journal of Accounting Information Systems*, 37, 2020.

[52] S. Eichstädt, M. Gruber, A. P. Vedurmudi, "Integrating metrological principles into the Internet of Things: a digital maturity model for sensor network metrology", tm-Technisches Messen, 91(1), pp.17-31, 2024.

[53] M. Guarino, M. A. Di Palma, T. Menini, M. Gallo, "Digital transformation of cultural institutions: a statistical analysis of Italian and Campania GLAMs", Quality & quantity, 54, pp. 1445-1464, 2020.

[54] X. Han, M. Zhang, Y. Hu, Y. Huang, "Study on the Digital Transformation Capability of Cost Consultation Enterprises Based on Maturity Model", Sustainability, 14(16), 10038, 2022.

[55] I. Herceg, V. Kuč, V. M. Mijušković, T. Herceg, "Challenges and driving forces for industry 4.0 implementation", Sustainability, 12(10), 4208, 2020.

[56] V. Hrosul, D. Galoyan, T.Mkrtchyan, A. Volosov, H. Balamut, A. Kolesnyk, "Assessment of digital maturity, the transformation of business models in the context of digital transformation", REICE: Revista Electrónica de Investigación en Ciencias Económicas, 11(32), pp. 81-105, 2023.

[57] A. Imboden, S. Grèzes-Bürcher, D. Fumeaux, E. Fragnière, M. Fux, "Piloting a digital maturity model for smart destinations", *International Journal of Technology Marketing*, 16(4), pp.304-317, 2022.

[58] T. Kafel, A. Wodecka-hyjek, R. Kusa, "Multidimensional public sector organizations' digital maturity model", Administration & Public Management Review, (37), 2021.

[59] K. J. Kupilas, V. Rodríguez Montequín, M. Díaz Piloñeta, C. Alonso Álvarez, "Sustainability and digitalisation: Using Means-End Chain Theory to determine the key elements of the digital maturity model for research and development organisations with the aspect of sustainability", *Advances in production engineering and management*, 2022.

[60] L. Ladu, C. Koch, P. A. Ashari, K. Blind, P. Castka, "Technology adoption and digital maturity in the conformity assessment industry: Empirical evidence from an international study", *Technology in Society*, 77, 102564, 2024.

[61] Y. Liu, F. Pan, "Digital transformation value creating of manufacturing enterprises based on the Internet of Things and data encryption technology", *Soft Computing*,, pp.1-12, 2023.

[62] I. Merzlov, E. Shilova, "A Digital Maturity Model for Organizations: An Approach to Assessment and Case Study", *International Journal of Systematic Innovation*, 7(2), pp. 22-36, 2022.

[63] C. Minonne, R. Wyss, K. Schwer, D. Wirz, C. Hitz, "Digital maturity variables and their impact on the enterprise architecture layers", Problems and perspectives in management,16, Iss. 4, pp. 141-154, 2018.

[64] P. Phiri, H. Cavalini, S. Shetty, G. Delanerolle, "Digital Maturity Consulting and Strategizing to Optimize Services: Overview", *Journal of Medical Internet Research*, 25, e37545, 2023.

[65] W. G. Ryan, A. Fenton, W. Ahmed, P. Scarf, "Recognizing events 4.0: The digital maturity of events", *International Journal of Event and Festival Management*, 11(1), pp. 47-68, 2020.

[66] Á. Sándor, Á. Gubán, "A measuring tool for the digital maturity of small and medium-sized enterprises", Management and Production Engineering Review,14, 2021.

[67] Á. Sándor, Á. Gubán, "A multi-dimensional model to the digital maturity life-cycle for smes", *International Journal of Information Systems and Project Management*, 10(3), pp. 58-81, 2022.

[68] F. J. Santos, C. Guzmán, P. Ahumada, "Assessing the digital transformation in agri-food cooperatives and its determinants", *Journal of Rural Studies*, pp.105, 103168, 2024.

[69] P. P. Senna, A. C. Barros, J. B. Roca, A. Azevedo, "Development of a digital maturity model for Industry 4.0 based on the technology-organization-environment framework", *Computers & Industrial Engineering*,, pp. 185, 109645, 2023.

[70] R. Teichert, "A Model for Assessing Digital Transformation Maturity for Service Provider Organizations", *European Journal of Business Science and Technology*, pp. 205, 2023.

[71] T. Thordsen, M. Bick, "A decade of digital maturity models: much ado about nothing?" Information Systems and e-Business Management, pp. 1-30, 2023.

[72] C. E. Tømte, J. Smedsrud, "Governance and digital transformation in schools with 1: 1 tablet coverage", 2023.

[73] M. Voss, D. Jaspert, C. Ahlfeld, L. Sucke, "Developing a digital maturity model for the sales processes of industrial projects", *Journal of Personal Selling & Sales Management*, 44(1), pp.7-28, 2024.

[74] C. Williams, B. Krumay, D. Schallmo, E. Scornavacca, "Digital Maturity Model for SMEs: Validation Through a Mixed-Method Approach", Pacific Asia Journal of the Association for Information Systems, 16(1), 2, 2024.

[75] C. A. Williams, D. Schallmo, E. Scornavacca, "How applicable are digital maturity models to SMEs?: A conceptual framework and empirical validation approach." *International Journal of Innovation Management*, 26(03), 2240010, 2022.

[76] L. Woods, R. Eden, R. Duncan, Z. Kodiyattu, S. Macklin, C. Sullivan, "Which one? A suggested approach for evaluating digital health maturity models", Frontiers in Digital Health, 4, 246, 2022.

[77] F. Zaoui, N. Souissi, "Digital Maturity Assessment – A Case Study", *Journal of Computer Science*, 18 (8), pp. 724-731, 2022.

[78] T. Von Leipzig, M. Gamp, D. Manz, K. Schöttle, P. Ohlhausen, G.Oosthuizen, D. Palm, K. Von Leipzig, "Initialising customer-orientated digital transformation in enterprises", Procedia Manufacturing, 8, pp. 517-524, 2017.

[79] D. Fayon, M. Tartar, "Transformation digitale: 5 leviers pour l'entreprise", Pearson Education France, 2014.

EK.1 Dijital Olgunluk Modellerinin İncelenmesi

Veri tabanı	Makale	Alan	Dijital olgunluk modeli Boyut /Faktör	Ele Alınan Seviye	Açıklama
Scopus WoS	[38]	Eğitim Sektörü	1. Liderlik, planlama ve yönetim, 2.Kalite güvencesi, 3. Bilimsel araştırma, 4. Dijital öğrenme ve öğretme, 5. Toplum Hizmeti, 5. Ekipman ve teknolojik altyapı, 6. Teknolojik kültür	Temel Başlangıç e-Etkin (e-enabled) e-Güvenli (e-confident) e-Olgun (e-mature)	Üniversiteler için dijital dönüşüm olgunluk modeli önerilmiştir.
Scopus	[25]	Eğitim Sektörü	1. DT Vizyon, strateji, liderlik ve iletişim, 2. DT gezegeni, beceri ve bilgi, 3. DT süreçleri, kontrolleri ve teknolojileri, 4. DT teknoloji altyapısı, 5. Müşteri ve paydaşları anlama ve onlarla iletişim kurma yaklaşımı		Eğitim kurumlarında dijital dönüşüm olgunluk değerlendirmesini ölçmek için yeni bir çerçeve önerilmiştir.
Scopus WoS	[39]	Havacılık Endüstrisi	1.Teknoloji, 2.Süreç, 3.İnsanlar, 4.Yapı	Temel Orta Seviye Standartlaştırılmış Dönüştürülmüş	Havacılık sektörü için 4 boyutlu bir model önerilmiştir.
Scopus WoS	[40]	Sektör Bağımsız	1. Dijital Strateji, 2. Dijital Değer, 3. Dijital Süreçler, 4. Dijital Teknoloji ve Veri, 5. Dijital Çalışma, 6. Dijital Yönetişim	Niyet Başlangıç Adapte Uygulayıcı Dönüştürücü	Literatür analizi yapılmış ve yeni bir dijital olgunluk modeli önerilmiştir.
Scopus	[41]	Kamu Yönetimi	1. Teknoloji, 2. Süreçler, 3. Yapı, 4. Örgütsel Kültür, 5.İnsanlar		Çalışmanın sonuçları, Slovenya'daki bakanlıkların, özellikle BİT geliştirme ve entegrasyonunun çeşitli yönlerinde, genel olarak belediyelerden daha gelişmiş bir dijital altyapıya sahip olduğunu göstermiştir.
Scopus	[42]	Sektör Bağımsız			Belirlenen 8 dijital olgunluk modelinin güçlü ve zayıf yönlerinin belirlenmesi amacıyla yapılan bir çalışmadır.
Scopus	[43]	Sağlık Sektörü	1. BT ve Dijital Teknolojiler, 2. Strateji ve Süreç Yönetimi, 3. İnsan Kaynakları ve Kültür, 4. Tedarik Zinciri ve Hasta Yolculukları Yönetimi	Seviye 0: Dijitalleşme mevcut değil Seviye 1: Dijitalleştirmenin başlatılması Seviye 2: Dijitalleşmenin entegre edilmesi Seviye 3: Tam Dijitalleşme	Dijital Olgunluk Değerlendirme Modeli sunulmuştur. Model 4 boyut, 4 olgunluk seviyesi ile değerlendirilebilecek toplam 65 alt boyut olarak belirlenmiştir.
Scopus	[44]	Küçük ve Orta Ölçekli Aile İşletmeleri			Aile şirketlerinde yönetici ve dijital dönüşüm yaklaşımı ile ilgili bulgulara yer verilmiştir. Çalışma, KOBİ aile şirketlerinin dijital dönüşümlerinde daha düşük düzeyde strateji oluşturma ve pragmatik-artırmacı bir yaklaşım sergilediklerini ortaya koymuştur.
Scopus	[45]	Küçük ve Orta Ölçekli İşletmeler	1. Strateji, 2. Liderlik, 3. Projeler, 4. Operasyonlar, 5. Kültür, 6. İnsanlar, 7. Yönetişim, 8. Teknoloji	Habersiz Kavramsal Tanımlanmış Entegre	[78] çalışmasından uyarlanmış model ele alınmıştır.

				Dönüştürülmüş	
Scopus WoS	[46]	Bankacılık Sektörü	1. Kültür, 2. Organizasyon, 3. Teknoloji, 4. Anlayış (Insights)		Önerilen model, bankacılık sektöründe test edilmiştir.
Scopus WoS	[47]	Küçük ve Orta Ölçekli İşletmeler			Çek Cumhuriyeti'nde seçilen KOBİ'lerde stratejik yönetim yaklaşımı ve stratejik yönetimin dijital olgunluk üzerindeki algılanan etkisine ilişkin bir araştırma yapılmıştır.
Scopus	[48]	Sektör Bağımsız			Manuel ve yarı otomatik benzerlik değerlendirme çerçeveleriyle IMPULS ve PwC dijital olgunluk modelleri karşılaştırılmıştır.
Scopus	[49]	Sektör Bağımsız			Önerilen çerçeve, 12 boyut ve 58 alt boyuta göre yapılandırılmış, 263 anahtar kelime kullanılarak ve 451 KPI'nın tanımlanması yoluyla 13 olgunluk modeli üzerinde uygulanmıştır.
Scopus	[50]	Sigorta Sektörü	1. Kültür, 2. Teknoloji, 3. Organizasyon, 4. İçgörüler		Çalışmanın amacı, Hırvat hayat dışı sigorta şirketlerinin dijital olgunluğu ile verimlilikleri arasındaki mevcut ilişkiyi analiz etmektir.
Scopus WoS	[51]	Vergi Danışmanlığı	Faktör 1: İş modeli dönüşümü Faktör 2: Dijital işbirliği Faktör 3: Uzaktan erişim Faktör 4: Dijital iletişim Faktör 5: Ara Bağlantılık		Vergi danışmanlarının Büyük-Beşli (Big-Five) kişilikleri ile dijitalleşme düzeyleri arasındaki ilişki araştırılmıştır.
Scopus WoS	[52]	Sensör Ağları			Sensör ağı metrolojisinin gelişimi için yol haritasının geliştirilmesinde temel olarak kullanılabilecek farklı dijital olgunluk seviyeleri ele alınmıştır.
Scopus	[53]	Kültürel kurumlar (Galeriler, Kütüphaneler, Arşivler ve Müzeler.)			İtalya ve Kampaniya bölgesindeki kültürel kurumların (galeriler, kütüphaneler, arşivler ve müzeler) dijital tesisler ve hizmet sunma konusundaki yeteneklerine ilişkin bir ön tanımlayıcı analiz ortaya konulmuştur.
Scopus	[26]	Havalimanı	1. Teknolojik 2. Organizasyonel	Analog süreçler Dijitalleştirme Dijitalleşme Dijital dönüşüm	Yolcu deneyimi perspektifine odaklanan bir havalimanı dijital olgunluk modeli geliştirilmiş ve kavramsallaştırılmıştır.
Scopus WoS	[54]	Maliyet Danışmanlığı İşletmeleri	1. Üst düzey tasarım, 2. Altyapı, 3. Maliyet danışmanlığı iş süreci, 4. Profesyonel yönetim,5. Kapsamlı entegrasyon, 6. Dijital maliyet performansı	Yeni Başlayan Başlangıç Uygulayıcı Öncü Lider	Maliyet danışmanlığı işletmeleri için 7 boyutlu ve 5 seviyeli bir dijital olgunluk modeli önerilmiştir.
Scopus WoS	[10]	Sektör bağımsız	1. Yapılanma ve Organizasyon, 2. Teknoloji, 3. Strateji, 4. Müşteri, 5. Çalışan, 6. Kültür,7. Dönüşüm Süreci	Tamamlanmamış Gerçekleştirilmiş Yönetilen Kurulmuş	Dijital Dönüşüm Öz Değerlendirme olgunluk modeli geliştirilmiştir.

				Öngörülebilir En Uygun Hale Getirme (Optimizing)	
Scopus WoS	[55]	İmalat işletmeleri	1. Müşteri, 2. Strateji, 3. Teknoloji, 4. Operasyonlar, 5. Organizasyon ve Kültür	1. Düşük dijital olgunluk, 2. Ortalama dijital olgunluk, 3. Yüksek dijital olgunluk	İmalat işletmeleri için 5 boyutlu ve 3 seviyeli bir dijital olgunluk modeli önerilmiştir.
WoS	[56]	Lojistik sektörü	1. Yönetim, 2. Malzeme Akışları, 3. Bilgi Akışları	Görmezden Gelme Tanımlama Sahiplenme Yönetme Entegre olma	Lojistik sektörü için 3 boyutlu ve 5 seviyeli bir dijital olgunluk modeli önerilmiştir.
Scopus	[57]	Turizm Sektörü			Olgunluk modeli, İsviçre'nin az bilinen üç turizm merkezinin dijital dönüşümlerini nasıl yönettiklerine, hangi zorluklarla ve fırsatlarla karşılaştıklarına ve bu süreçte operasyonlarının hangi yönlerine öncelik verdiklerine yanıt getirmek üzere tasarlanmıştır.
Scopus	[58]	Kamu Sektörü	1. Dijitalleşme odaklı yönetim, 2. Paydaşların (ortakların) ihtiyaçlarına açıklık, 3. Çalışanların dijital yetkinlikleri, 4. Süreçlerin dijitalleştirilmesi, 5. Dijital teknolojiler, 6. E-yenilikçilik	IDM - yetersiz dijital olgunluk derecesi VLDM - çok düşük dijital olgunluk derecesi LDM - düşük dijital olgunluk derecesi MDM - orta derecede dijital olgunluk HDM - yüksek derecede dijital olgunluk FDM - tam dijital olgunluk; VHDM - çok yüksek derecede dijital olgunluk;	Kamu sektörü kuruluşlarının dijital olgunluk derecesini değerlendirmek için kullanılabilecek bir dijital olgunluk modeli sunulmuştur.
Scopus	[30]	Sektör bağımsız			Çalışma, literatürde dijital dönüşümü engelleyen faktörler olarak bilinen kurum kültürü ve insan faktörlerine odaklanmıştır. Literatürde en çok kullanılan boyutlar teknoloji, strateji, insan, organizasyon, ürün, kültür, yönetim/liderlik, iş süreçleri ve müşteriler olarak belirlenmiştir.
Scopus WoS	[59]	Araştırma ve geliştirme kuruluşları	1. Akıllı operasyonlar ve araştırma, 2.İnsan, 3.Sürdürülebilirlik, 4.Strateji ve organizasyon, 5.Akıllı Tesisler, 6.Akıllı ürünler ve hizmetler		Araştırma ve geliştirme kuruluşları için oluşturulan olgunluk modeli 6 boyut üzerinden ele alınmıştır.
Scopus	[60]	Uygunluk değerlendirme kuruluşları	1. Strateji, 2. Bilgi Teknolojisi ve Süreç Dijitalleşmesi, 3. Müşteriler, 4. Kültür ve Uzmanlık, 5. Organizasyon ve Değişim Yönetimi, 6. İnovasyon ve 7. İşbirliği.	Aşama 1- Başlangıç Aşama 2 Aşama 3 Aşama 4 Aşama 5 - Uzman	Uygunluk değerlendirme kuruluşlarına uyarlanmış bir dijital olgunluk modeli ortaya konulmuş ve uygulanmıştır.

Scopus WoS	[61]	İmalat İşletmeleri			Üç dijital dönüşüm yolu ve bunların uygulama çerçeveleri özetlenmiştir. 1. Akıllı(Smart) + Dijital İleri Üretim Dönüşümü ve Yükseltme Yolu, 2. Akıllı+ Dijital Endüstriyel Hizmet Dönüşümü ve Yükseltme Yolu, 3. Akıllı + Büyük Veri Sektörünün Dönüşümü ve Yükselme Yolu
Scopus WoS	[31]	Türk İşletmeleri	1. Strateji, 2. Müşteriler, 3. Çalışanlar, 4. Süreç Yönetimi, 5. Teknoloji ve Veri Yönetimi, 6. Kurumsal Kültür, 7. Yenilik		Çevresel faktörleri (sürdürülebilirlik) değerlendiren bir dijital olgunluk modelinin geliştirilmiş ve geliştirilen modelin doğrulama testleri yapılmıştır. Model her ölçekten ve her sektörden işletmeye uygulanabilir.
Scopus	[62]	Sektör bağımsız		Seviye 1: Yerel Dijitalleşme Seviye 2: Kısmi Dijitalleşme Seviye 3: Kapsamlı Dijitalleşme Seviye 4: Akıllı Organizasyon Seviye 5: Dijital Ekosistem	Model, iç ve dış temel iş süreçlerinin dijitalleşme seviyesinin değerlendirilmesine dayanmaktadır.
Scopus	[63]	Sektör bağımsız			Literatür taraması sonucunda, toplam 147 değişken içeren 15 dijital olgunluk modeli “strateji, iş ortamı, uygulamalar, teknoloji, fiziksel ve uygulama ve geçiş” kurumsal mimari katmanlarıyla eşleştirilmiş ve analiz edilmiştir.
Scopus WoS	[27]	Savunma Sanayi Sektörü	1.Liderlik, 2.Organizasyon ve Dijital kültür, 3.Strateji, 4.Teknoloji, 5.Operasyonlar		Savunma Sanayi Sektörü için 5 boyutlu bir dijital olgunluk modeli önerilmiştir.
WoS	[64]	Sağlık Sektörü			Dijital olgunluğu etkileyen faktörler belirlenmiştir (İnsan, Altyapı, BT Sistemleri)
Scopus WoS	[65]	Etkinlik Endüstrisi		Etkinlik 1.0 Temel Etkinlik 2.0: Gelişmekte Olan Etkinlik 3.0: Gelişme Etkinlik 4.0: Entegre	Etkinlik endüstrisinin dijital açıdan olgunlaşması ele alınmıştır. 4 seviyeden oluşan bir dijital olgunluk ölçeği önerisi sunulmuştur.
Scopus WoS	[66]	Küçük ve Orta Ölçekli İşletmeler	1. Bilgi Teknolojileri (IT), 2. Kurumsal		2 boyutlu bir çerçeve (framework) belirlenmiştir. BT boyutu (Technical solutions, Hardware, Software); Kurumsal boyut (Orgware, online presence, peopleware)
Scopus	[67]	Küçük ve Orta Ölçekli İşletmeler	1. Dijital olgunluk, 2. Kurumsal olgunluk, 3. BT yoğunluğu	Seviye 1: Başlangıç Seviye 2: Yol Gösterici Seviye 3: İleri düzey Seviye 4: Yönetilen Seviye 5: Optimize edilmiş	Geliştirilen dijital olgunluk yaşam döngüsü modeli, şirketlerin dijital olgunluğunu, organizasyonel özelliklerini ve faaliyet alanlarının BT yoğunluğunu ele almaktadır.
Scopus WoS	[68]	Tarımsal gıda sektörü	1. Altyapı, 2. Süreçler, 3. Organizasyon ve çalışanlar, 4.Ürünler ve hizmetler, 5.Müşteriler		Tarımsal gıda sektörü için 5 boyutlu bir model önerilmiştir.
Scopus WoS	[69]	İmalat Sektörü	1. Teknoloji, 2. Organizasyon, 3. Çevre	Seviye 1. Dijitalleştirme Seviye 2. İletişim	İmalat sektörü için 3 boyutlu bir model önerilmiştir.

				Seviye 3. Görünürlük Seviye 4. Şeffaflık Seviye 5. Tahmin Edilebilirlik Seviye 6. Esneklik/Uyarlanabilirlik	
Scopus	[6]	Sektör bağımsız			Sistematik literatür taraması yapılmış, 22 dijital olgunluk modeli ele alınmıştır ve en yaygın kullanılan olgunluk boyutları belirlenmiştir.
Scopus	[70]	Hizmet Sağlayıcı Kuruluşlar	Faktör 1: Dijital Hizmet Stratejisi, Faktör 2: Dijital Yetkinlik, Faktör 3: Müşteri Deneyimi, Faktör 4: Dijital Teknoloji, Faktör 5: Dijital Hizmet İş Modeli, Faktör 6: Akıllı Hizmetler, Faktör 7: Dijital Liderlik ve Organizasyon, Faktör 8: Dijital Kültür	Seviye 0: Her zamanki gibi (business-as-usual) Seviye 1: Test Etme ve Öğrenme Seviye 2: Resmileştirilmiş ve Etkinleştirici Seviye 4: Stratejik ve Entegre Seviye 5: Birleşik ve Yönetilen Seviye 6: Yenilikçi ve Uyarlanabilir	Dijital Hizmet Dönüşümü Olgunluk Öz Değerlendirme Modeli sunulmuştur. Model 8 faktör, 27 alt faktör ve 403 ilgili özel değerlendirme maddesinden oluşmaktadır.
Scopus WoS	[71]	Sektör bağımsız			2011 ve 2022 yılları arasında geliştirilen dijital olgunluk modelleri için bir karşılaştırma yapılmıştır.
Scopus WoS	[72]	Eğitim Sektörü			İskandinavya'daki okullarda 1:1 tablet kullanımının artma eğilimi sonucu, ulusal rehber dokümanların, okullarda dijital dönüşümü gerçekleştirmelerine yardımcı olup olamayacağı tartışılmıştır.
WoS	[28]	Telekomünikasyon Servis Sağlayıcılar	1. Strateji, 2. Organizasyon, 3. Müşteri, 4. Değer zinciri/ekosistem, 5. Operasyonlar, 6. Teknoloji, 7. Yenilik	Başlatılmadı Başlatılıyor Etkinleştiriliyor Bütünleştiriliyor Optimize ediliyor Öncü	Telekomünikasyon Sektörü için 7 boyutlu ve 6 seviyeli bir dijital olgunluk modeli önerilmiştir.
Scopus WoS	[73]	İşletmeler arası satış süreçleri	1. Dijital beceriler, 2. Dijital iş kültürü, 3. Dijital iş organizasyonu, 4. Dijital araçlar, 5. Lider sorumluluğu	Başlangıç Temel dijitalleştirme Ortalama dijitalleşme Gelişmiş dijitalleştirme Dijital odaklı	Geliştirilen olgunluk modelinin, farklı sektörlerdeki şirketlere satış süreçlerinin dijital dönüşüm faaliyetlerinde rehberlik etmesi öngörülmüştür.
Scopus	[74]	Küçük ve orta ölçekli işletmeler			Küçük ve orta ölçekli işletmeler için önerilen model ele alınmıştır.
Scopus	[75]	Küçük ve orta ölçekli işletmeler		Düşük Olgunluk Orta Olgunluk Yüksek Olgunluk	Yarı sistematik bir literatür taraması ve yarı yapılandırılmış mülakatlar yapılarak küçük ve orta ölçekli işletmeler (KOBİ'ler) için dijital olgunluk modellerinin mevcut durumu ve modellerde yer alan yetenekler araştırılmıştır.
Scopus WoS	[76]	Sağlık Sektörü	1. Strateji, 2. BT yeteneği, 3. Birlikte çalışabilirlik, 4. Yönetişim ve yönetim, 5. Hasta merkezli bakım, 6. İnsanlar, beceriler ve davranışlar,7. Veri analitiği		
Scopus	[77]	Sektör bağımsız	1. Strateji, 2. Organizasyon, 3. Personel, 4. Teklif, 5. Teknoloji, 6. İnovasyon, 7. Çevre	Başlatılmış Yönetilen Tanımlanmış	[79] tarafından tanımlanan Dijital İnternet Olgunluk Modeli ele alınmıştır. Model, çeşitli tesislere sahip büyük bir endüstri şirketinde test edilmiştir.

				Nicel Olarak Yönetilen Optimize Edilmiş	

Turkish Cyberbullying Detection with Fine-Tuned Pre-Trained Language Models

Araştırma Makalesi/Research Article

Metin BİLGİN[1*], Bilge Nur BEKAR[1]

[1]Bursa Uludağ Üniversitesi, Mühendislik Fakültesi, Bilgisayar Mühendisliği Bölümü, Bursa, Türkiye
metinbilgin@uludag.edu.tr , bilgenbekar@gmail.com
(Geliş/Received:05.08.2024; Kabul/Accepted:04.01.2025)
DOI: 10.17671/gazibtd.1528238

***Abstract*—** With the rapid increase in internet usage and its pervasive presence in all aspects of life, social media platforms have seen a rise in negative behaviors alongside their positive contributions. One such negative behavior is cyberbullying, which refers to the misuse of information and communication technologies to harm others. Cyberbullying is becoming a significant social problem. This study aims to detect and classify Turkish sentences containing cyberbullying using deep learning models. To achieve this, the BERT model, known for its ability to understand the context of language, was chosen. Specifically, the BERTurk, DistilBERTurk, and ConvBERTurk models—designed for the Turkish language—were fine-tuned and retrained using a dataset of 3,388 tweets labeled as racist, sexist, offensive language, or neutral. The primary goal of this study is to perform a comprehensive comparison of multi-class Turkish cyberbullying detection models and to develop an Artifical Intelligence (AI) model that delivers highly accurate results on real-world data. According to the results, BERTurk achieved the highest F1 score of 0.88, while the DistilBERTurk model showed the lowest performance.

***Keywords*—** natural language processing, transformers, BERT, cyberbullying, pre-trained language models

İnce Ayar Yapılmış Ön Eğitimli Dil Modelleri ile Türkçe Siber Zorbalık Tespiti

Özet— İnternet kullanımının hızla artması ve hayatın her alanında yaygın hale gelmesiyle birlikte, sosyal medya platformlarında olumlu katkıların yanı sıra bazı olumsuz davranışlar da artış göstermiştir. Bu olumsuz davranışlardan biri, başkalarına zarar vermek amacıyla bilgi ve iletişim teknolojilerinin kötüye kullanılmasını ifade eden siber zorbalıktır. Siber zorbalık, önemli bir toplumsal sorun haline gelmektedir. Bu çalışma, derin öğrenme modelleri kullanarak siber zorbalık içeren Türkçe cümleleri tespit etmeyi ve sınıflandırmayı amaçlamaktadır. Bu amaç doğrultusunda, dilin bağlamını anlama yeteneğiyle bilinen BERT modeli tercih edilmiştir. Özellikle, Türkçe dilini destekleyen BERTurk, DistilBERTurk ve ConvBERTurk modelleri, ırkçı, cinsiyetçi, saldırgan dil veya nötr olarak etiketlenen 3.388 tweet içeren bir veri kümesiyle ince ayar yapılarak yeniden eğitilmiştir. Çalışmanın temel hedefi, çok sınıflı Türk siber zorbalığını tespit eden modellerin kapsamlı bir karşılaştırmasını yapmak ve gerçek dünya verileri üzerinde yüksek doğrulukla sonuçlar üreten bir yapay zeka modeli geliştirmektir. Sonuçlara göre, BERTurk 0,88 F1 puanı ile en yüksek başarıyı elde ederken, DistilBERTurk modeli en düşük performansı göstermiştir.

***Anahtar Kelimeler*—** doğal dil işleme, transformers, BERT, siber zorbalık, ön eğitimli dil modelleri

1. INTRODUCTION

The internet continues to permeate every aspect of our daily lives. Its widespread use directly impacts social life as well. Information technologies, with their advantages, have led to various changes in our way of life [1]. For instance, many areas such as shopping, communication, education, and entertainment are now actively conducted online. One of the areas where the internet is used most in recent times is social media platforms. Through social media, individuals can easily reach large audiences, assume desired identities, and engage in various activities. While there are advantages to this, the intensive use of the internet has also led to the proliferation of behaviors such as joy, sadness, anger, and bullying within social communication networks [2]. The misuse of information and communication technologies to harm others is referred to as cyberbullying [3]. The concept was first coined by Bill Belsey [3]. Cyberbullying refers to the reflection of harmful behaviors such as insults, ridicule, racist remarks, and judgmental attitudes on social networks. The rate of cyberbullying is increasing rapidly worldwide, and according to 2021 data, it has a rate of approximately 16% among types of bullying. [4]. The anonymity of online communication fosters a belief that actions will go unpunished, contributing to the growing prevalence of this phenomenon [5]. The platforms where cyberbullying is most frequently observed include social networking sites such as Facebook and Twitter [6]. The rapidly increasing presence of cyberbullying on the internet poses threats to societies. Therefore, detecting cyberbullying is crucial to mitigating its harmful effects. The AI-based cyberbullying detection project aims to utilize AI models to provide reliable detection of cyberbullying.

Numerous studies have been conducted on cyberbullying in recent years, with many employing traditional machine learning methods [7][8][9]. These studies have explored various aspects of cyberbullying detection, yet they often fall short in terms of addressing the complexities introduced by nuanced language and the diversity of cyberbullying behaviors. However, the growing use of deep learning models in natural language processing (NLP) can be attributed to advances in hardware, the increasing volume of available data, and the rise of open-source projects. Recent innovations in deep learning, particularly the development of transformative architectures, have significantly enhanced NLP by improving both the speed and accuracy of word context understanding. Despite these advancements, a universally effective tool for detecting cyberbullying has yet to be developed. As a result, further research in this area is essential.

A review of the existing literature reveals a lack of Turkish-specific data related to cyberbullying, with most available datasets focusing on binary classification. Our study seeks to fill this gap by applying multi-class classification techniques using widely-used, high-performance pre-trained models such as BERT, DistilBERT, and ConvBERT. We evaluate the performance of these models by calculating precision, recall, confusion matrix, PR curves, and F1 score metrics. Additionally, user interaction is incorporated into this process. The system we developed generates appropriate labels and scores to indicate whether the input text contains instances of cyberbullying.

The objectives of this study are as follows:

- To achieve a comprehensive study by performing multi-class classification for cyberbullying detection using a multi-class dataset.
- To fine-tune and train deep learning-based pre-trained models and compare the performance of the developed models based on evaluation metrics.
- To contribute to research in Turkish natural language processing, particularly in text classification and cyberbullying detection. Turkish-supported BERT models, which have been insufficiently explored in the literature, using a multi-class dataset.
- To enhance cyberbullying awareness through the evaluation of data containing cyberbullying.
- To develop a system that provides users with an easy way to detect Turkish cyberbullying.

2. LITERATURE REVIEW

The study focuses on a multi-class text classification problem, with an emphasis on Turkish research and recent developments in the field. The methods discussed in the literature review are categorized under relevant headings and critically analyzed in the context of this study. There are many machine learning methods in the literature. In addition, studies involving deep learning methods are also included in the literature review. In contrast, our research contributes to the existing literature by using transformer-based architectures, known for their high performance in natural language processing tasks.

In this study, the methods are presented by incorporating changes based on the accumulated knowledge obtained from the studies observed in the literature. A dataset containing neutral, offensive language, sexism, and racism was selected to comprehensively address the detection of cyberbullying in Turkish. A multi-class Turkish dataset was used to implement BERT-based models, and BERTurk, ConvBERTurk, and DistilBERTurk models were applied to develop a Turkish cyberbullying detection system. In this study, for the first time in the literature, BERTurk, ConvBERTurk, and DistilBERTurk models were evaluated using the Turkish multi-class cyberbullying dataset, using the F1 score, recall, precision, confusion matrix, and PR curve.

2.1 Machine Learning Approaches in Existing Literature

Sevli and Sezgin used machine learning methods to detect and categorize cyberbullying in social media posts. The study used a dataset consisting of 47,692 English tweets. The dataset was balanced and contained six classes: bullying based on belief, gender, age, ethnicity, other characteristics, and non-cyberbullying. Confusion matrix,

F1 score, precision, and recall measures were used to compare KNN, SVM, and Random Forest algorithms. The best result was obtained with the SVM model, which achieved 83% accuracy. The best-performing category was non-cyberbullying tweets. It was suggested that the relatively smaller size of this class (16%) and the presence of fewer meaningful expressions could explain this result. The study suggested the use of deep learning techniques to address the problem more effectively [7]. This study was examined because it classified cyberbullying in multi-class, and the dataset contained similar categories. Although the balanced dataset provides a high accuracy rate with machine learning methods, the fact that the dataset is in English limits its applicability in detecting multi-class cyberbullying in Turkish, which has not yet been widely addressed in the literature.

In their study, Rohini and Ramchander focused on detecting cyberbullying in digital forums using machine learning methods. Two datasets were employed. The first dataset contains 10,000 comments, with 80% of the data not involving cyberbullying and 20% involving it. The second dataset includes 20,000 comments, where 60% of the data does not contain cyberbullying and 40% does, making it an imbalanced dataset. The best results were achieved using the Random Forest approach, with an accuracy of 99% [8]. Despite the imbalance, high performance was observed across both datasets. The study demonstrates high effectiveness in detecting cyberbullying, even with the imbalanced nature of the dataset.

Bozyiğit et al.'s study emphasizes the role of social media features in cyberbullying detection, creating a balanced dataset of 5,000 labeled posts and using the chi-square test to explore the relationship between features like the sender's follower count and cyberbullying. They tested machine learning algorithms on two datasets: one with only text features and the other with social media features. The latter showed better performance [9]. This binary classification study contributes to Turkish cyberbullying detection but differs methodologically from ours, as we use BERT-based deep learning models, multi-class classification, and a broader set of performance metrics. Both studies advance the field, yet our research highlights the evolution from traditional machine learning to deep learning and provides deeper insights into different types of cyberbullying.

In this study, Çöltekin developed a dataset to automatically detect offensive language in Turkish social media posts and conducted experiments using various machine learning methods. His primary focus was the classification of offensive language. The dataset was derived from Turkish tweets posted between 2018 and 2019, consisting of 36,232 manually labeled tweets. Each tweet was classified as either offensive or non-offensive, with a third annotator making the final decision in cases of disagreement. Çöltekin also evaluated the dataset in terms of cyberbullying analysis across different regions of Turkey and assessed the general rate of tweets containing cyberbullying, offering valuable insights. He employed SVM, n-gram, and BM25 models to automatically classify offensive tweets. The study achieved an F1 score of 77.3% for identifying offensive tweets, 77.9% for determining whether a specific tweet was targeted, and 53.0% for classifying targeted offensive tweets into three subcategories [10]. The study also addressed the multi-label classification problem. In contrast, our study tackles the multi-class classification problem with greater success.

2.2 Hybrid and Advanced Approaches

Sel and Hanbay explored various algorithms, including machine learning-based TFIDF+SVM, deep learning-based CNN and LSTM, and pre-trained Turkish language models like BERT, DistilBERT, and Electra. Their study focused on binary text classification using a partially balanced dataset of 5,292 tweets, evaluating models based on accuracy, F1 score, specificity, and sensitivity. BERT achieved the highest accuracy at 80%. In gender determination, the SVM classifier was compared to pre-trained language models. Although SVM requires more parameters and doesn't account for word meanings, it performed similarly to pre-trained models when contextual understanding was less important. While the study is valuable for exploring different models, further research is needed to achieve higher accuracy with balanced and larger datasets [11].

Nergiz and Avaroğlu trained an LSTM neural network with three different word embedding models using a dataset of 180,000 comments collected from three different social media platforms and set the epoch value to 10. The model using the Fasttext method achieved the highest accuracy of 93%. It was evaluated that the factors contributing to this high accuracy were the use of balanced data for binary classification and the implementation of data preprocessing steps. The limited increase in success was attributed to the non-standardized structure of social media spelling rules and the insufficient content of the comments coming from the Instagram platform [12]. The study stands out in terms of evaluating the adequacy of social media data. The dataset we used in the study shed light on the interpretation of the adequacy of social media data.

2.3 BERT and Transformer Model Approaches

Karaman used the BERT model for the binary classification of discriminatory-exclusionary tweets against Syrian refugees in his study. The study utilized the pre-trained BERT-BASE-TURKISH-UNCASED model, trained on Turkish documents. The dataset consisted of 2,264 tweets, and the model was trained for 12 epochs, achieving an accuracy of 0.8562 during training and 0.81 on the test set. It was suggested that increasing the sample size could enhance the model's sensitivity and accuracy [13]. While the study is similar to ours in using the BERT model for cyberbullying detection, there are key differences, such as the use of different models and a focus on binary classification.

In this study, Beyhan et al. conducted experiments on both binary and multi-class classification problems using the Istanbul Convention and Refugees datasets. They retrained the model with a 5-class dataset representing hate speech, aiming to accurately classify data on topics like hate speech and cyberbullying using various text classification techniques. The BERTurk model was employed, and results were evaluated using 5-fold cross-validation. On the Istanbul Convention dataset, the binary classification achieved an average accuracy of 77.06%, and multi-class classification reached 72.22%. The F1 scores were 77.86% and 72.22%, respectively. However, challenges arose as the BERT model, trained on formal sources, struggled with informal and short texts like those from Twitter. Additionally, the unbalanced nature of the Istanbul Convention dataset impacted classification results [14]. This study offers a comprehensive approach to detecting both multi-class and binary Turkish cyberbullying. In contrast, our research addresses the gap by comparing the performance of different BERT-based models, achieving a higher F1 score using a different dataset. Furthermore, while this study focused on tweets related to specific events (such as the Istanbul Convention or refugee issues), our study was trained on tweets from a dynamic, independent cyber environment.

In their study, Çelikten and Bulut developed a model using the BERT model to classify ten diseases in a dataset consisting of Turkish medical texts. BERTurk models were preferred due to the fact that Turkish is an agglutinative language and the morphological difficulties it presents in natural language processing. In addition, a multilingual BERT model developed by Google was used. The evaluation metrics of the study included precision, recall, F1 score, and weighted and macro averages. It was seen that the BERTurk model outperformed the multilingual BERT model [15]. This study shows that the BERTurk model stands out in terms of comparing BERT models supporting Turkish. The study provides resources for examining the BERTurk model with evaluation metrics and examining its performance. Aytan and Sakar conduct a comparative analysis of transformer-based models for Turkish natural language processing problems in their study. The study examined BERT, ConvBERT, and Electra models on sentiment analysis, named entity recognition (NER), and text classification problems using pre-trained Turkish models, while the RoBERTa model was trained with a large Turkish corpus. For text classification with a seven-category dataset, the BERTurk model achieved the highest classification performance with an accuracy of 94%. The ConvBERT model achieved a performance of 93.9% [16]. The study provides us with resources to examine the use of BERTurk and ConvBERTurk models in text classification problems.

In their study, Özkan and Görkem apply the BERT model to a multi-class classification problem. They preprocess the dataset with steps such as tokenization, case transformation, removal of stop words, deletion of numbers, and stemming, and then apply two different training-validation ratios to the dataset. The first ratio is 80% training and 20% testing, while the second is 85% training and 15% testing. The AdamW optimization method was selected as the optimizer. It was observed that AdamW showed less overfitting compared to models trained with the Adam optimization algorithm. The model with an 85% training ratio achieved an F1 score of 96%. The 85% training ratio gave better results compared to the 80% training ratio [17]. The study guides the use of AdamW optimization and the BERT model in a multi-class classification problem.

Arzu and Aydoğan performed a comparative analysis of BERTurk models for Turkish sentiment classification. In their study, they achieved the highest accuracy of 83% using the preprocessed ConvBERTurk mc4 (without case) model with a balanced dataset of 150,000 samples. The lowest performance was observed in the DistilBERTurk cased model [18]. The models used in their study were also applied to cyberbullying detection in our research. The study provided insight into the BERTurk, ConvBERTurk, and DistilBERTurk models.

Table 1. Literatıre Review Summary

Study	**Methods/Approaches**	**Classification Type**	**Dataset Language**	**Topic**	**Best Performance**
Sevli & Sezgin [7]	SVM, KNN, Random Forest	Multi-class	English	Cyberbullying detection (6 categories, balanced)	SVM: 83% accuracy
Rohini & Ramchander [8]	Random Forest	Binary	English	Cyberbullying detection (imbalanced 2 datasets)	Random Forest: 99% Accuracy
Bozyiğit et al. [9]	Traditional ML	Binary	Turkish	Cyberbullying detection with social media features	ML with social features: High correlation
Çöltekin [10]	SVM, BM25	Binary	Turkish	Offensive language detection (multi-label)	SVM: 77.9% F1 score (targeted tweets)
Sel & Hanbay [11]	TFIDF+SVM, CNN, LSTM, BERT, DistilBERT, Electra	Binary	Turkish	Text (gender) classification with pre-trained models	BERT: 80% accuracy

Nergiz & Avaroğlu [12]	LSTM with Fasttext	Binary	Turkish	Classification of Cyberbullying in Social Media Comments	Fasttext+LSTM: 93% accuracy
Karaman [13]	BERTurk (BASE-TURKISH-UNCASED)	Binary	Turkish	Discriminatory tweets about Syrian refugees	BERTurk: 81% test accuracy
Beyhan et al. [14]	BERTurk	Multi-class/Binary	Turkish	Hate speech and cyberbullying detection	BERTurk: 77.06% accuracy (binary)
Çelikten & Bulut [15]	BERTurk, Multilingual BERT	Multi-class	Turkish	Medical text classification (10 diseases)	BERTurk: Outperformed Multilingual BERT
Aytan & Sakar [16]	BERTurk, ConvBERT, Electra	Multi-class	Turkish	Sentiment analysis and text classification	BERTurk: 94% accuracy
Özkan & Görkem [17]	BERT	Multi-class	Turkish	Multi-class classification with optimization	BERT+AdamW: 96% F1 score
Arzu & Aydoğan [18]	ConvBERTurk mc4, DistilBERTurk	Binary	Turkish	Sentiment classification with balanced data	ConvBERTurk mc4: 83% accuracy

Upon examining the existing literature, it is evident that various approaches have been employed in the field of cyberbullying detection. An evolution from machine learning methods to deep learning techniques, particularly transformer-based models, can be observed. However, significant gaps still exist in the area of Turkish cyberbullying detection.

Firstly, studies on multi-class cyberbullying detection in the Turkish language are limited. Most existing research has focused on binary classification problems. In this context, our study aims to fill this gap in the literature by adopting a multi-class approach (neutral, offensive language, sexism, and racism).

Secondly, there is a lack of comparative analysis of the performance of BERT models specifically developed for the Turkish language in cyberbullying detection. Our study evaluates the effectiveness of BERTurk, ConvBERTurk, and DistilBERTurk models in Turkish cyberbullying detection by comparing them on the same dataset. This comparison allows us to determine which model is most suitable for this task, providing a valuable reference point for future research.

Thirdly, most existing studies have evaluated the performance of the models used with limited metrics. Our study employs a comprehensive set of metrics, including F1 score, recall, precision, confusion matrix, and PR curve, to evaluate the models' performance from multiple angles. This approach enables a more in-depth understanding of the strengths and weaknesses of the models.

Finally, the methodological innovation of our study lies in its systematic comparison of different BERT-based models' performance in Turkish cyberbullying detection. This comparison not only identifies the best-performing model but also reveals the success of each model in detecting different types of cyberbullying. This detailed analysis can guide future studies and assist researchers in selecting the most appropriate model for cyberbullying detection in Turkish natural language processing.

In conclusion, our study aims to fill existing gaps in the field of Turkish cyberbullying detection, comparatively evaluate the performance of the most up-to-date BERT-based models, and improve methodological approaches in this area. The findings of this study will contribute to the advancement of research in the field of Turkish cyberbullying detection by providing a solid foundation for future investigations.

3. MATERIAL AND METHODS

This section will provide information about the materials and methods used in the study.

The stages of the study are outlined below:

- Three distinct language models were trained on the same dataset to classify sentences containing racism, sexism, offensive language, and neutral expressions related to cyberbullying.
- The trained models and tokenizer files were shared on the HuggingFace platform, providing resources and opportunities for users to test the model for Turkish cyberbullying detection.
- To ensure rapid and reliable interaction of the cyberbullying detection with software, an API was developed using FastAPI.
- A user-friendly and lightweight interface was developed to facilitate easy cyberbullying detection.
- This study contributes to the field of Turkish natural language processing, which has limited research, by providing reliable and high-accuracy results in cyberbullying detection using deep learning-based models.

The flow diagram of the study can be seen in Figure 1.

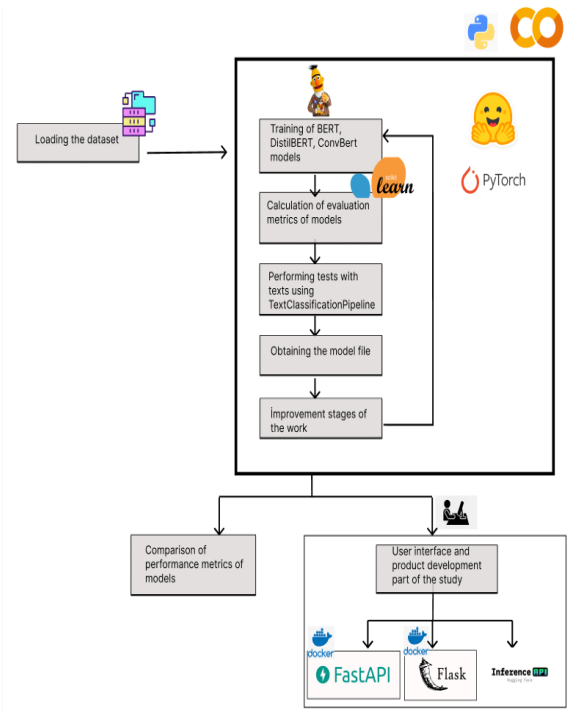


Figure 1. Flow diagram of the study

3.1. Dataset

Finding a labeled, multi-class dataset for Turkish cyberbullying is quite challenging. In this study, the Turkish-social-media-offensive-bullying dataset [19] provided by the HuggingFace platform was chosen, as it is the only available dataset. This dataset consists of 3,388 pre-processed samples, labeled for cyberbullying detection across four classes: racism, sexism, offensive language, and neutral. These classes cover significant aspects of cyberbullying detection and represent the only multi-class labeled dataset available in this domain. To address the imbalance in the dataset, data reduction techniques were applied to the most populated class, Neutral. The dataset contains 490 samples for Racism, 601 for Sexism, 910 for Offensive Language, and 980 for Neutral.

The initial and final distribution graphs of the dataset are shown in Figure 2.

a. b.

Figure 2. Distribution Graph of Classes

a. Distribution graph of the old dataset b. Distribution graph of the new dataset

Word clouds of the classes of the data set are shown in Figure 3.

a. b.

c. d.

Figure 3. Word clouds for each class: a. Word cloud for the Sexism class b. Word cloud for the Racism class c. Word cloud for the Offensive language d. Word cloud for the Neutral class.

One of the main reasons for choosing this dataset is that it contains four key classes that help detect cyberbullying. It is also the only Turkish open-source dataset with these categories. The Offensive Turkish Dataset [20], another open-source, multi-class cyberbullying dataset, provides an ideal resource for studies focused on multiple classifications and labeling. Cyberbullying data can be categorized into several types: non-aggressive language with swearing or untargeted attacks, attacks against a group, individual attacks on a person, and attacks targeting a non-human entity, such as an event or organization. Additionally, data can belong to more than one class. For these reasons, we decided to proceed with the Turkish-social-media-offensive-bullying dataset for our study.

3.2. Pre-Trained Models

This section provides information about the models used in the study.

1) BERT: Bidirectional Encoder Representations from Transformers (BERT) is a pre-trained language model based on the evolving transformer architecture. The BERT model generates numerical representations by evaluating context-appropriate words. Specifically, the numerical vectors of words with similar meanings are closely aligned, whereas vectors for different meanings of the same word are less similar. BERT's architecture simultaneously considers both left and right contexts, which contributes to its effectiveness in interpreting word meanings even in complex language processing tasks [21]. The BERT model is composed of a 12-layer transformer structure [22]. The architecture of the BERT model is presented in Figure 4.

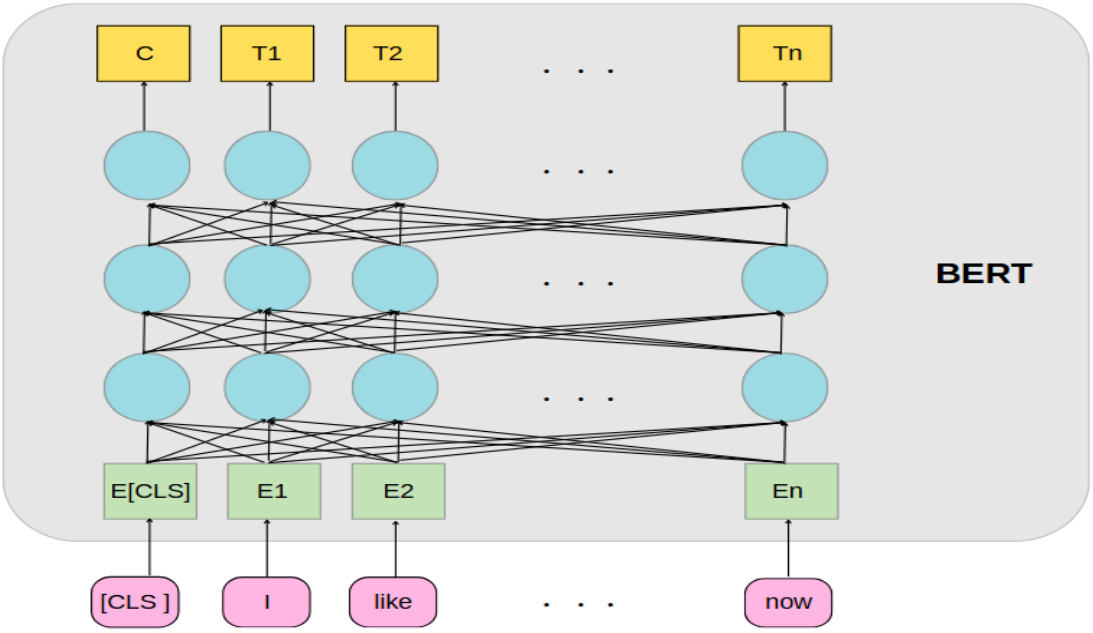


Figure 4. BERT model structure [21]

2) ConvBERT: The ConvBERT model is developed by replacing the self-attention layer of the BERT architecture with a span-based dynamic convolution. While the self-attention layer in the BERT model requires querying all inputs, the span-based self-attention allows for the examination of local dependencies. This modification aims to reduce the high memory usage and computational cost associated with the self-attention layer in BERT, thereby achieving better performance [16]. The ConvBERT model maintains the general structure of BERT while enhancing the attention mechanism by introducing convolution, thus improving the model.

ConvBert avoids the bottleneck of the BERT model and uses mixed attention. It is important to emphasize that ConvBERT uses both self-attention and convolution mechanisms. This shows that combining the two approaches yields better results [23]. The approaches presented in Figure 5 are discussed.

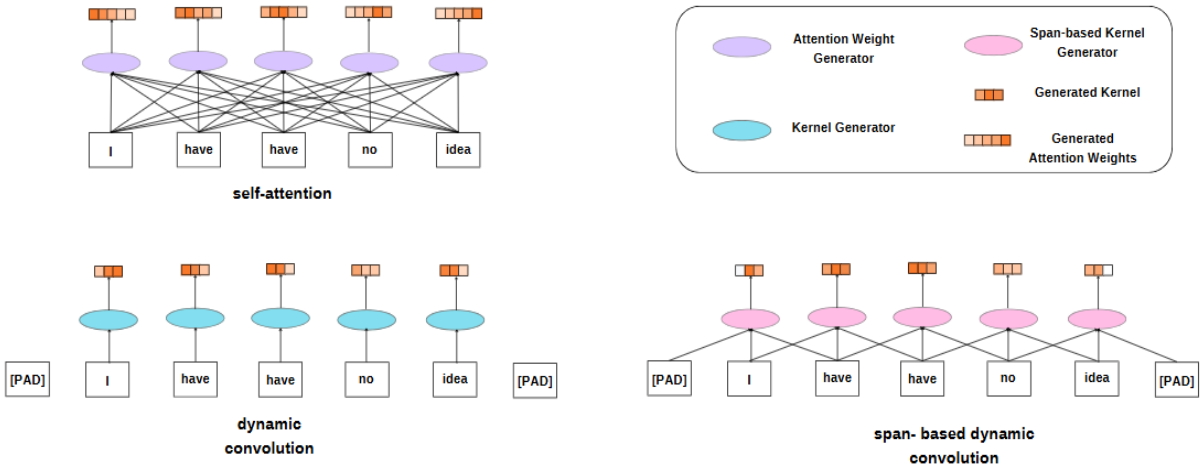


Figure 5. Approaches of self-attention, dynamic convolution, and span-based dynamic convolution [24]

3) DistilBERT: DistilBERT, introduced in a 2019 paper [24] and developed by Hugging Face [25], is a distilled version of larger models like BERT, designed to address the challenge of limited computational resources. The model claims a 40% reduction in size while retaining 97% of the linguistic understanding of BERT and increasing processing speed by 60% [24]. Its compact nature makes it especially well-suited for deployment on mobile devices and real-time applications. DistilBERT has fewer transformer layers compared to BERT and, in some studies, has been observed to have lower accuracy than BERT [26].

4) BERTurk: The pre-trained transformer models used in this study were pre-trained on Turkish data using a specialized corpus provided by Kemal Oflazer, a filtered version of the Turkish OSCAR corpus Wikipedia dumps, and various OPUS corpora [27]. The models employed in the study include bert-base-turkish-uncased, distilbert-base-turkish-cased, and convbert-base-turkish-mc4-cased, all trained with the specified datasets.

3.3. Fine-Tuning Training Phase

These models were trained for Turkish cyberbullying detection in a Google Colab environment, using a Tesla T4 GPU with CUDA. The GPUs enabled parallel computation, which allowed the models to be fine-tuned more quickly and with better memory management. The Turkish-social-media-offensive-bullying dataset was split into 80% for training and 20% for testing. To address the dataset imbalance, where the "Neutral" class was the most frequent with 1387 samples, the number of samples was reduced to 980 for better results. The distribution of word counts, including the minimum, maximum, and average counts, was also examined through a box plot (Figure 4). Based on this analysis, a maximum length (max_len) of 100 was selected to accommodate subword tokenization in the models.

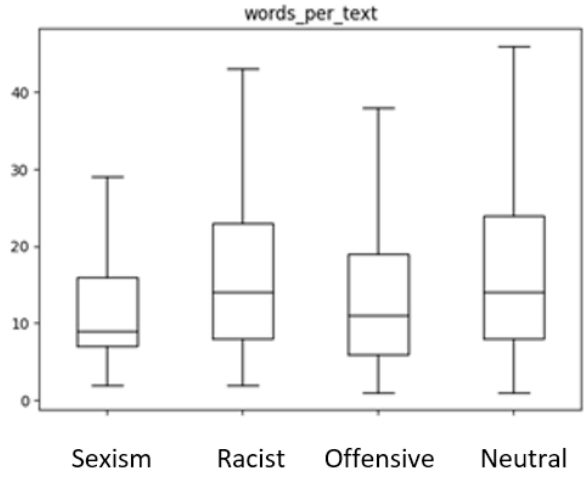


Figure 6. Box plot of word counts by class

The BERTurk model was trained using the pre-trained bert-base-turkish-uncased model and tokenizer through the Transformers library with 7 epochs and a batch size of 16. The AdamW algorithm was employed for optimization. The created model was trained using the initial weights of the BERTurk model. For training the DistilBERTurk model, 10 epochs and a batch size of 16 were used. The ConvBERTurk model was trained with 9 epochs and a batch size of 16. Although the training duration of the ConvBERTurk model is similar to that of the BERTurk model, it is somewhat shorter.

3.4. Evaluation Metrics

Precision: Precision, also known as sensitivity, measures how many of the positive predictions are correctly identified as true positives.

$$\text{Precision} = TP/(TP+ FP) \quad (1)$$

Recall: Recall, also known as sensitivity, measures the proportion of actual positive instances that are correctly identified as true positives in classification.

$$\text{Recall} = TP/(TP+FN) \quad (2)$$

F1 Score: The F1 score is used to measure classification performance. It is computed as the harmonic mean of precision and recall values. The F1 score is particularly useful in cases where the distribution is imbalanced.

$$F1 = 2 \times ((PxR)/(P+R)) \quad (3)$$

Confusion Matrix: The confusion matrix provides a summary of a classification problem by comparing the true labels with the predicted labels [26].

Precision-Recall Curves (PR Curves): It allows for the examination of the performance of precision and recall values in models with class imbalance

4. APPLICATION AND RESULTS

In the study developed for the detection of cyberbullying in Turkish, the dataset was divided into training and test data in equal proportions (80%-20%). The results obtained from the BERTurk, DistilBERTurk, and ConvBERTurk models were compared. The confusion matrices regarding the class prediction performance in the test dataset are presented in Figures 7-9. When the confusion matrices were examined, it was revealed that the models showed more errors in predicting the Offensive Language and Neutral classes. The ConvBERTurk and BERTurk models produce similar results. The results of the models are given in detail in Tables 2-4. The study achieved high success in the detection of cyberbullying in Turkish with the multi-class transformative model by obtaining an F1 score of 0.88. When the research results and project development process were evaluated, it was seen that reducing the Neutral category reduced the bias in the model outputs and improved the generalization ability. 2385 training samples and 596 test samples were used for model training. Among

the BERTurk, DistilBERTurk, and ConvBERTurk models, the BERTurk model achieved the highest F1 score of 0.884, while the DistilBERTurk model achieved the lowest F1 score of 0.83. PR curves are given in Figure 10 to examine the effects of class imbalance on the model and its performance. The graphs show that the Sexism class achieved the highest performance in all models with balanced precision and recall values. High locality and steep curves show good performance in accurate predictions and identifying positive examples. The BERTurk model is the most balanced in terms of class distribution, followed by the ConvBERTurk model. Comparing the balanced structure of the PR curve is useful in comparing the success of the models.

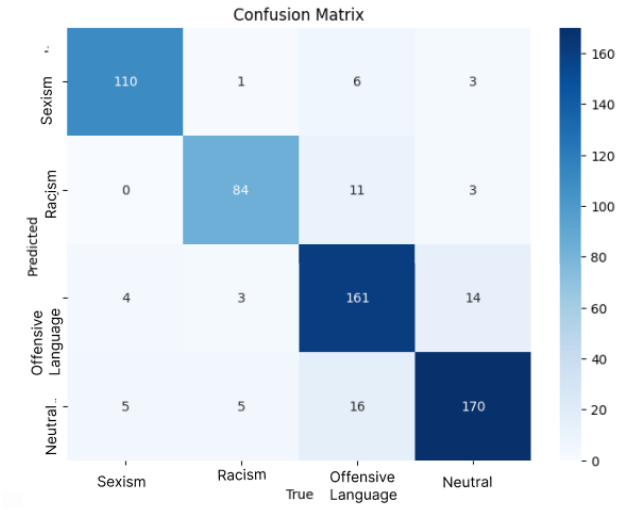


Figure 7. BERTurk confusion matrix

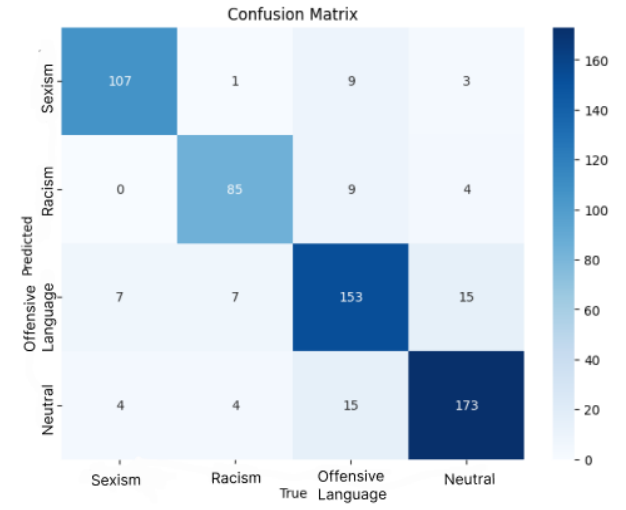


Figure 8. ConvBERTurk confusion matrix

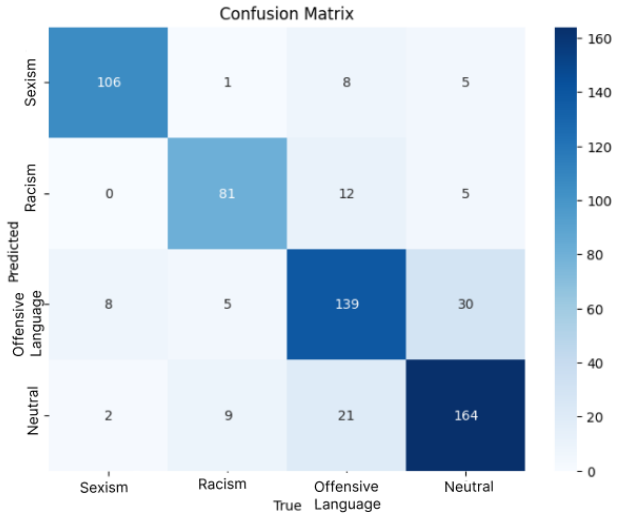


Figure 9. DistilBERTurk confusion matrix

Table 2. BERTurk model evaluation results

Classes	Precision	Recall	F1 score
All Classes	0.888	0.881	**0.884**
Sexism	0.924	0.917	**0.920**
Racist	0.903	0.857	0.880
Offensive	0.830	0.885	0.856
Neutral	0.895	0.867	0.881

Table 3. ConvBERTurk model evaluation results

Classes	Precision	Recall	F1 score
All Classes	0.873	0.871	0.872
Sexism	0.907	0.892	**0.899**
Racist	0.876	0.867	0.872
Offensive	0.823	0.841	0.831
Neutral	0.887	0.883	0.885

Table 4. DistilBERTurk model evaluation results

Classes	Precision	Recall	F1 score
All Classes	0.833	0.828	0.830
Sexism	0.914	0.833	**0.898**
Racist	0.844	0.826	0.835
Offensive	0.772	0.764	0.768
Neutral	0.804	0.837	0.820

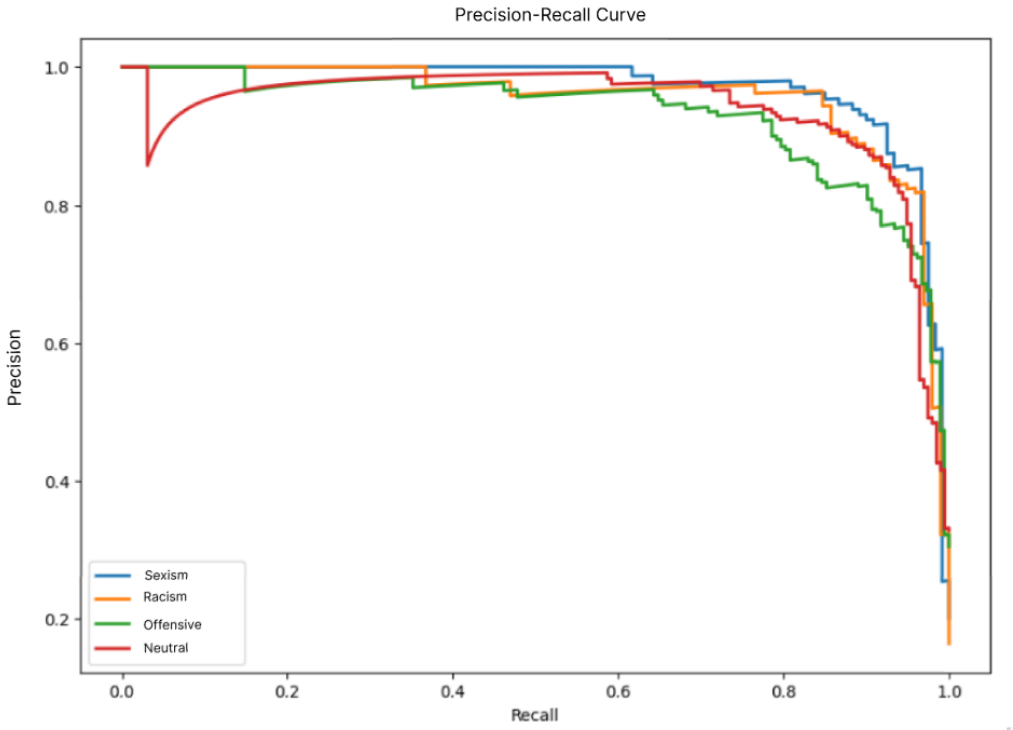


Figure 10. BERTurk PR curves

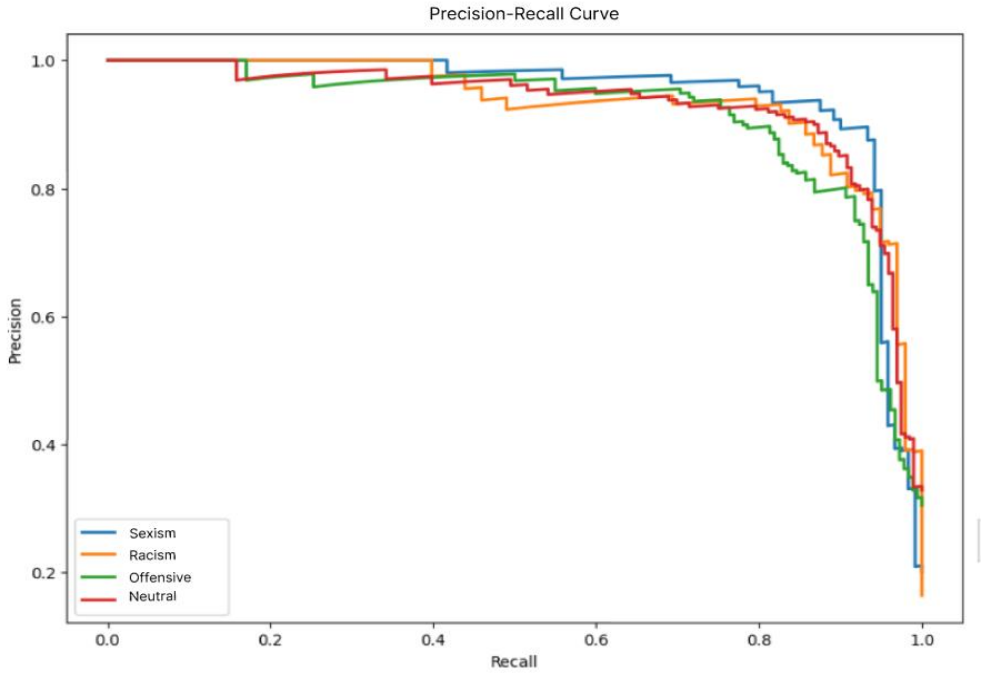


Figure 11. ConvBERTurk PR curves

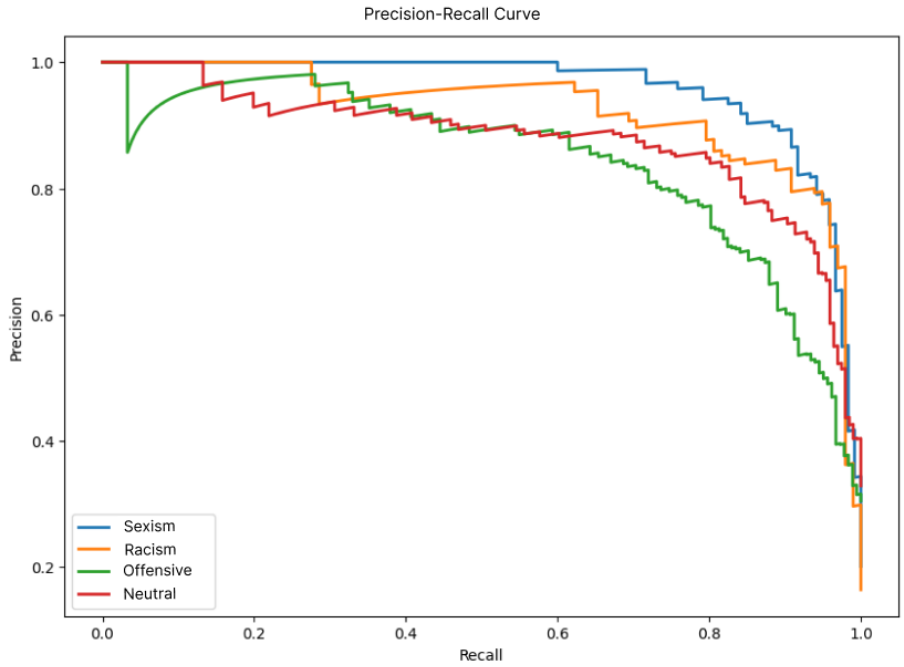


Figure 12. DistilBERTurk PR curves

The BERTurk model achieves high performance; however, the ConvBERTurk model provides similar results while offering advantages such as a shorter training duration and a smaller model file size.

5. CONCLUSIONS

Cyberbullying poses a significant threat to both the present and future world. To address this issue, raising awareness and automating processes are essential. Recent advancements in natural language processing techniques enable the categorization of text in a language by learning contextual information from large corpora. Turkish natural language processing is still developing in this field, and there is a need for using large language models in detecting cyberbullying. The developed cyberbullying detection system can be adapted for individual use, social media environments, internal corporate communication platforms, and various other settings. This contributes to raising awareness on the issue and advancing Turkish natural language processing research.

In the conducted study, several challenges were encountered. The first challenge is the scarcity of multi-class and labeled datasets in the domain of cyberbullying. The second challenge pertains to the interpretability of the language. For instance, during the testing phase, the sentence "sen de ben de ne dediğimizi bilmiyoruz" (both you and I do not know what we are saying) could be classified as either Neutral or Offensive Language by different models. Both labels could be considered correct for this sentence. Another issue is that the labeled data for the Offensive language class might be insufficiently representative of the general scope. The third issue is that in the Turkish-social-media-offensive-bullying dataset, sentences containing racial terms classified as Racism can also produce Racism outputs in contexts where they might be considered Neutral. The interpretability of the language complicates the detection of cyberbullying. Higher accuracy could be achieved with a more comprehensive and less biased Turkish dataset. The results indicate that BERT emerged as the most successful model. BERT performs well in producing strong and accurate results but operates more slowly. ConvBERT, on the other hand, provides a better balance between power and speed compared to BERT. The DistilBERT model, being a more distilled version, achieved lower accuracy compared to other models. It may be preferred in scenarios where speed and memory management are more critical. The ConvBERT model, with its balanced architecture, is suitable for both cases. Additionally, to enable users to experiment with the trained models and obtain results, as well as to analyze the models with current data, a Flask-based, user-friendly, portable, and platform-independent application was created. This API, easily integrated with software, was developed using FastAPI. Furthermore, the models can be accessed for development and testing through the Hugging Face platform [28]. The developed system provides support for those interested in model development and usage. With further development, this work could offer effective and realistic solutions for real-time detection of cyberbullying.

REFERENCES

[1] O. Zorbaz, **"Lise Öğrencilerinin Problemli İnternet Kullanımının Sosyal Kaygı ve Akran İlişkileri Açısından İncelenmesi."** Yüksek lisans tezi, Hacettepe Üniversitesi, Sosyal Bilimler Enstitüsü, Ankara, 2013.

[2] F. Gültekin, **"Saldırganlık ve Öfkeyi Azaltma Programının İlköğretim İkinci Kademe Öğrencilerinin Saldırganlık ve Öfke Düzeyleri Üzerindeki Etkisi"**, Doktora Tezi, Hacettepe Üniversitesi, 2008

[3] M. Tuncer, M. Dikmen, **"Sosyal Ağlarda Bekleyen Yeni Tehlike: Siber Zorbalık"**, 4. International Instructional Technologies and Teacher Education Symposium, 94-104, 2016.

[4] İ. Yıldırım, "Sosyal Medya, Dijital Bağımlılık ve Siber Zorbalık Ekseninde Değişen Aile İlişkileri Üzerine Bir Değerlendirme" . *Anemon Muş Alparslan Üniversitesi Sosyal Bilimler Dergisi*, 9.5: 1237-1258, 2021.

[5] E. V. Altay, B. Alataş, **"Detection of Cyberbullying in Social Networks Using Machine Learning Methods"** International Congress on Big Data, Deep Learning and Fighting Cyber Terrorism (IBIGDELFT). IEEE, p. 87-91, 3-4 Dec. 2018.

[6] V. Balakrishnan, S. Khan, H. R. Arabnia, "Improving Cyberbullying Detection Using Twitter Users' Psychological Features and Machine Learning*.", Computers & Security* 90, 101710, 2020.

[7] O. Sevli, & S. Sezgin, "Sosyal Medya Paylaşımlarında Siber Zorbalığın Tespiti ve Kategorizasyonuna Yönelik Makine Öğrenmesine Dayalı Bir Sınıflandırma". **Bursa 3rd International Scientific Research Congress,** Bursa, 626-637, 2022.

[8] D. S. Rohini, M. Ramchander, "A Comparative Study of Machine Learning Approaches for Cyberbullying Detection in Digital Forums", **International Conference on Advances in Computation, Communication and Information Technology** (ICAICCIT) (pp. 332-338). IEEE, 23-24 Nov. 2023.

[9] A. Bozyiğit, S. Utku, E. Nasibov, "Cyberbullying Detection: Utilizing Social Media Features", *Expert Systems with Applications*, 179, 115001, 2021.

[10] Ç. Çöltekin, "A Corpus of Turkish Offensive Language on Social Media**." In Proceedings of the Twelfth Language Resources and Evaluation Conference** (pp. 6174-6184). Marseille, 11–16 May 2020

[11] İ. Sel, İlhami, D. Hanbay. "Ön Eğitimli Dil Modelleri Kullanarak Türkçe Tweetlerden Cinsiyet Tespiti" *Fırat Üniversitesi Mühendislik Bilimleri Dergisi,* 33.2: 675-684, 2021.

[12] G. Nergiz, E. Avaroğlu. "Türkçe Sosyal Medya Yorumlarındaki Siber Zorbalığın Derin Öğrenme ile Tespiti." *Avrupa Bilim ve Teknoloji Dergisi* 31 :77-84, 2021.

[13] E. Karaman, "Suriyeli Mültecilere Uygulanan Ayrımcı Dışlayıcı Twitlerin BERT Modeli ile Sınıflandırılması". *Ortadoğu Ve Göç*, 12(2), 428-456, 2022.

[14] F. Beyhan, B. Çarık, I. Arın, A. Terzioğlu, B. Yanıkoğlu, & R. A. Yeniterzi, Turkish Hate Speech Dataset and Detection System. **In Proceedings of the Thirteenth Language Resources and Evaluation Conference** (pp. 4177-4185). (2022, June).

[15] A. Çelıkten, H. Bulut "Turkish Medical Text Classification Using Bert." , **29th Signal Processing and Communications Applications Conference** (SIU). IEEE, 9-11 June 2021.

[16] B. Aytan, C. O. Sakar. "Comparison of Transformer-based Models Trained in Turkish and Different Languages on Turkish Natural Language Processing Problems." **30th Signal Processing and Communications Applications Conference** (SIU). IEEE, 15-18 May 2022.

[17] M. Özkan, G. Kar, "Türkçe Dilinde Yazılan Bilimsel Metinlerin Derin Öğrenme Tekniği Uygulanarak Çoklu Sınıflandırılması*". Mühendislik Bilimleri ve Tasarım Dergisi*, 10.2: 504-519, 2022.

[18] M. Arzu, M. Aydoğan, "Türkçe Duygu Sınıflandırma İçin Transformers Tabanlı Mimarilerin Karşılaştırılmalı Analizi", *Computer Science*, (IDAP-2023), 1-6, 2023

[19] Internet: Nanelimon, Huggingface Datasets, https://huggingface.co/datasets/nanelimon/turkish-social-media-offensive-dataset, 1.03.2024.

[20] Internet: A Corpus of Turkish Offensive Language, https://coltekin.github.io/offensive-turkish, 16.10.2024.

[21] J. Devlin, M. W. Chang, K. Lee, K. Toutanova, "Bert: Pre-training of Deep Bidirectional Transformers for Language Understanding" , arXiv preprint arXiv:1810.04805, 2018.

[22] S. K. Behera, R. Dash, "A Novel Feature Selection Technique for Enhancing the Performance of Unbalanced Text Classification Problem". *Intelligent Decision Technologies*, 16(1), 51-69, 2022.

[23] Z. Jiang, W. Yu, D. Zhou, Y. Chen, J. Feng, S. Yan, "Convbert: Improving Bert with Span-Based Dynamic Convolution." Advances in Neural Information Processing Systems, 33: 12837-12848, 2020.

[24] V. Sanh, L. Debut, J. Chaumond, T. Wolf, "DistilBERT, a Distilled Version of BERT: Smaller, Faster, Cheaper and Lighter", arXiv preprint arXiv:1910.01108, 2019.

[25] T. Wolf, L. Debut, V. Sanh, J.Chaumond, C. Delangue, A. Moi, P. Cistac, T. Rault, R. Louf, M. Funtowicz, J. Davison, S. Shleifer, P. V. Platen, C. Ma, Y.Jernite, Julien Plu, C. Xu, T. L. Scao, S. Gugger, M.Drame, Q. Lhoest, A., "Rush, Transformers: State-of-the-art Natural Language Processing"., **Conference on Empirical Methods in Natural Language Processing: System Demonstrations** (pp. 38-45), October 2020.

[26] M. Turan, "**Derin Öğrenme ile Beklenti Tabanlı Duygu Analizi**", Yüksek Lisans Tezi, Bursa Uludağ Üniversitesi, Fen Bilimleri Enstitüsü, 2022.

[27] H. A. Ardaç, P. Erdoğmuş, "Question-Answering System with Text Mining and Deep Networks". Evolving Systems, 1-13, 2024.

[28] İnternet: B. N. Bekar, HuggingFace, https://huggingface.co/AIZinu, 21.7.2024.

Tinc VPN ile Güçlendirilmiş Merkeziyetsiz Dosya Depolama ve Paylaşım Uygulaması

Araştırma Makalesi/Research Article

Alihan ÖZEN[1], Derya YILTAŞ KAPLAN[1]*

[1]Bilgisayar Mühendisliği Bölümü, İstanbul Üniversitesi-Cerrahpaşa, İstanbul, Türkiye.
alihanozen76@gmail.com, dyiltas@iuc.edu.tr
(Geliş/Received:25.10.2024; Kabul/Accepted:28.01.2025)
DOI: 10.17671/gazibtd.1573426

Özet— Kuruluşlar ve özellikle şirketler için veri güvenliği kritik bir öneme sahiptir. Verilerin paylaşımı sırasında kullanılan ağın gizliliği, hata toleransı ve sistemin dayanıklılığı gibi faktörler, güvenlik ve gizliliği doğrudan etkilemektedir. Geleneksel istemci-sunucu modelinin bazı önemli eksiklikleri bulunmaktadır. Bu modelde genellikle bir veya iki sunucu bulunur ve istemciler bu sunuculara dosya paylaşımı veya erişimi talepleri gönderirler. Ancak, merkezi bir sunucuya bağlı olan bu sistemler, sunucunun arızalanması veya saldırıya uğraması durumunda işlevsiz hale gelebilir. Ayrıca, sistem bileşenlerinden birinin çökmesi, tüm sistemi devre dışı bırakabilir ve güvenlik açıklarına yol açabilir. Kullanılan ağ genellikle sanal, gizli veya özel değildir, bu da veri paylaşımını güvenlik açısından riskli hale getirmektedir. Bu çalışmada amaç, merkeziyetsiz bir özel ağ içerisinde eşten eşe dosya depolama ve paylaşım sisteminin oluşturulmasıdır. Çalışma, IPFS ve Tinc VPN kullanılarak güvenli ve dayanıklı bir veri paylaşım altyapısı sunmayı hedeflemektedir. Tasarlanan sistem, veri paylaşımı sırasında gizlilik, dayanıklılık ve güvenliği ön planda tutan bir yapı sunmaktadır.

Anahtar Kelimeler— depolama ve paylaşım, ipfs, merkeziyetsiz, sanal özel ağ, tinc, veri aktarımı

Decentralized File Storage and Sharing Implementation Powered by Tinc VPN

Abstract— Data security is of critical importance for organizations and especially companies. Factors such as the confidentiality of the network used during data sharing, fault tolerance and system durability directly affect security and privacy. The traditional client-server model has some important deficiencies. In this model, there are usually one or two servers and clients send file sharing or access requests to these servers. However, these systems, which are connected to a central server, can become dysfunctional if the server malfunctions or is attacked. In addition, a crash of one of the system components can disable the entire system and lead to security vulnerabilities. The network used is usually not virtual, private or private, which makes data sharing risky in terms of security. The aim of this study is to create a peer-to-peer file storage and sharing system within a decentralized private network. The study aims to provide a secure and durable data sharing infrastructure using IPFS and Tinc VPN. The designed system offers a structure that prioritizes confidentiality, durability and security during data sharing.

Keywords— storage and sharing, ipfs, decentralized, virtual private network, tinc, data transfer

1. GİRİŞ (INTRODUCTION)

Kurum ve kuruluşlar için dosya paylaşımının güvenli ve dayanıklı bir şekilde merkezi olmayan bir yapıda gerçekleştirilmesi, iş verimliliği ve iletişim açısından kritik bir rol oynamaktadır. Geleneksel merkezi dosya paylaşım sistemleri, etkin merkezi kontrol ve kaynak yönetimi ile kullanıcıya özerklik sağlasa da [1] çeşitli dezavantajlara sahiptir. Bu dezavantajlar, sistemin dayanıklılığı, hata toleransı ve güvenlik açığı ile ilgili bazı sorunları içermektedir [2]. Merkezi sistemlerde sıklıkla karşılaşılan sorunlar şunlardır:

- **Dayanıklılık Eksikliği:** Merkezi sistemler, tek bir veya sınırlı sayıda sunucuya bağlıdır. Sunuculardan birinin çökmesi durumunda sistemin işlerliği kesintiye uğrar. Ayrıca, hedef belirli olduğu için saldırılar karşısında savunmasız kalırlar. Bu tür sistemlerde tek bir sunucunun devre dışı kalması bile sistemin genel işleyişini olumsuz etkiler.
- **Düşük Hata Toleransı:** Merkezi sistemlerde sadece ana düğüme (node) talepler gönderilir. Ana düğümde herhangi bir hata oluştuğunda sistem yanıt vermez ve işleyiş kesintiye uğrar. Merkezi sistemlerin sınırlı hata toleransı nedeniyle güvenlik riski yüksektir.
- **Ağın Güvenlik Açığı:** Kullanılan ağ sanal, gizli veya özel olmadığı için güvenlik zayıftır. Bu durum, verilerin güvenli bir şekilde iletilmesini zorlaştırır ve saldırılara karşı açık hale getirir. Böylece veri bütünlüğü ve gizlilik riskleri de ortaya çıkmış olur [1].

Bu çalışmada, merkezi sistemlerdeki bu problemlerin üstesinden gelmek amacıyla, Gezegenler Arası Dosya Sistemi (InterPlanetary File System, IPFS) ve Tinc VPN (Virtual Private Network, Sanal Özel Ağ) kullanılarak merkeziyetsiz ve güvenli bir dosya paylaşım altyapısı oluşturulmuştur. IPFS, dosya paylaşımını merkezi olmayan bir ağ üzerinde gerçekleştirirken Tinc VPN, ağın güvenliğini ve gizliliğini sağlar. Önerilen sistemin temel avantajları şunlardır:

- **Yüksek Dayanıklılık:** Sistem, beklenen veya beklenmeyen değişikliklere karşı kendi işleyişini sürdürme yeteneğine sahiptir. Bu, dosya paylaşımının sürekli bir şekilde devam etmesini sağlar.
- **Güvenilir Veri Saklama:** Veriler, güvenli bir şekilde saklanır ve sadece yetkili kişilerle paylaşılır. Verilerin yedeklenme ihtiyacı ortadan kalkar, çünkü her dosya birçok farklı düğümde depolanır.
- **İstemci-Sunucu Modelinin Ötesinde:** Sistemdeki her düğüm hem istemci hem de sunucu olabilir. Bu, dosya paylaşımında esneklik sağlar.
- **Kesintisiz Güncellemeler:** Sistem güncellemeleri, işleyişi kesintiye uğratmadan yapılabilir.
- **Ölçeklenebilirlik:** IPFS ve Tinc VPN, sisteme daha fazla düğüm eklenmesine olanak tanır, böylece ağ genişledikçe performans da artar.
- **Güvenli ve Gizli İletişim:** Tinc VPN, düğümler arasındaki iletişimi güvenli hale getirir, böylece veri aktarımı sırasında gizlilik sağlanır.
- **Merkezi Denetimden Bağımsızlık:** Sistem, merkezi bir otoriteye ihtiyaç duymadan işleyebildiği için sansür ve izleme riskleri minimize edilmiş olur.

IPFS, içerik adreslemesi ile dosyalara ağ üzerinde hızlı ve güvenilir bir şekilde erişimi sağlar. Dosyalar, içeriklerine göre benzersiz kimlikler ile tanımlanır [2] ve her dosyanın bütünlüğü IPFS'nin oluşturduğu özet (hash) fonksiyonları ile korunur [3-4]. Bu sistem, veri güvenliği ve erişim hızını artırarak merkeziyetsiz dosya paylaşım altyapısını sağlamlaştırır.

Sonuç olarak, bu çalışmada önerilen sistem, veri güvenliği ve dayanıklılık açısından merkezi sistemlere göre üstünlük sağlar. Ağ üzerindeki her düğümün eşit derecede işlevsel olduğu bu modelde, sistemin herhangi bir parçası arızalansa dahi işleyiş kesintiye uğramaz. Ayrıca, sistemin hata toleransı yüksektir ve sistem saldırılara karşı dirençlidir. IPFS ve Tinc VPN teknolojileri ile desteklenen bu yapı, güvenli ve sürdürülebilir bir dosya paylaşım altyapısı sunmaktadır.

Bölüm 2'de sanal özel ağ yapısı gibi çalışmada geçen temel kavramlarla ilgili açıklamalar sunulmaktadır. Bölüm 3'te literatürdeki farklı çalışmalar incelenmektedir. Bölüm 4'te uygulamada kullanılan yazılım teknolojileri ile altyapılar ele alınmaktadır. Bölüm 5'te ağ testlerinin sonuçları ve uygulama arayüzü sunulmaktadır. Bölüm 6'da önerilen sistemin avantajları ve kısıtları tartışılmaktadır. Son olarak Bölüm 7'de gelecek çalışmalarla ilgili öneriler verilmektedir.

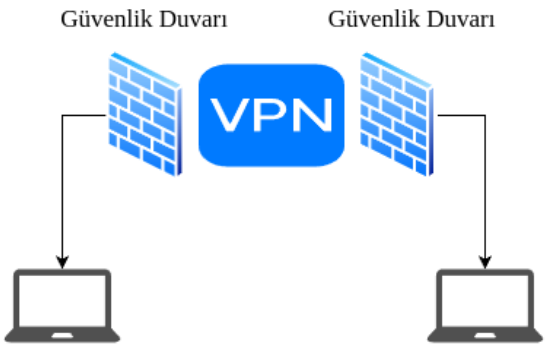
2. GENEL TANIMLAR (GENERAL DEFINITIONS)

Bu çalışmada ele alınan temel konular, sanal özel ağlar ve merkeziyetsiz depolama sistemleridir.

2.1. Sanal Özel Ağlar (Virtual Private Networks)

VPN, kullanıcıların internete açık bir ağ üzerinden güvenli ve gizli bağlantılar kurmalarını sağlayan bir ağ teknolojisidir. Bir VPN, kullanıcıyı fiziksel olarak o ağa bağlı olmadan, farklı bir ağa sanal bir şekilde bağlar. Bu bağlamda VPN, veri transferi sırasında kullanıcıların konum bilgilerini gizler ve güvenli veri aktarımı sağlar. Özellikle şirket içi ağlara uzaktan erişim sağlamak için kullanılan VPN teknolojisi, internet üzerinden güvenli bir tünel oluşturarak özel ağlara erişim imkânı sunar [5].

VPN'nin sağladığı en büyük avantajlardan biri, verilerin şifrelenerek aktarılması ve bu sayede yetkisiz erişimlerin önlenmesidir. VPN aynı zamanda esneklik, güvenlik ve yüksek üretkenlik gibi faydalara sahiptir [6]. Şekil 1, temel bir VPN yapısını şematik olarak göstermektedir.



Şekil 1. VPN şeması
(Figure 1. VPN schema)

Tablo 1'de VPN'nin güvenlik ve erişim avantajlarının yanı sıra yapılandırma hataları ve hız düşüşü gibi dezavantajları vurgulanmaktadır.

Tablo 1. VPN'nin avantajları ve dezavantajları
(Table 1. Advantages and disadvantages of VPN)

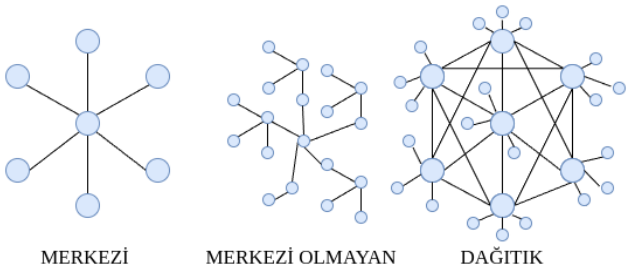
VPN'nin Avantajları	*VPN'nin Dezavantajları*
Veriler, şifrelenerek güvenli bir şekilde iletilir. Böylece internet tehditlerine karşı koruma sağlanır.	Yanlış yapılandırılması durumunda DNS ve IP sızıntıları gibi güvenlik açıkları oluşabilir.
Kablosuz ağlar üzerinden güvenli bağlantılar kurulmasına olanak tanır.	Kurulumu ve yönetimi zor olabilir. Özellikle büyük ağlarda yapılandırma hataları güvenlik sorunlarına yol açabilir.
Uzak cihazlar, şirket içi ağlara güvenli bir şekilde bağlanabilir.	İnternet hızında düşüş yaşanabilir; bu, şifreleme ve veri yönlendirme süreçlerinden kaynaklanabilir.

2.2. Merkeziyetsiz Depolama ve Paylaşım (Decentralized Storage and Sharing)

Merkeziyetsiz depolama, verilerin tek bir merkezi sunucu yerine çoklu otoriter düğümlerin bulunduğu bir ağ üzerinde saklandığı ve kullanıcılar arasında paylaşıldığı bir sistemdir. Bu sistemde veriler, merkezi olmayan bir ağ yapısı üzerinde birçok farklı düğümde saklanır ve bu düğümler, verilerin erişilebilir kalmasını sağlar. Merkeziyetsiz depolama, geleneksel bulut depolama sistemlerinin aksine, tek bir şirkete veya veri merkezine bağlı kalmadan veri depolama ve paylaşımını mümkün kılar [7]. Merkeziyetsiz sistemde her bir düğüm, servislerin bir alt kümesini veya belirli bir bölgeyi yönetebilir ve bunu yaparken diğer düğümlerle iletişimini devam ettirir. Bu durum sistem yükünün dengelenmesini, veri senkronizasyonunu ve ağ direncini getirir [1].

Merkez düğümü ortadan kaldırarak bir düğümün diğer birçok düğüme bağlantı sağladığı farklı bir yapı da dağıtık sistemlerdir [1]. Dağıtık sistemlerde veriler, birçok düğüm arasında eşit olarak dağıtılır. Merkezi sistemlerdeki gibi tek bir otoritenin kontrolü altında olmayan bu yapı, kullanıcıların veri depolama ihtiyaçlarını daha güvenli ve esnek bir şekilde karşılamalarını sağlar. Merkeziyetsiz depolama, aynı zamanda veri artıklığı ve yedekliliği sağlayarak veri kaybı riskini minimize eder [8]. Şekil 2, merkezi ve merkeziyetsiz ağ yapıları arasındaki farkları göstermektedir [1].

Merkeziyetsiz depolama teknolojileri, kullanıcıların veri kontrolünü artırırken merkezi sistemlerin yarattığı güvenlik ve gizlilik risklerini azaltır. Merkeziyetsiz bir uygulama, genellikle blok zinciri gibi dağıtık sistemler üzerine inşa edilir ve verilerin güvenli ve şeffaf bir şekilde işlenmesini sağlar [9].



Şekil 2. Ağ yapıları
(Figure 2. Network structures)

3. LİTERATÜR TARAMASI (LITERATURE REVIEW)

Merkeziyetsiz depolama ve veri paylaşım sistemleri ile ilgili literatürde birçok farklı çalışma mevcuttur. Bu bölümde özellikle son yıllarda yapılan çalışmalardan bazılarına yer verilmektedir. Çalışmaların çoğunun blok zinciri konusuna bağlı kaldıkları, bazılarının ise burada önerilen yapıdaki gibi IPFS teknolojisinden yararlandığı görülmektedir. Literatür taramasında Tinc VPN kullanılarak hazırlanan ve uygulama testleri için farklı senaryoların tasarlandığı bir çalışma ile karşılaşılmamıştır. Aşağıda literatürdeki çalışmaların her biri için ilgili açıklamalardan sonra bu çalışmada önerilen sistem ile karşılaştırmasına da yer verilmektedir.

Nevpurkar vd. (2020) IPFS ve Ethereum blok zinciri entegrasyonu ile merkeziyetsiz bir dosya depolama ve paylaşım sistemi geliştirmişlerdir [10]. Çalışma, IPFS'te depolanan dosyaların özet değerlerini Ethereum blok zincirinde saklayarak veri güvenliği ve bütünlüğünü sağlamayı amaçlamaktadır. Blok zinciri tabanlı bu yaklaşım, veri manipülasyonuna karşı güçlü bir koruma sağlarken, sistemdeki blok zinciri kullanımı ağ performansını ve ölçeklenebilirliği olumsuz etkileyebilir.

Önerilen sistemde ise Tinc VPN ile oluşturulan güvenli ağ sayesinde benzer bir güvenlik sağlanmakla birlikte daha düşük bir işlem maliyeti ve ağ gecikmesi avantajı sunulmaktadır.

Ghosh vd. (2023) oldukça büyük veri miktarları için erişim kontrolünü de yöneten merkeziyetsiz bir bulut depolama ve paylaşım sistemi tasarlamışlardır [11]. Çalışmada IPFS yapısı kullanılarak büyük boyuttaki multimedya dosyalarının parçaları ağ üzerindeki çeşitli düğümlerde depolanmaktadır. IPFS vekil sunucuları aracılığıyla kullanıcıların verilere erişim aşamaları için gerekli olan kimlik doğrulama adımlarında blok zinciri ağı çalıştırılmaktadır. Böylece güvenli bir paylaşım sistemi ile dosyaların kaç kişi tarafından görüntülendiği, dosyalar üzerinde bir değişiklik olup olmadığı kontrol edilebilmektedir. Önerilen sistem, Ghosh vd.nin (2023) çalışmalarındaki erişim kontrolüne dayalı güvenlik mekanizmalarına ek olarak, Tinc VPN ile sağlanan şifreli ağ yapısı sayesinde veri gizliliğini artırırken IPFS-Cluster entegrasyonu ile ölçeklenebilirliği ve ağ verimliliğini optimize etmektedir.

Peng vd.nin (2023) geliştirdikleri modelde, konsorsiyum blok zinciri tabanlı bir eşler arası dosya paylaşım sistemi önerilmiştir [12]. Bu sistem, kimlik doğrulama ve erişim kontrolü mekanizmalarıyla güvenli veri paylaşımını önceliklendirmektedir. Çalışma, özellikle çapraz organizasyonel veri paylaşımı için etkili çözümler sunsa da konsorsiyum blok zincirinin sınırlı katılımcı yapısı, ölçeklenebilirlik açısından bir dezavantaj oluşturabilir. Önerilen sistemde ise Tinc VPN ile oluşturulan şifreli ağ yapısı ve IPFS-Cluster entegrasyonu sayesinde daha geniş ölçekli ve dinamik bir cihaz ağı desteklenmektedir.

Shamdasani vd. (2023) IPFS ve blok zinciri teknolojilerini entegre ederek merkeziyetsiz bir dosya depolama sistemi oluşturmuşlardır [13]. Çalışma, dosya paylaşımında güvenliği artırmak için akıllı sözleşmelerle erişim kontrolünü entegre etmektedir. Bu sistemin avantajı, IPFS'nin dosya bölme ve kriptografik özetleme özelliklerini kullanarak yüksek güvenlik sunmasıdır. Ancak, merkezi bir yapıdan tamamen bağımsız olmaması ve IPFS'nin performans sorunları, özellikle büyük dosya işlemlerinde sınırlayıcıdır. Önerilen yöntemin Shamdasani vd.nin (2023) çalışmalarından farkı, Tinc VPN ile güvenli bir ağ kurulması ve IPFS-Cluster ile küme oluşturulmasının veri paylaşımında ek bir güvenlik ve performans avantajı sağlamasıdır.

Du vd. (2024) IPFS tabanlı sistemlerde maliyet ve güvenlik dengesini sağlamak için DWare isimli bir ara katman sunmuşlardır [14]. Çalışmada veri tekrarını önleme ve kontrol süreçleri iyileştirilmiş, ancak otomasyon çözümlerine yer verilmemiştir. Önerilen sistemde Tinc VPN ve IPFS-Cluster ile sağlanan güvenlik, DWare'in yalnızca veri tekrarını önlemeye odaklanan yapısından daha kapsamlıdır.

Lee vd. (2024) IoT cihazlarının depolama sorunlarını çözmek için IPFS ile merkezi bir veri yönetim sistemini birleştiren bir model sunmuşlardır [15]. Sistem, MQTT protokolü ve bir veritabanı ile veri organizasyonu sağlamaktadır. Öte yandan otomasyon çözümleri daha geniş kapsamda uygulanabilirlik sunmaktadır. Önerilen sistemde IPFS-Cluster altyapısı ile ağ etkileşimi optimize edilmiştir.

Zang vd.nin (2024) çalışmalarında merkeziyetsiz bir depolama sistemi tasarımı için blok zinciri ve bulut yerel teknolojiler birleştirilmiştir [16]. Bu sistem, kenar bilişim altyapılarında yüksek veri aktarım hızı ve esneklik sunar. Sistem, veri erişiminde ölçeklenebilirliği artırırken blok zinciri teknolojisini kullanarak veri bütünlüğü ve güvenilirlik sağlar. Bununla birlikte, IPFS'ye kıyasla veri aktarım performansında avantaj sağlasa da, sistemin kurulum ve işletim karmaşıklığı bir dezavantaj olarak değerlendirilebilir. Önerilen sistemde Tinc VPN kullanımı ile veri güvenliğinin artırılması, özellikle hassas verilerin korunması açısından üstünlük sağlamaktadır.

Han ve Son (2025), IPFS ve blok zinciri kullanarak belgelerin daha güvenli ve şeffaf bir şekilde yönetilmesini sağlayan bir sistem tasarlamışlardır [17]. Shamir'in Gizli Paylaşım Yöntemi ile belge yetkilendirme yapılmış, fakat otomasyon çözümlerine yer verilmemiştir. Önerilen sistem, ağ yapısı oluşturmada daha geniş bir esneklik ve kapsam sunmaktadır.

Samuel ve Kasturi (2025), bulut tabanlı bir blok zinciri ağı üzerinde derin öğrenme uygulamasıyla anahtar üretimi yaparak veri güvenliğini ve gizliliğini sağlamışlardır [18]. SpinalNet yapısının kullanımı, şifreleme süreçlerinde yüksek doğruluk ve güvenlik sağlarken sistemin karmaşık altyapısı ve yüksek işlem maliyetleri bir dezavantaj olarak öne çıkmaktadır. Önerilen sistemde ise ağ yapısının sadeliği ve düşük işlem maliyeti sayesinde daha pratik ve ölçeklenebilir bir çözüm sunulmaktadır.

Bu çalışmalar, merkeziyetsiz veri paylaşımı konusunda çeşitli yaklaşımlar sunmakta ve güvenlik, performans ile ölçeklenebilirlik gibi farklı parametreler üzerinde durmaktadır. Önerilen sistem, Tinc VPN ve IPFS-Cluster entegrasyonu ile düşük maliyetli, güvenli ve esnek bir altyapı sunarak mevcut çalışmalardan ayrılmaktadır.

4. MATERYAL VE METOT (MATERIALS AND METHODS)

Bu çalışma süresince kullanılan teknolojiler, altyapı bileşenleri ve sistemin gerçekleştirilme adımları bu bölümde detaylandırılmaktadır.

4.1. Kullanılan Teknolojiler (Technologies Used)

Bu çalışmada kullanılan yazılım ve altyapı teknolojileri aşağıda sıralanmıştır:

4.1.1. Ubuntu (Ubuntu)

Çalışmada kullanılan ana bilgisayarlar ve test için kullanılan sanal sunucuların işletim sistemi, özgür ve açık kaynaklı bir Linux dağıtımı olan Ubuntu'dur. Ubuntu, kullanıcıların ihtiyaçlarına göre özgürce geliştirilebilir ve ayarlanabilir bir işletim sistemidir. Linux çekirdeği üzerine inşa edilen Ubuntu, hem masaüstü hem de sunucu ortamlarında yaygın bir şekilde kullanılmaktadır. Altyapı, servis ve sistem çalışmalarında geniş bir uyarlanabilirlik sunar.

4.1.2. Tinc VPN (Tinc VPN)

Bu çalışmada kullanılan VPN teknolojisi, açık kaynaklı ve dağıtık ağları destekleyen Tinc VPN'dir. Tinc, 1998 yılında başlatılan ve örgü (mesh) ağ yapısını destekleyen, şifreleme ve sıkıştırma gibi özelliklere sahip bir VPN yazılımıdır. Tinc, trafiği fırsatçı bir şekilde yönlendiren ve ağın otomatik olarak yeniden yapılandırılmasına olanak tanıyan bir yapıya sahiptir [19].

Bu çalışmada Tinc'in hata toleransı ve kaynak tüketimi test edilmiştir. Bu testlerde, bir düğümün kapanması durumunda ağın çalışmaya devam ettiği gözlemlenmiştir. Ayrıca Tinc, CPU ve bellek kullanımında son derece verimli olup özellikle son kullanıcı cihazlarında düşük kaynak tüketimi ile çalışabilmektedir.

Tablo 2, farklı VPN teknolojilerinin gizlilik, veri bütünlüğü ve kimlik doğrulama açısından karşılaştırmasını göstermektedir [5].

Tablo 2. VPN teknolojilerinin kıyaslaması
(Table 2. Comparison of the VPN technologies)

VPN	Kısıtlar		
	Gizlilik	*Veri Bütünlüğü*	*Kimlik Doğrulama*
Tinc	Evet	Evet	Evet
Htun	Hayır	Hayır	Hayır
Cipe	Evet	Evet	Hayır
PPTP	Evet	Evet	Hayır
L2tpd	Hayır	Hayır	Hayır
OpenVPN	Evet	Evet	Evet
VTUN	Evet	Hayır	Hayır
Vpnd	Evet	Evet	Evet
Yavipin	Evet	Evet	Hayır

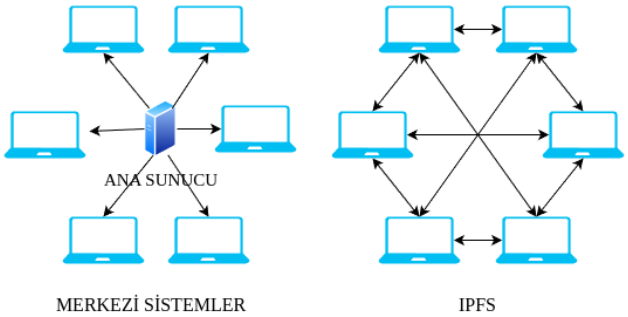
4.1.3. Gezegenler Arası Dosya Sistemi (InterPlanetary File System)

IPFS, merkeziyetsiz dosya depolama ve paylaşımı sağlayan, eşten eşe (peer-to-peer, P2P) bir ağ protokolü ve içerik adreslemeli bir dosya sistemi olarak tasarlanmıştır. 2015 yılında Juan Benet tarafından geliştirilen IPFS, Web'in mevcut HTTP tabanlı merkezi yapısından farklı olarak, daha güvenli, hızlı ve dağıtık bir yapı sunmayı hedeflemektedir [20]. IPFS, dosyaların saklanması ve paylaşılmasında merkezi sunuculara ihtiyaç duymayan, içerik adresleme yöntemini kullanan bir sistem olarak bilinir. Bu yöntem, dosyaların benzersiz bir İçerik Tanımlayıcı (CID) ile tanımlanmasını sağlar. CID, dosyanın içeriğine dayalı olarak oluşturulan kriptografik bir özet fonksiyonudur. Böylece her dosya, benzersiz bir adresle temsil edilir ve içeriğin güvenilirliğini sağlar. Uyarlanabilirliğini artırmak için çoklu aktarım/ağ protokollerini veya kriptografik karma işlevlerini destekler [21].

IPFS'nin temel çalışma prensibi, veriyi bir düğümde saklayıp, diğer düğümlerin bu veriye erişebilmesidir. Bir IPFS düğümü, başka bir düğümden dosya talep ettiğinde, dosyayı içeren blokları alır ve bu blokları birleştirerek dosyanın tam haline ulaşır. Bu yapı, geleneksel merkezi sunucuların aksine, verilerin birden fazla düğümde saklanmasına olanak tanır. Sonuç olarak, veri kaybı veya sunucu çökmesi gibi durumlar minimize edilir [22].

IPFS'nin sunduğu diğer önemli avantaj, sansüre dayanıklı olmasıdır. Geleneksel Web yapısında, içerik tek bir sunucu üzerinde barındırıldığında, bu sunucuya erişim engellendiğinde içeriğe erişim de engellenmiş olur. Ancak IPFS'de içerik, birden fazla düğümde depolandığı için bu düğümler aracılığıyla içerik erişilebilir hale gelir. Bu, özellikle internet sansürünün yoğun olduğu bölgelerde önemli bir avantaj sağlar [23].

Şekil 3'te görüldüğü gibi IPFS, diğer dosya paylaşım sistemlerinden farklı olarak merkeziyetsiz bir yapıda çalışır. Dosyaların depolanmasında ve paylaşımında, bir merkezi sunucu ya da herhangi bir hizmet sağlayıcısına ihtiyaç duymaz. Bu nedenle diğer dosya paylaşım sistemlerine göre daha güvenli ve daha özgür bir platform sunar.



Şekil 3. Veri saklama sistemlerinin karşılaştırması
(Figure 3. Comparison of the data storage systems)

IPFS aynı zamanda verilerin dağıtık bir şekilde depolanmasını sağladığı için, ölçeklenebilirlik açısından büyük avantajlar sunar. Merkezi sistemlerde verinin büyümesi, depolama kapasitesi veya sunucu maliyetlerini artırırken, IPFS'de bu maliyetler ağdaki düğümler arasında dağıtılır. Bu, IPFS'nin büyük veri setlerinin depolanması ve paylaşımı için ideal bir çözüm olmasını sağlar [24]. Tablo 3'te, IPFS'in içerik adresleme, P2P ağ yapısı, sansür direnci ve verimli veri dağıtımı gibi temel özellikleri açıklanmaktadır.

Tablo 3. IPFS'nin temel özellikleri
(Table 3. Basic features of IPFS)

Özellikler	*Açıklama*
İçerik Adresleme	IPFS, verileri sunucunun bulunduğu adresle değil, verinin kendisiyle adresler. Bu, verilerin her zaman benzersiz bir CID ile tanımlanmasını sağlar.
P2P Ağ	Veriler merkezi bir sunucuda tutulmaz, bunun yerine bir P2P ağda saklanır ve paylaşılır.
Yedeklilik ve Güvenlik	Veriler, birden fazla düğümde saklanarak yedeklenir. Bu yapı, veri kaybını önleyerek güvenliği artırır [25].
Sansür Direnci	Merkezi sunuculara bağlı olmadığı için, içerik erişimi engellenemez ve sansüre karşı dayanıklıdır.
Veri Dağıtımı	İçeriğe en yakın düğümden erişim sağlanır, bu da daha hızlı ve verimli veri dağıtımı sunar.

IPFS ile HTTP arasındaki farklılıklar Tablo 4'te yer almaktadır.

Sonuç olarak IPFS, dosya depolama ve paylaşım sistemleri için veri güvenliğini ve esnekliğini sağlayan, sansüre karşı dayanıklı, hızlı ve ölçeklenebilir bir çözüm sunar. Özellikle büyük veri setleri ve merkeziyetsizlik arayışı içinde olan projeler için önemli bir alternatif olarak öne çıkmaktadır.

Tablo 4. HTTP ve IPFS kıyaslaması
(Table 4. HTTP and IPFS comparison)

Özellikler	*HTTP*	*IPFS*
İstemci-Sunucu Modeli	Merkezi sunuculara dayanır.	Merkezi olmayan, P2P yapı.
Adresleme	Veriler sunucu adresiyle istenir.	Veriler içerik adresiyle (CID) istenir.
Erişim Güvenilirliği	Sunucu çöktüğünde verilere erişilemez.	Veriler birden çok düğümde saklanır.
Bant Genişliği Kullanımı	Sunuculara yüklenir, yavaş olabilir.	Yakındaki düğümlerden hızlı erişim.
Sunucu İhtiyacı	Barındırma sunucusu gerektirir.	Ana sunucuya gerek kalmaz.

4.1.4. Sanal Makine (Virtual Machine)

Sanal sunucu teknolojisi, bir fiziksel sunucu üzerinde birden fazla işletim sisteminin barındırılmasına olanak tanıyan bir yaklaşımdır. Bu teknoloji, sunucunun kaynaklarını etkin bir şekilde kullanarak birden fazla sanal sunucu oluşturulmasına imkân verir. Tüm sanal sunucular, tek bir fiziksel sunucu üzerinde yer almakta olup birbirlerinden bağımsız olarak çalışabilmektedir.

Bu çalışmada uygulama testlerinin gerçekleştirilmesi amacıyla, sanal sunucu merkezi olarak Linode platformu tercih edilmiştir.

4.1.5. Ansible (Ansible)

Ansible, sistemlerin kurulumu, yönetimi ve yapılandırması için kullanılan bir otomasyon dilidir. Genellikle makinelerin istenilen konfigürasyonla yapılandırılması ve yönetilmesi amacıyla kullanılır. Ansible, belirlenen görevlerin gerçekleştirilmesi ve makine konfigürasyon yönetimini birleştirerek kullanıcı dostu bir yapı sunar. Sistemler, çoğunlukla Güvenli Kabuk (SSH) protokolü üzerinden yönetilmektedir. Ansible'ın kullanımı için öncelikle bir Python ortamının kurulmuş olması gereklidir.

4.1.6. Terraform (Terraform)

Terraform, HashiCorp tarafından geliştirilen açık kaynaklı bir altyapı otomasyon aracıdır. Bu araç, kullanıcıların veri merkezi altyapısını tanımlayıp sağlamak için HashiCorp Yapılandırma Dili (HCL) veya isteğe bağlı olarak JavaScript Nesne Gösterimi (JSON) gibi bildirim temelli yapılandırma dillerini kullanmasına olanak tanır. Terraform, sanal sunucu merkezinde yeni makinelerin oluşturulmasında kullanılacak önemli bir araçtır.

4.1.7. Kabuk Betiği (Bash Script)

Uygulamadaki kodların çalıştırılması ve yapılandırılması, Kabuk Betiği ile gerçekleştirilmektedir. Kabuk Betiği, Unix tabanlı bir işletim sisteminde çalışan bir komut satırı yorumlayıcısıdır ve kullanıcıların çeşitli otomasyon görevlerini kolaylıkla gerçekleştirmesine imkân tanır.

4.1.8. JavaScript (JavaScript)

JavaScript, internet sayfalarına dinamiklik kazandırmak amacıyla yaygın olarak kullanılan bir programlama dilidir. Genellikle, Çoklu Metin İşaretleme Dili (HTML) ve Basamaklı Stil Şablonları (CSS) ile birlikte kullanılır. JavaScript, HTML ve CSS ile uyumlu çalışarak kullanıcı etkileşimini artırır ve internet sitelerinin daha dinamik ve etkileşimli hale gelmesine katkıda bulunur. Bu sayede, kullanıcı deneyimini zenginleştiren, etkileşimli ve kullanışlı web uygulamaları oluşturulmasına olanak sağlar.

4.2. Sistemin Gerçeklenmesi (Implementation of the System)

Bu çalışmada geliştirilen sistem, öncelikle makineleri özel bir ağ içine alarak, maskelenmiş IP adresleriyle IPFS'ye entegre edilmekte ve IPFS-Cluster aracılığıyla bir küme yapısı oluşturmaktadır. Altyapı, JavaScript kullanılarak bir web sayfasına aktarılmaktadır. Sistem kurulumu ve konfigürasyonu, Ansible, Terraform ve Kabuk Betikleri gibi otomasyon araçları aracılığıyla gerçekleştirilmektedir. Uygulamanın işleyişine dair detaylar Şekil 4'te sunulmaktadır.

4.2.1. Sunucu Makinelerinin Oluşturulması (Creating Server Machines)

Uygulamanın altyapısı, sanal sunucu merkezinde kiralanan makinelerde test edilmiştir. Bu testler için Linode adlı sanal sunucu sağlayıcısında Terraform kodu ve ilgili Kabuk Betikleri kullanılarak üç adet sanal makine oluşturulmuştur. Uygulamanın diğer bileşenleri Ansible ile geliştirilmiştir.

4.2.2. Envanter Dosyasının Oluşturulması (Creating the Inventory File)

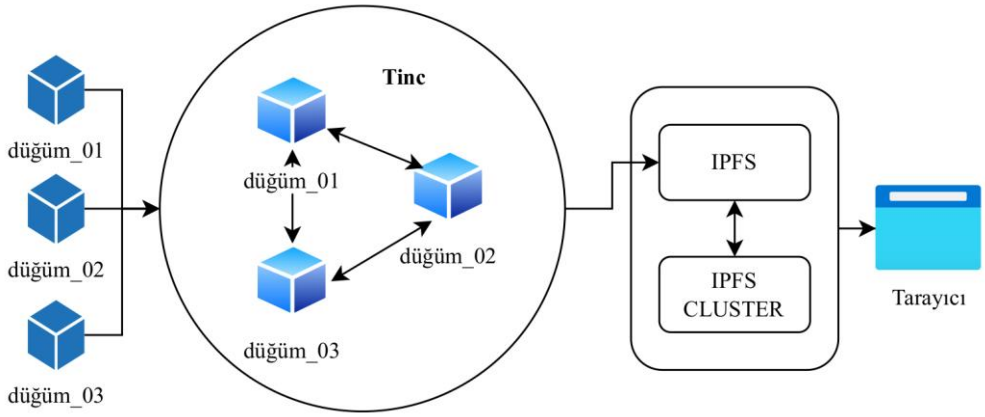
Ansible ile kurulacak makinelerin bilgilerini içeren envanter dosyası, sistemin yapılandırılması için kritik bir bileşendir. Bu dosyada makinelerin IP adresleri, kullanıcı adları ve bilgisayar adları gibi bilgiler yer almalıdır. Özel IP adresleri, yalnızca ağ içinde diğer cihazlarla veri transferinde kullanılırken, genel IP adresleri internete erişim için kullanılmaktadır. Özel IP adresleri, internete yönlendirilmediği için makinelerin iç IP adresleri olarak da adlandırılabilir.

4.2.3. Sanal Özel Ağın Oluşturulması (Creating the Virtual Private Network)

Bu aşamada Tinc VPN teknolojisi kullanılmaktadır. Tinc, tüm makinelerin birbirleriyle güvenli bir şekilde iletişim kurmasını sağlamak için bir rol aracılığıyla yapılandırılmaktadır. Güvenlik önlemleri için Kolay Ateş Duvarı (UFW) kullanılmaktadır. Böylece belirli portların açık olması sağlanarak güvenlik duvarı yönetimi basitleştirilmektedir. Ağın güvenliği için gerekli yapılandırmalar, Tinc ile sağlanmakta ve UFW üzerinden gerekli izinler verilmektedir.

4.2.4. IPFS Yapılandırması (IPFS Configuration)

Bu yapılandırmada dosya paylaşım sistemi olarak IPFS kurulumu gerçekleştirilmektedir. IPFS, Go programlama diline dayalı olarak çalıştığı için öncelikle Go'nun sistemde bulunması gerekmektedir. Daha sonra IPFS'nin yapılandırması, ilgili kullanıcı ve grup tanımlamaları ile devam etmektedir. IPFS, yüklenen dosyaların diğer makinelerde sabitlenmesini sağlamakta ve böylece merkeziyetsiz bir veri depolama çözümü sunmaktadır.



Şekil 4. Uygulama şeması
(Figure 4. Application schema

4.2.5. IPFS-Cluster Yapılandırması (IPFS-Cluster Configuration)

IPFS-Cluster, IPFS ile entegre çalışarak makineleri bir küme halinde organize eder. Bu yapı, yüklenen dosyaların ağdaki diğer tüm makinelerde sabitlenmesini sağlamaktadır. IPFS-Cluster kurulumunda, makinelerin aynı kümede olduğunu belirten anahtarlar ve diğer yapılandırma bilgileri bir değişken dosyasında tanımlanmaktadır. Bu şekilde, sistemin sürdürülebilirliği ve verimliliği artırılmaktadır.

4.2.6. IPFS-App Yapılandırması (IPFS-App Configuration)

Son olarak, oluşturulan altyapı JavaScript/Node.js ile bir web arayüzüne aktarılmaktadır. Node.js'nin açık kaynaklı yapısı, geliştirme sürecinde esneklik sağlamaktadır. Uygulama, belirli dizinlerin oluşturulması ve gerekli Node Package Manager (NPM) paketlerinin yüklenmesi ile başlatılmaktadır. Web uygulaması, kullanıcılara http://localhost:3000 adresinden erişim imkânı sunarak sistemin işleyişini kolaylaştırmaktadır.

Bu yapı, merkeziyetsiz ve özel bir sanal ağ içinde dosya paylaşım ve depolama sisteminin oluşturulmasını sağlamaktadır.

5. BULGULAR (FINDINGS)

Bu bölümde, oluşturulan VPN için gerçekleştirilen testlerin sonuçları ve sistemin işlevselliği ele alınmaktadır. Elde edilen bulgular, altyapının performansını ve kullanımını daha kapsamlı bir şekilde açıklamaktadır.

Testler için Tablo 5'teki özelliklere sahip 3 sanal makine ve 1 fiziksel makine kullanılmıştır.

Tablo 5. Test ortamındaki makinelerin teknik özellikleri
(Table 5. Technical properties of the machines in the test environment)

Özellik	Kategori	
	Fiziksel Makine	*Sanal Makine*
İşletim Sistemi	Ubuntu 20.04	Ubuntu 20.04
İşlemci (CPU)	4	1
Bellek	16 GB RAM	1 GB RAM
Depolama	512 GB SSD	25 GB
Makine Bölgesi	İstanbul, Türkiye	Frankfurt, Almanya (eu-central)
Bağlantı Hızı (Giriş/Çıkış)	33/25 Mbps	40/1 Gbps

5.1. Ağ Testleri (Network Tests)

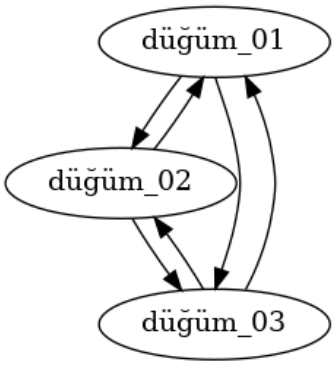
Kurulum tamamlandıktan sonra, makineler arası bağlantı ve iletişim testleri gerçekleştirilmiştir. Makinelerin portunun kapatılması, makinelerin tekrar başlatılması ya da kapatılması gibi durumlarda ağın durumu grafiklerle izlenmiştir. Bu grafiklerin oluşturulabilmesi için Tinc konfigürasyon dosyasında grafik özelliğinin aktif hale getirilmesi gerekmektedir. Tinc servisi, bu süreçte ağın topolojisini analiz ederek makineler arasındaki bağlantıları .dot formatında grafikler hâlinde üretmiştir. Bu sayede, ağın farklı durumlara nasıl tepki verdiği görsel olarak analiz edilebilmiştir.

Tablo 6'da makine port durumları ve örnek senaryolardaki koşullar belirtilmektedir. Daha sonra her bir örnek senaryo sonucunda elde edilen bulgular sunulmaktadır.

Tablo 6. Makine port durumları ve örnek senaryolar
(Table 6. Machine port states and sample scenarios)

Makineler	Örnek Senaryolar			
	Senaryo 1	*Senaryo 2*	*Senaryo 3*	*Senaryo 4*
düğüm_01	Port kapalı	Port kapalı	Port kapalı	Port açık
düğüm_02	Port açık	Port kapalı	Port açık	Port açık
düğüm_03	Port açık	Port açık	Port kapalı	Port açık

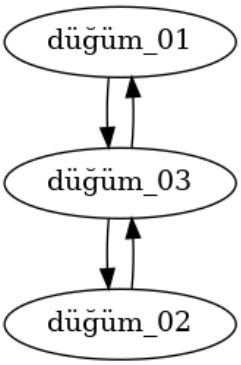
Örnek Senaryo 1: Tüm makineler yeniden başlatıldığında, ağ durumu Şekil 5'te gösterildiği gibi gözlemlenmiştir. Bu senaryo, ağın temel bağlantı durumunu ve iletişim işlevselliğini doğrulamaktadır.



Şekil 5. Örnek senaryo 1
(Figure 5. Sample scenario 1)

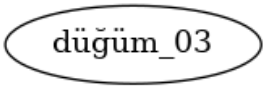
Örnek Senaryo 2: Tüm makineler yeniden başlatıldığında ağ durumu Şekil 6'da yer aldığı gibi kaydedilmiştir. Bu

durum, ağın iki düğümün kapalı olması durumundaki performansını incelemektedir.

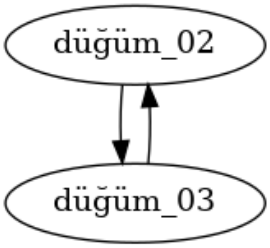


Şekil 6. Örnek senaryo 2
(Figure 6. Sample scenario 2)

Örnek Senaryo 3: düğüm_01 ve düğüm_02 kapatıldığında ağın durumu Şekil 7'deki gibi gözlemlenmiştir. Daha sonra düğüm_02 açıldığında ağ durumu Şekil 8'deki gibi kaydedilmiştir. Bu senaryo, ağın dinamik durum değiştirme kabiliyetini test etmektedir.

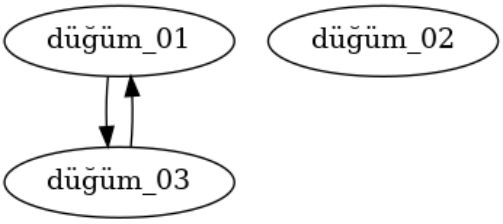


Şekil 7. Örnek senaryo 3 - 1. durum
(Figure 7. Sample scenario 3 - case 1)



Şekil 8. Örnek senaryo 3 - 2. durum
(Figure 8. Sample scenario 3 - case 2)

Örnek Senaryo 4: düğüm_02 kapalı olduğunda ve düğüm_01 ile düğüm_03 açık kaldığında ağın durumu Şekil 9'da gösterildiği gibi gözlemlenmiştir. Bu durum, tüm düğümlerin açık olduğu bir senaryoda ağın genel durumunu değerlendirmektedir.



Şekil 9. Örnek senaryo 4
(Figure 9. Sample scenario 4)

Sistemin kurulumu tamamlandıktan sonra, ağın işleyişini ve makineler arasındaki bağlantıyı doğrulamak amacıyla ağ testleri gerçekleştirilmiştir. Bu testlerde, düğüm_01 isimli makineden diğer ağ makineleri olan düğüm_02 ve düğüm_03'e ping denemesi yapılmıştır. Bu testler sayesinde makineler arasındaki bağlantının düzgün çalıştığı ve ağın işlerliğini koruduğu gözlemlenmiştir.

Şekil 10'da görüldüğü gibi düğüm_01 makinesinden düğüm_02 ve düğüm_03'e yapılan ping denemesi başarılı olmuştur. Bu sonuçlar, VPN üzerinden makinelerin birbirleriyle iletişim kurabildiğini ve verilerin sorunsuz şekilde iletilebildiğini doğrulamaktadır.

```
root@localhost:~# ping 172.16.1.2
PING 172.16.1.2 (172.16.1.2) 56(84) bytes of data.
64 bytes from 172.16.1.2: icmp_seq=1 ttl=64 time=1.23 ms
64 bytes from 172.16.1.2: icmp_seq=2 ttl=64 time=1.70 ms
64 bytes from 172.16.1.2: icmp_seq=3 ttl=64 time=9.62 ms
64 bytes from 172.16.1.2: icmp_seq=4 ttl=64 time=1.80 ms
^C
--- 172.16.1.2 ping statistics ---
4 packets transmitted, 4 received, 0% packet loss, time 3007ms
rtt min/avg/max/mdev = 1.227/3.585/9.615/3.488 ms

root@localhost:~# ping 172.16.1.3
PING 172.16.1.3 (172.16.1.3) 56(84) bytes of data.
64 bytes from 172.16.1.3: icmp_seq=1 ttl=64 time=1.38 ms
64 bytes from 172.16.1.3: icmp_seq=2 ttl=64 time=1.73 ms
64 bytes from 172.16.1.3: icmp_seq=3 ttl=64 time=1.77 ms
64 bytes from 172.16.1.3: icmp_seq=4 ttl=64 time=1.57 ms
^C
--- 172.16.1.3 ping statistics ---
4 packets transmitted, 4 received, 0% packet loss, time 3005ms
rtt min/avg/max/mdev = 1.384/1.614/1.768/0.152 ms
```

Şekil 10. Ağ testi - ping denemesi
(Figure 10. Network test - ping attempt)

Ping testleri sırasında elde edilen zaman değerleri, iki cihaz arasındaki veri iletiminin gecikme süresini yansıtmaktadır. Aşağıdaki analiz, düğüm_01'den düğüm_02 ve düğüm_03 makinelerine yapılan ping denemelerine dayanmaktadır.

Gecikme Süresi (Latency) İncelemesi:

Ping testleri esnasında kaydedilen gecikme süreleri şu şekildedir:

- **düğüm_02 için ping sonuçları**: 1.23 ms, 1.70 ms, 9.62 ms, 1.80 ms
- **düğüm_03 için ping sonuçları**: 1.38 ms, 1.73 ms, 1.77 ms, 1.57 ms

Bu değerler üzerinden analiz yapıldığında, düşük gecikme sürelerinin (örneğin 1-20 ms aralığı) ağın hızlı ve verimli çalıştığını, yüksek gecikme sürelerinin ise (örneğin 100 ms üzeri) ağda olası sorunlar veya yavaşlık olduğunu işaret ettiği bilinmektedir.

Gecikme Süresi Karşılaştırması:

düğüm_02'ye atılan pinglerde gecikme süresi, ortalama olarak 3 ms civarındadır. Bu sonuç, ağ bağlantısının sorunsuz ve hızlı olduğunu göstermektedir.

düğüm_03'e atılan pinglerde gecikme süresi, ortalama 1.5 ms olarak kaydedilmiştir. Bu değer, düğüm_03'ün ağ bağlantısının da istikrarlı olduğunu gösterir, ancak düğüm_02 ile kıyaslandığında biraz daha düşük bir gecikme süresi gözlemlenmiştir. Bu durum, düğüm_02'nin ağ topolojisi içinde daha uzakta yer alıyor olmasından veya daha yoğun bir yük altında olmasından kaynaklanabilmektedir.

Gecikme ve Performansın Önemi:

Gecikme sürelerinin düşük olması, dosya paylaşımı ve veri iletimi açısından sistemin performansını olumlu yönde etkiler. Özellikle IPFS gibi merkeziyetsiz dosya paylaşım sistemlerinde düşük gecikme süreleri, dosyaların daha hızlı bir şekilde erişilebilir olmasını sağlar. VPN ağı üzerinden çalışan bu sistemde, gecikmenin düşük olması güvenli veri iletimini hızlandırır ve sistem performansını artırır.

5.1.1 Yanlış Yapılandırılmış Ağ Testi (Misconfigured Network Testing)

Bu bölümde, bilerek yanlış yapılandırılmış bir düğüm eklenerek ağın güvenlik ve dayanıklılığı test edilmiştir. Temel amaç, bir cihazın ağ üzerindeki IP ve alt ağ (subnet) bilgilerinin yanlış yapılandırılması durumunda ağın bu hatalı düğüme nasıl tepki verdiğini ve veri iletişiminde nasıl bir kesinti yaşandığını incelemektir. Böylece sistemin hataya karşı dayanıklılığını ve ağın işleyişine etkisini değerlendirmek mümkün olacaktır.

Test Aşamaları:

- düğüm_02'den yapılandırma hataları içeren bir düğüme ping gönderme denemesi gerçekleştirilmiştir. Yanlış alt ağ tanımlaması ve /etc/hosts dosyasında gerekli girişlerin eksik olduğu bir senaryo oluşturulmuştur.
- Şekil 11'de test sırasında alınan çıktı verilmektedir. Bu çıktı, düğümün yanlış yapılandırılması nedeniyle veri iletişiminin gerçekleşmediğini ve "Destination Net Unknown" hatası aldığını göstermektedir.

```
root@localhost:/etc/tinc/vpn0/hosts# ping 172.16.1.5
PING 172.16.1.5 (172.16.1.5) 56(84) bytes of data.
From 172.16.1.5 icmp_seq=1 Destination Net Unknown
From 172.16.1.5 icmp_seq=2 Destination Net Unknown
From 172.16.1.5 icmp_seq=3 Destination Net Unknown
From 172.16.1.5 icmp_seq=4 Destination Net Unknown
From 172.16.1.5 icmp_seq=5 Destination Net Unknown
From 172.16.1.5 icmp_seq=6 Destination Net Unknown

--- 172.16.1.5 ping statistics ---
6 packets transmitted, 0 received, +6 errors, 100% packet loss, time 5104ms
```

Şekil 11. Test Çıktısı
(Figure 11. Test Output)

Bu test aracılığıyla yanlış yapılandırılmış bir düğümün ağda nasıl tespit edilebileceği ve veri iletişimine nasıl engel oluşturabileceği gösterilmiştir. Yanlış alt ağ veya eksik ana bilgisayarlar (hosts) dosyası gibi bilgiler nedeniyle düğümler arasında iletişim sağlanamamış ve paket kayıpları meydana gelmiştir. Bu durum, ağdaki hatalı yapılandırmaların hızla tespit edilebileceğini ve düzeltilmezse güvenlik ve veri bütünlüğü sorunlarına yol açabileceğini göstermektedir.

Bu testin güvenlik açısından önemi büyüktür. Yanlış yapılandırılmış bir düğüm, potansiyel bir güvenlik açığı oluşturabilir ve izinsiz veri erişimine veya bilgi sızıntısına neden olabilir. Örneğin, Tinc VPN ve IPFS tabanlı bir dosya sistemi üzerinde düğümlerde hatalı IP veya alt ağ bilgileri tanımlandığında, ağ üzerindeki trafiğin düzgün yönlendirilememesi, dosya transferlerinin kesintiye uğramasına veya yetkisiz düğümler tarafından ele geçirilmesine yol açabilmektedir. Ayrıca, kimlik doğrulama mekanizmalarının yanlış yapılandırılması, kötü niyetli aktörlerin ağa izinsiz erişim sağlamasına zemin hazırlayabilmektedir. Ancak ağın sağlam yapısı, bu tür hataları doğru bir şekilde tespit ederek sistemin bütünlüğünü koruma yeteneğini göstermektedir. Şifreleme protokollerinin etkin kullanımı ve ağ parametrelerinin doğru tanımlanması gibi önlemler, olası güvenlik risklerini azaltabilir. Bu deney, ağın hataya dayanıklı olduğunu, yanlış yapılandırmalardan kaynaklanan olumsuz etkilerin minimize edilebileceğini ve sistemin veri güvenliği ile operasyonel sürekliliği koruma kapasitesini kanıtlamaktadır.

5.2. Tekrarlanabilirlik (Repeatability)

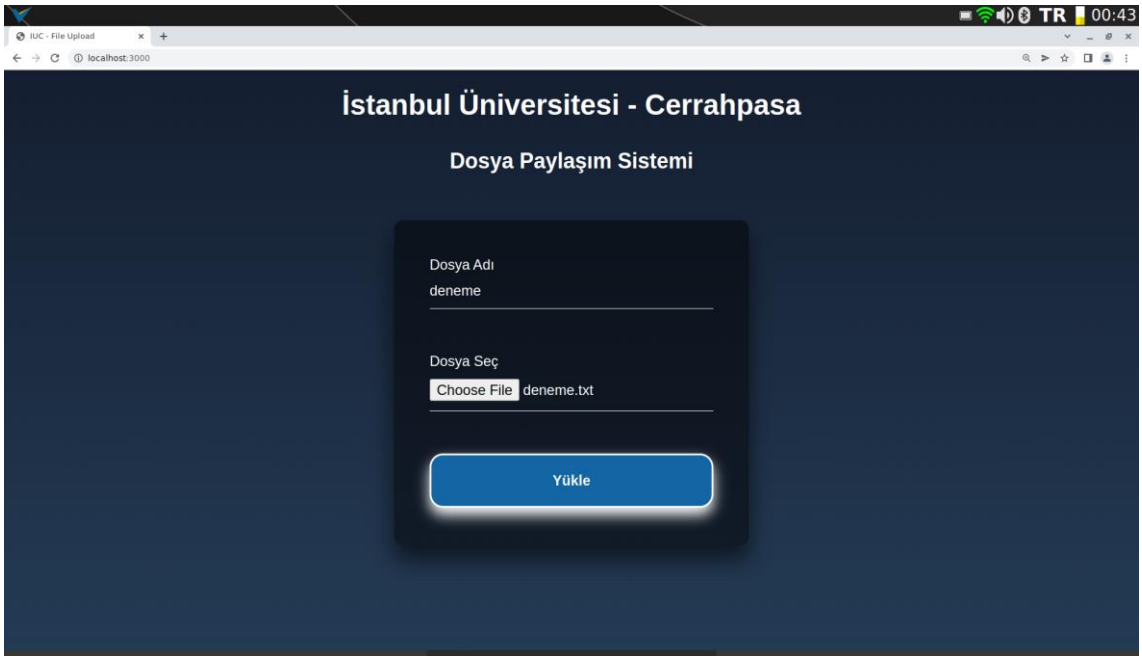
Ağ testlerinde elde edilen bulguların güvenilirliğini sağlamak için farklı senaryolar kullanılmış ve her bir senaryo birden fazla kez tekrarlanmıştır. Özellikle, Tinc VPN ve IPFS tabanlı sistemin dinamik koşullarda nasıl davrandığını anlamak amacıyla cihazların açılıp kapatılması, portların kapatılması gibi çeşitli durumlar uygulamaya dâhil edilmiştir. Bu testlerin her biri, benzer koşullarda tekrarlanarak sonuçların tutarlılığı doğrulanmıştır.

Tekrarlanabilirlik, ağın sadece bireysel testlerde değil, uzun vadede de kararlı ve güvenilir bir performans sergilediğini ortaya koymaktadır. Örneğin, iki düğümün portlarının kapatıldığı bir senaryoda, ağın fonksiyonelliği her testte aynı şekilde etkilenmiş ve yeniden yapılandırma süreçleri tutarlı sonuçlar vermiştir. Ayrıca, ağın dayanıklılığı ve otomatik toparlanma yetenekleri, farklı senaryoların tekrarıyla daha net bir şekilde gözlemlenmiştir. Bu yaklaşım, elde edilen bulguların sistemin genel davranışını temsil ettiğini ve güvenilirlik açısından güçlü bir temel oluşturduğunu göstermektedir.

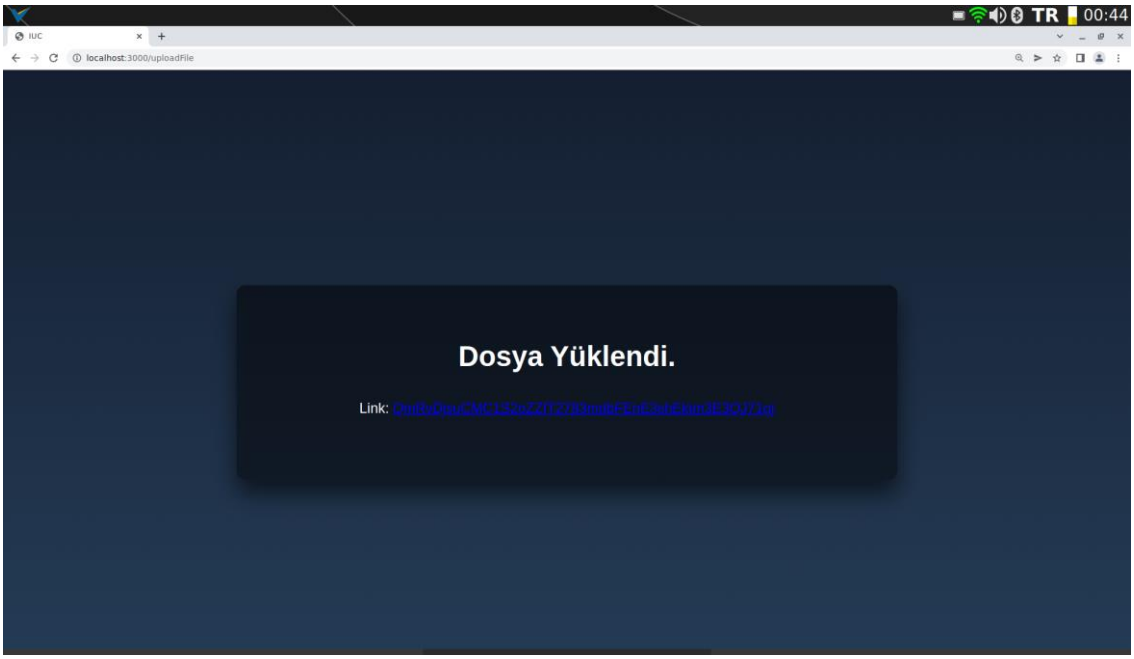
5.3. Uygulama Arayüzü (Application Interface)

Uygulama arayüzü, dosya yükleme sürecini kolaylaştırmak için kullanıcı dostu bir tasarım ile geliştirilmiştir. Şekil 12'de görülen ana sayfada kullanıcı, dosya adını belirleyip yüklemek üzere dosyasını sistemden seçebilir. "Yükle" butonuna tıklandığında, dosya güvenli ve gizli sisteme yüklenmiş olur.

Yüklenen dosyanın erişim linki, Şekil 13'te belirtilmektedir. Ağ içerisindeki diğer makineler bu bağlantıyı kullanarak yüklenen dosyaya erişebilir. Link açıldığında dosyanın içeriği kullanıcıya sunulmaktadır. Bu sistem ile farklı formatlardaki dosyaların da yüklenmesi mümkün olmaktadır. Ayrıca, dosya paylaşım sürecinin hızlandırılması ve erişimin kolaylaştırılması adına, sistemin kullanılabilirliği sürekli izlenmektedir.



Şekil 12. Uygulama arayüzü - dosya yükleme portalı
(Figure 12. Application interface - file upload portal)



Şekil 13. Uygulama arayüzü - dosya linkinin gösterilmesi
(Figure 13. Application interface - showing file link)

Şekil 14'te, yüklenen dosyanın içerik ekranı gösterilmektedir. Kullanıcı, bağlantıyı açarak dosyanın içeriğini doğrudan görüntüleyebilir. Bu ekran, dosyanın adını, türünü ve içeriğini açık bir şekilde sunarak kullanıcıların dosya hakkında hızlı bir değerlendirme yapmalarına olanak tanımaktadır. Uygulamanın bu özelliği, dosya paylaşım sürecini daha şeffaf ve kullanıcı dostu hale getirerek kullanıcıların dosyalarına kolayca erişimini sağlamaktadır.

Şekil 14. Uygulama arayüzü - dosyanın açılması
(Figure 14. Application interface - opening the file)

Yapılan testler sonucunda, dosya yükleme işlemi tamamlandığında, IPFS ağına bağlı her bir düğümün dosyayı başarıyla sabitlediği (pinned) gözlemlenmiştir. Yukarıdaki terminal çıktısında, üç farklı düğüm üzerinde dosyanın sabitlenme süreci ayrıntılı olarak verilmektedir (Şekil 15). Bu, verinin ağ üzerinde kalıcılığını ve erişilebilirliğini sağlayan önemli bir adımı temsil etmektedir.

```
root@localhost:~# ipfs-cluster-ctl add deneme.txt
added
QmRvDjsuCMC1S2oZZfT2783mdbFEnE3shEktm3E3QJ71gj
deneme.txt

root@localhost:~# ipfs-cluster-ctl status
QmRvDjsuCMC1S2oZZfT2783mdbFEnE3shEktm3E3QJ71gj
> düğüm_03      : PINNED | 2024-10-11T20:24:35Z | Attempts: 0 |
Priority: false
> düğüm_02      : PINNED | 2024-10-11T20:24:35Z | Attempts: 0 |
Priority: false
> düğüm_01      : PINNED | 2024-10-11T20:24:35Z | Attempts: 0 |
Priority: false

root@localhost:~# ipfs cat
QmU6iYxgJbZqQ8MAD5aqXwC1tbnoC8hGUDgsnbabcb5G8q
Bu bir deneme dosyasıdır.
```

Şekil 15. Terminal çıktısı - dosyanın paylaşımı
(Figure 15. Terminal output - sharing the file)

5.4. Ek Bulgular (Additional Findings)

Bu sistemin sağladığı güvenlik ve erişim kolaylıkları, kullanıcıların dosyalarını güvenli bir ortamda paylaşmalarını teşvik etmektedir.

Uygulamanın geliştirilmesi sırasında elde edilen kullanıcı geri bildirimleri, arayüzün kullanılabilirliğini artırmak için önemli bir kaynak olmuştur. Kullanıcılar tarafından sağlanan başlıca geri bildirimler şunlardır:

- **Arayüz sadeliği:** Kullanıcılar, arayüzün basit olduğunu belirtmiş, ancak dosya yükleme sırasında belirsizlik yaşadıklarını ifade etmişlerdir. Bu doğrultuda, arayüz sadeleştirilmiş ve geçişler iyileştirilerek kullanıcı deneyimi geliştirilmiştir.
- **Erişim linkinin görünürlüğü:** İlk sürümlerde erişim linkinin yerleşimi kullanıcılar tarafından yeterince belirgin bulunmamıştır. Bu nedenle, linkin daha görünür bir şekilde tasarlanması sağlanmıştır.

Test süreçleri sırasında, kullanıcıların sistemle olan etkileşimlerinin analizi, uygulamanın sürekli geliştirilmesine olanak sağlamıştır. Ayrıca, sistemin güvenlik özellikleri incelendiğinde, şifreleme ve veri bütünlüğü sağlama yöntemlerinin etkili olduğu gözlemlenmiştir. Kullanıcıların, dosyalarını güvenli bir şekilde paylaşarak merkezi sistemlerin risklerinden uzak durmaları teşvik edilmiştir.

Böylece bu çalışma kapsamında sunulan merkeziyetsiz dosya paylaşım sistemi, Tinc VPN ve IPFS altyapılarını kullanarak güvenli ve erişilebilir bir çözüm sunmaktadır.

6. TARTIŞMA (DISCUSSION)

Bu çalışma, kurum veya kuruluşlar arasında dosya paylaşımının güvenli ve gizli bir ağ üzerinden gerçekleştirilmesine olanak tanıyan bir sistemin tasarımını açıklamaktadır. Çeşitli teknolojilerin entegrasyonu ile oluşturulan bu uygulama, iş yönetimi ve akışı için güvenli ve merkeziyetsiz bir altyapı sunmaktadır. Böylece, sistemin işleyişi, farklı ülkelerde kullanılabilir hale getirilmiştir.

Önerilen altyapıda, öncelikle farklı ağlardaki tüm bilgisayarlar Tinc ile güvenli ve şifrelenmiş bir sanal ağa dâhil edilmiştir. Farklı ağlardaki makineler, Tinc ile yapılandırılmış ve ağlar maskelendikten sonra oluşturulan VPN ağına entegre edilmiştir. Daha sonra, IPFS konfigürasyonu ile bu VPN ağındaki cihazlar bir küme haline getirilmiş ve aralarında dosya paylaşımı gerçekleştirilmiştir. Ağda bulunan herhangi bir cihazın yüklediği dosya, şifrelenmiş biçimde diğer cihazlar tarafından özet bilgisiyle görüntülenip kaydedilebilmektedir. Sistemin testleri, sanal sunucu merkezinden sağlanan makinelerle gerçekleştirilmiştir. Uygulamanın sanal makinelerinin oluşturulmasında Terraform, sistemin kurulum kodu için Ansible, ortam ve sistemin yapılandırılmasında ise Kabuk Betikleri kullanılmıştır. Uygulamanın web arayüzüne aktarımı ise JavaScript ile gerçekleştirilmiştir. IPFS ve Tinc teknolojileriyle oluşturulan dosya paylaşım sisteminin birçok avantajı bulunmaktadır:

- **Merkezi Olmayan Yapı:** IPFS ve Tinc, merkeziyetsiz bir yapıda çalıştığından, dosyaların depolanması ve paylaşımında merkezi bir sunucuya ihtiyaç duyulmamaktadır. Bu durum, dosya paylaşım sisteminin güvenliğini artırmaktadır.

- **Güvenlik:** IPFS, dosyaların özet tabanlı adresleme sistemi kullanarak orijinalliğini korumakta ve dosyaların değiştirilmesini zorlaştırmaktadır. Tinc, güvenli ve şifreli bir VPN protokolü olarak iletilen verilerin güvenliğini sağlamaktadır.
- **Ölçeklenebilirlik:** Hem IPFS hem de Tinc, ölçeklenebilir yapıları sayesinde dosyaların depolanması ve paylaşımında genişleme imkânı sunmaktadır.
- **Hız:** IPFS, dosyaların bloklar halinde depolanmasını sağlayarak hızlı bir paylaşım imkânı sunmakta; Tinc ise hızlı bir VPN bağlantısı temin etmektedir.
- **Uygun Maliyet:** IPFS ve Tinc, açık kaynak kodlu ve ücretsiz olarak sunuldukları için dosya paylaşım sisteminin maliyet etkin bir şekilde kurulup yönetilmesine olanak tanımaktadır.

Ayrıca, sistemin öne çıkan özellikleri arasında dayanıklılık, yüksek hata toleransı, güvenli ve özel bir ağda çalışabilme yeteneği, güncellemelerin sistemin işleyişini aksatmadan yapılabilmesi, verilerin yedekleme ihtiyacının ortadan kalkması, arıza durumunda sistemin işlevselliğini yitirmemesi ve her makinenin hem tedarikçi hem de tüketici olabilmesi sayılabilir. Bununla birlikte, çalışmanın bazı kısıtları aşağıdaki gibi özetlenebilir:

- **Veri Erişimi Hızı:** IPFS, dosyaların yakın düğümlerden alınmasını teşvik etse de uzak düğümlerden veri alımında yavaşlık yaşanabilir. Özellikle büyük dosyaların paylaşımı ve indirilmesi durumunda hızlı erişim sağlamak zor olabilir.
- **Veri Bütünlüğü Sorunları:** IPFS, dosya bütünlüğünü sağlamak için kriptografik özet fonksiyonları kullanmasına rağmen, bu özet fonksiyonların kaybolmaması ve verilerin doğru bir şekilde saklanması gereklidir. Dosyaların kaybolması veya değiştirilmesi durumunda bütünlük bozulabilir [26].
- **Bağlantı ve Performans:** Tinc, bir VPN çözümü olarak bazı bağlantı ve performans sınırlamalarına sahip olabilir. Özellikle büyük veri miktarlarının iletimi performans sorunlarına yol açabilir.
- **Kullanım Zorlukları:** Tinc, diğer VPN çözümlerine kıyasla daha karmaşık bir yapıya sahip olabilir. Bu da ayarların yönetilmesi, kurulumu ve yapılandırması için daha fazla teknik bilgi gerektirebilir.

7. SONUÇ VE ÖNERİLER (CONCLUSION AND RECOMMENDATIONS)

Sonuç olarak, bu çalışma, merkeziyetsiz ve P2P mimarisiyle, makinelerin bir veya birkaçının devre dışı kalması durumunda bile işlevselliğini sürdüren bir dosya saklama ve paylaşım sistemi sunmakta ve verilerin paylaşımını daha güvenli ve verimli bir şekilde gerçekleştirmektedir.

Gelecek çalışmalar açısından sistemin daha geniş bir perspektifte değerlendirilebilmesi için şu alanlarda ileri araştırmalar yapılması önerilmektedir:

- **Performans Optimizasyonu:** Farklı ağ yükleri ve kullanıcı yoğunlukları altında sistemin yanıt süreleri ve işlem kapasiteleri detaylı olarak analiz edilmelidir. Özellikle yüksek trafiğe maruz kalan düğümlerin veri işleme hızlarının artırılması için yeni algoritmalar ve protokoller geliştirilebilir.
- **Mobil Entegrasyonu:** Sistem, mobil cihazlarda kullanılabilirlik açısından test edilmelidir. Bu entegrasyon, merkeziyetsiz dosya paylaşımının daha fazla kullanım alanına ulaşmasını sağlayabilir.
- **Çoklu Dil Desteği ve Kullanıcı Deneyimi:** Sistem arayüzü, farklı dilleri destekleyecek şekilde genişletilmelidir. Bu, uluslararası kullanıcıların sisteme erişimini kolaylaştıracaktır. Ayrıca, kullanıcı deneyimini artırmak için farklı yaş ve beceri seviyelerine uygun arayüz tasarımları oluşturulabilir.
- **Akademik ve Endüstriyel Testler:** Sistem, daha geniş bir kullanıcı kitlesi ile gerçek dünyadaki senaryolar altında test edilmelidir. Örneğin, akademik araştırmalarda büyük veri setlerinin paylaşımı veya endüstriyel projelerde hassas veri transferleri gibi senaryolar değerlendirilebilir.
- **Yapay Zekâ Destekli Veri Yönetimi:** Sistemin verimliliğini artırmak adına, yüklenen dosyaların sınıflandırılması ve erişim istatistiklerinin analiz edilmesi için yapay zeka algoritmaları entegre edilebilir. Bu, özellikle ağ üzerindeki yoğunluğun optimize edilmesine ve kullanıcı ihtiyaçlarının daha iyi karşılanmasına olanak tanıyacaktır.
- **Yasal ve Düzenleyici Uyum:** Merkeziyetsiz dosya paylaşımının yasal yönleri araştırılmalı ve farklı ülkelerdeki düzenlemelerle uyumlu bir sistem tasarımı yapılmalıdır. Böylece sistemin daha geniş bir kullanıcı tabanı tarafından benimsenmesi kolaylaşacaktır.

Bu öneriler, sistemin hem teknik hem de pratik açıdan daha işlevsel ve kapsamlı bir çözüm haline gelmesine katkı sağlayacaktır.

KAYNAKLAR (REFERENCES)

[1] M. M. Merlec, H. P. In, "Blockchain-Based Decentralized Storage Systems for Sustainable Data Self-Sovereignty: A Comparative Study", *Sustainability*, 16(17), 7671, 2024.

[2] C. Li, M. Xu, J. Zhang, H. Guo, X. Cheng, "SoK: Decentralized storage network", *High-Confidence Computing*, 4(3), 100239, 2024.

[3] H. R. Hasan, K. Salah, R. Jayaraman, M. Omar, I. Yaqoob, S. Pesic, T. Taylor, D. Boscovic, "A Blockchain-Based Approach for the Creation of Digital Twins", IEEE Access, 8, 34113–34126, 2020.

[4] Z. Fauziah, N. P. Anggraini, Y. P. A. Sanjaya, T. Ramadhan, "Enhancing Cybersecurity Information Sharing: A Secure and Decentralized Approach with Four-Node IPFS", *International Journal of Cyber and IT Service Management*, 3(2), 153–159, 2023.

[5] S. Khanvilkar, A. Khokhar, "Virtual Private Networks: An Overview with Performance Evaluation", IEEE Communications Magazine, 42(10), 146–154, 2004.

[6] L. Ibrahim, "Virtual Private Network (VPN) Management and IPSec Tunneling Technology", *Multi-Knowledge Electronic Comprehensive Journal for Education and Science Publications* (MECSJ), (1), 2017.

[7] T. Sabrina, A. Widjajarto, A. Budiyono, "Memory Analysis for IPFS Implementation on Ethereum Smart Contract", **International Conference on Rural Development and Entrepreneurship 2019: Enhancing Small Business and Rural Development Toward Industrial Revolution** 4.0, Yogyakarta, Endonezya, 1333–1343, 20-21 Ağustos, 2019.

[8] Internet: Cryptopedia Staff, An Overview of Decentralized Cloud Storage Services, https://www.gemini.com/cryptopedia/crypto-cloud-storage-decentralized-cloud-storage-providers, 18.12.2024.

[9] W. Cai, Z. Wang, J. B. Ernst, Z. Hong, C. Feng, V. C. M. Leung, "Decentralized Applications: The Blockchain-Empowered Software System", IEEE Access, 6, 53019–53033, 2018.

[10] M. Nevpurkar, C. Bandgar, R. Deshmukh, J. Thombre, R. Sadafule, S. Bhat, "Decentralized File Storing and Sharing System Using Blockchain and IPFS", *International Research Journal of Engineering and Technology* (IRJET), 7(5), 560-563, 2020.

[11] R. Ghosh, P. Yadav, Z. Khan, L. Panika, "Decentralized File Storing and Sharing System", *International Research Journal of Modernization in Engineering Technology and Science*, 5(5), 5847-5856, 2023.

[12] S. Peng, W. Bao, H. Liu, X. Xiao, J. Shang, L. Han, S. Wang, X. Xie, Y. Xu, "A Peer-to-Peer File Storage and Sharing System Based on Consortium Blockchain", Future Generation Computer Systems, 141, 197-204, 2023.

[13] N. Wadile, J. Shamdasani, S. Deshmukh, M. Sayyed, S. Khandare, "Decentralized File Storage (Interplanetary File System) Using Blockchain", *International Journal of Engineering Research & Technology* (IJERT), 12(3), 252-256, 2023.

[14] Y. Du, A. Zhou, C. Wang, "DWare: Cost-Efficient Decentralized Storage With Adaptive Middleware", IEEE Transactions on Information Forensics and Security, 19, 8529-8543, 2024.

[15] C. Lee, J. Kim, H. Ko, B. Yoo, "Addressing IoT Storage Constraints: A Hybrid Architecture for Decentralized Data Storage and Centralized Management", Internet of Things, 25, 101014, 2024.

[16] H. Zang, H. Kim, J. Kim, "Blockchain-Based Decentralized Storage Design for Data Confidence Over Cloud-Native Edge Infrastructure", IEEE Access, 12, 50083-50099, 2024.

[17] J. Han, Y. Son, "Design and Implementation of a Decentralized Document Management System", Expert Systems with Applications, 262, 125516, 2025.

[18] S. B, K. K, "A Novel Secure Privacy-Preserving Data Sharing Model With Deep-Based Key Generation on the Blockchain Network in the Cloud", Computer Standards & Interfaces, 92, 103932, 2025.

[19] A. A. Pardina, **Comparison of virtual networks solutions for community clouds, Lisans Tezi**, KTH Royal Institute of Technology, Stockholm, İsveç, 2014.

[20] Internet: J. Benet, IPFS - Content Addressed, Versioned, P2P File System, https://arxiv.org/pdf/1407.3561, 16.04.2025.

[21] N. Sangeeta, S. Y. Nam, "Blockchain and Interplanetary File System (IPFS)-Based Data Storage System for Vehicular Networks with Keyword Search Capability", Electronics, 12(7), 1545, 2023.

[22] E. Daniel, F. Tschorsch, "IPFS and Friends: A Qualitative Comparison of Next Generation Peer-to-Peer Data Networks", IEEE Communications Surveys & Tutorials, 24(1), 31–52, 2022.

[23] Internet: zK Capital, IPFS: A Complete Analysis of the Distributed Web, https://medium.com/hackernoon/ipfs-a-complete-analysis-of-the-distributed-web-6465ff029b9b, 18.12.2024.

[24] R. Kumar, R. Tripathi, "Implementation of Distributed File Storage and Access Framework using IPFS and Blockchain", **2019 Fifth International Conference on Image Information Processing** (ICIIP), Shimla, Hindistan, 246–251, 15–17 Kasım, 2019.

[25] R. R. Stewart, Q. Xie, Stream Control Transmission Protocol (SCTP): **A Reference Guide, Addison-Wesley,** A.B.D., 2001.

[26] B. Guidi, M. Conti, A. Passarella, L. Ricci, "Managing Social Contents in Decentralized Online Social Networks: A Survey", *Online Social Networks and Media*, 7, 12–29, 2018.

X (Twitter) Sentiment Analysis Based on Hybrid Approach: An Application for Online Food Ordering

Araştırma Makalesi/Research Article

Yıldırım GÜNEŞ[1], Murat ARIKAN[2]

[1]Gazi University, Graduate School of Natural and Applied Sciences, Department of Supply Chain and Logistcs Management, Ankara, Türkiye
[2]Gazi University, Engineering Faculty, Department of Industrial Engineering, Ankara, Türkiye
yildirimgunes1973@gmail.com, marikan@gazi.edu.tr
(Geliş/Received:09.01.2025; Kabul/Accepted:27.02.2025)
DOI: 10.17671/gazibtd.1616709

Abstract— For sentiment analysis of user opinions on online platforms such as X (formerly known as Twitter), dictionary-based approaches and machine learning methods are generally used. Recent studies emphasize that hybridizing these approaches improves model performance. In this study, we propose a hybrid classification model for sentiment analysis of texts on food ordering. In addition, we suggest a feature selection method based on aggregating words for the high-dimensionality problem of text classification. The main problems in that domain are low number of words with distinctive features, complexity of interpretation of food ordering field, domain dependency of text classification. The use of classification algorithms and a domain lexicon-based approach will contribute to overcoming these difficulties. For this purpose, two domain-specific lexicons are developed using data from online users' opinions, one for sentiment analysis and the other for product-service systems classification, referred to as basic lexicons. Basic lexicons have been transformed into new lexicons with fewer words, referred to as boosted lexicons, by grouping the words in basic lexicons and representing the groups with a single word in boosted lexicons. 144 models of combinations of six classification algorithms, three term weighting methods, and the lexicons are created in a hybrid approach for sentiment analysis. The study used two datasets of 21 039 and 14 389 tweets obtained from X between January 1 and December 31, 2020. The models were trained, tested on the first dataset, and the best models were selected. The second dataset is analyzed with the selected models, we present proposals for the industry.

Keywords— X (twitter) sentiment analysis, lexicon-based classification, online food order, natural language process, feature selection

Hibrit Yaklaşıma Dayalı X (Twitter) Duygu Analizi: Çevrimiçi Yemek Siparişi Üzerine Bir Uygulama

Özet— X (eski adıyla Twitter) gibi çevrimiçi platformlardaki kullanıcı görüşlerinin duygu analizi için, genellikle sözlük tabanlı yaklaşımlar ve makine öğrenmesi yöntemleri kullanılır. Son çalışmalar, bu yaklaşımların hibrit kullanımının model performansını iyileştirdiğini vurgulamaktadır. Bu çalışmada, yemek siparişi ile ilgili metinlerin duygu analizi için hibrit bir sınıflandırma modeli öneriyoruz. Ayrıca, metin sınıflandırmanın yüksek boyutluluk problemi için kelimeleri toplulaştırmaya dayalı bir özellik seçim yöntemi öneriyoruz. Bu alandaki temel sorunlar, ayırt edici özelliklere sahip kelime sayısının düşük olması, yemek siparişi ile ilgili cümlelerin yorumlanmasının karmaşıklığı, metin sınıflandırmanın alan bağımlılığıdır. Sınıflandırma algoritmalarının ve alan sözlüğü tabanlı bir yaklaşımın birlikte kullanılması, bu zorlukların üstesinden gelinmesine katkıda bulunacaktır. Bu amaçla, çevrimiçi kullanıcıların görüşlerinden elde edilen veriler kullanılarak, biri duygu analizi için diğeri ise temel sözlükler olarak adlandırılan ürün-hizmet sistemleri sınıflandırması için olmak üzere iki alana özgü sözlük geliştirilmiştir. Temel sözlükler, bu sözlüklerdeki kelimelerin gruplandırılması ve sözkonusu gruplardan grubu temsil edecek bir kelimenin seçilmesiyle, daha az sayıda kelime içeren ve güçlendirilmiş sözlük olarak adlandırılan yeni sözlüklere dönüştürülmüştür. Duygu analizi için hibrit yaklaşımla, altı sınıflandırma algoritması, üç terim ağırlıklandırma yöntemi ve sözlüklerin kombinasyonlarından oluşan 144 model oluşturulmuştur. Çalışmada, 1 Ocak - 31 Aralık 2020 tarih aralığında X'ten paylaşılmış, 21 039 ve 14 389 tweetten oluşan iki veri seti kullanılmıştır. Modeller eğitilmiş, ilk veri seti üzerinde test edilmiş ve bunların arasından en iyi model seçimi yapılmıştır. İkinci veri seti seçilen modellerle analiz edilmiş ve sektör için öneriler sunulmuştur.

Anahtar kelimeler— X (twitter) duygu analizi, sözlük tabanlı sınıflandırma, çevrimiçi yemek siparişi, doğal dil işleme, özellik seçimi

1. INTRODUCTION

X (formerly known as Twitter) is a platform where users can share their opinions on any topic; new ideas are added to the shared ideas at any time; and millions of tweets are outdated with the newly added tweets. Considering that as of 2022, there are 368.4 million monthly active users [1], it can be estimated how many and varied the number of tweets shared will be. These tweets, which contain user opinions for every sector and every field, are an essential source where business owners can collect positive and negative opinions about the sector. During the pandemic period, there has been an increase in people's tendency to eat at home or work, and these trends and habits, which have become permanent, have been reflected in X posts [2]. Providing timely feedback to the customer by evaluating the customer opinions to be obtained from social media tools that enable the rapid dissemination of customer opinions ensures customer satisfaction. These piles of textual data, which are continuously generated by users, are converted into usable data with sentiment analysis and text classification methods using automated methods.

X (formerly Twitter), user-generated platform, is a useful source of texts that enable customer insights through text classification including sentiment analysis. However, classifications of short texts on a domain to understand the customer emotions continues to be an difficulty due to the low number of words with distinctive features in the texts. The topic, food, is also difficult to interpret and domain dependent. The domain-based studies that make significant contributions to the correct understanding of emotion in a text are at a low level in languages other than English. It is seen that academic studies employ machine and deep learning algorithms, natural language process, domain-based lexicons for text classifications. On the other hand pre-trained language models are also run on textual data analysis. However domain dependency continues especially on sentiment analysis. [3-5].

Sentiment analysis and text classification are performed using natural language processing (NLP) on X texts containing customer opinions written in colloquial chats. For sentiment analysis, tweet contents can be classified as binary, positive and negative; ternary, positive, negative, and neutral; or multiple, with additional emotions such as anger, satisfaction, distrust, etc. [6]. These classifications are carried out using various methods, including artificial intelligence (AI) and NLP methodologies. Text classification is generally based on dictionary-based (as general dictionaries, domain specific lexicons and corpus-based lexicons) approaches, machine learning (ML) and deep learning (DL), transformative, and hybrid approaches [7,8].

Shinde et al [9] performed sentiment analysis on 25 000 tweets with an hybrid method using lexicon and machine learning approach. They employed SVM as classification algorithms, and unigram, bigram and trigram as feature selection method. They achieved in the range 57-62% performances. Vatambeti [10] et al proposed a hybrid model, Convolutional Bi-directional Long Short Term Memory, for tweet texts. The model performed ranging between 83-92 % in text classification. Trust and Minghim [11] studied the performance of seven large language models that are generally successful in text generation but not thoroughly studied in sentiment analysis, on text classification. In this study, it was found that the models performed ranging from 62-99 % in sentiment analysis tasks on five different datasets.

This study based on text classification with an hybrid method on food industry, which is an unstudied domain, will contribute to the solution for classification problem of short text and filling the gap in the domain of Turkish. In the study, we propose a hybrid classification model, deploying classification algorithms and domain lexicon-based approach, for sentiment analysis of texts on food ordering. Two domain-specific lexicons are developed for the model using data from online users' opinions. A classification of each tweet (document level) was made with the hybrid model. The classification models in the study provide performance at a level that can compete with state-of-the-art models.

The models created in the study are used to classify tweets' product and service features as positive and negative. Two models created in the study were used to classify the positive and negative emotions of product and service features of tweets about online food ordering. The results obtained from the model's rapid assessment of many customer opinions that cannot be evaluated manually, can contribute to the industry in two ways. The first is to present the opinions of previous customers to new customers with information such as "the level of customer satisfaction with the product and service," thus facilitating and guiding customers' decision-making. The second is to use the results obtained to improve the product and service by transforming them into tasks for the stakeholders involved in the supply chain. Thus, customer satisfaction and competitive advantage can be achieved.

In order to generate the dataset for the study, data was extracted from X with the Turkish keywords yemek sipariş (food order), yemeksepeti sipariş (basket of food

order), döner sipariş (döner kebab order), lahmacun sipariş (thin Turkish pizza order), hamburger sipariş (hamburger order), and pide sipariş (pita order). The collected tweets were used in two separate datasets: 21 039 tweets and 14 389 tweets. The tweets belong to dates between 1st January and 31st December 0f 2020.

There are some problems encountered in text classification models. One of them is the domain dependency of the word and text [12]. The words used in a text have specific meanings related to the domain. Classifying with general dictionaries causes a decrease in classification performance since specific meanings are not taken into account [13,14]. Another challenge of text classification, such as short texts or tweet posts, is the small number of terms with distinctive features. If these distinctive terms in a document cannot be selected as features, the number of unclassified or misclassified documents increases. Another difficulty in classifying sentences on the topic of food ordering is the difficulty in interpreting conversations on this topic [3]. This study aims to develop models that contribute to solving these challenges through the hybrid use of ML and lexicon-based approaches. To this end, two domain-specific lexicons, called basic lexicons were developed using data from online users' opinions, one for sentiment and one for product-service system classification. Basic lexicons have been transformed into new lexicons with fewer words, referred to as lexicons, by grouping the words in the lexicon and representing the groups with a single word. The feature selection used in this transformation, based on the grouping of words, is a method that also contributes to solving the high-dimensionality problem of text classification.

As a result, we have contributed to the literature with two domain-specific lexicons, an approach that reduces the lexicon size by reducing the number of features for lexicon-based studies, and a model for sentiment analysis and classification of product-service systems.

The rest of the paper is organized as follows: In the second section, a literature review is conducted, focusing on sentiment analysis and text classification methods, as well as their applications in the food sector. The third section introduces the method used for developing the model, its stages, the dataset utilized, evaluation metrics, and the developed domain-specific basic and boosted lexicons. The fourth section evaluates and discusses the analysis, classification results, and the proposed model in detail. The final section presents the conclusions and outlines areas for future research.

2. LITERATUR REVIEW

The text contents in the posts of X users are observed to be unstructured, disorganized, ambiguous in meaning, suggestive, and varied in forms such as jargon and slang. The use of domain-specific emotion terms in such unformatted texts, differences in people's expressions of their emotions and thinking methods, spelling errors, implicit meaning, and ambiguities make sentiment analysis and text classification complicated [4]. The standard phases commonly used for analysis—understanding the task and data, data preparation, modeling, evaluation, and usage—also apply to text analyses [15]. The details of these analyses and phases, which can be used with different nomenclatures in different fields, may vary depending on the social media platform from which the data is drawn, the characteristics and content of the data set, and the analysis objective.

Machine learning and dictionary-based approaches, which are used as basic approaches for text classification and sentiment analysis, can be used in hybrid form as a third approach. Recent studies emphasize that hybrid approaches, which overcome the disadvantages of the basic approaches, improve classification model performance [7,8]. In addition, transformative approaches using advanced techniques such as deep learning also improve the performance of these basic approaches. Alongside the chosen methodologies, the characteristics of the dataset to be analyzed directly impact classification performance.

Text and sentiment classification is fundamentally a word-centric study, focusing on the characteristics of words. A text can be classified at three level—document, sentence, and aspect/feature level—using values derived from words. A commonly used method in these classifications is the aggregation of sentiment scores. These methods can be applied in various ways, such as combining the weight scores of words [16] or aggregating the classification results of different classifiers [17]. Mirtalaie et al. [18] aggregate sentiment polarities by considering the relationship between different features and the desired feature when determining the sentiment value of a specified target feature. In their study examining aggregation methods, Basiri et al. [19] determined the values of words and then performed aggregation at the sentence and general levels based on these values.

There are various challenges in determining the sentiment value of a word that is closest to its natural usage. To overcome these challenges, features such as the position and the specific meaning of the word can be analyzed

separately [13,20,21]. The values obtained from these features determine the polarity score of the word. In a review of 47 studies that looked at problems with classifying emotions, Hussein [12] found problems that were seen in most of them. These problems included the fact that sentiment classification is domain-dependent and it can be hard to figure out whether negative sentences have explicit or implicit meanings. Despite the challenges in text classification, high-performance models are being developed for different sectors [22-24]. Recent studies indicate a growing trend toward hybridizing ML and lexicon-based classification methods. The lexicon-based method utilizes a sentiment lexicon to measure the power of emotions. There are two ways to prepare the sentiment dictionaries. The first is lexicon-based, using general dictionaries as a source, and the second is corpus-based, using the dataset as a source [25,26]. When a ready-made lexicon (the first one) is used, the classification may fail because words not included in the lexicon are not taken into account [27] or the specific meaning of the word is ignored [13,14]. When using a corpus-based approach, the problem of high dimensionality [13] is also encountered, as irrelevant words remain in the corpus even after text cleaning. In addition, the need to update lexicons due to the constant production of content with new and different structures on online platforms, shifts in word meanings, and the derivation of new words can also increase the failure of sentiment classification [28].

In many studies in the literature on dictionary-based classification and sentiment analysis in different languages, dictionaries such as Wordnet, Sentiwordnet, Bing, Afinn, Laughran, SentiStrength, NRC, Bing Liu Opinion Lexicon, and Textblob are utilized. Furthermore, domain-specific dictionaries generated from a limited number of seed word lists, dictionaries developed through automatic or manual translation methods, and specialized dictionaries that include idioms and proverbs are also used in text classification studies [29-45]. In classifying Turkish texts, dictionaries such as TS Corpus, Turkish National Corpus, Spoken Turkish Corpus, SentiTürkNet, and Turkish WordNet are also available as intuition dictionaries prepared with specific methods [46-59]. In languages lacking sufficient resources in terms of dictionaries and training data for text classification, building accurate models and improving model performance can be a significant challenge. Domain-specific studies and transfer learning models such as cross-lingual embeddings can contribute to solving this problem [60,61]. In their study, Kılıcer et al. [62] reported that in sentiment analysis for Turkish, Turkish classification dictionaries did better than translation dictionaries, and hybrid approaches did better than other approaches. However, there were not many hybrid studies at the time.

For ML, various methods can be employed, including supervised, unsupervised, semi-supervised, reinforcement, multi-task, ensemble, and instance-based learning, as well as neural networks [63,64]. The selection of an algorithm in ML depends on factors such as the type of problem, the number of variables, and the appropriate model type for the problem. The superiority of the algorithms can vary, and new methods and algorithms are continually developed to strengthen the weaknesses of previous approaches, leading to the introduction of new versions of algorithms [65-69].

It is observed that text classification algorithms exhibit different performances on various domains and datasets. Naive Bayes (NB), Decision Trees, Artificial Neural Networks, Support Vector, Instance Based and Statistical Language Model Based Classifiers are widely preferred in text classification applications [70]. In a literature review on sentiment analysis, Metha [71] stated that ML methods such as SVM, NB, and neural networks have the highest accuracy, considering them as fundamental learning methods. Numerous studies in the literature suggest the superior performance of ML, including DL, through proposed models and comparative analyses [8,10,13,23, 72-78]. Furthermore, it is often mentioned that combining multiple classifiers generally yields better experimental results than using a single classifier. However, in some cases, dictionary-based methods are also noted to be highly effective [79, 80]. When used together in a hybrid approach, ML and dictionary-based approaches can strengthen each other's weaknesses, allowing the development of higher-performing models. It is noted that hybrid models, with appropriate architecture and precise hyperparameter selection, can outperform all models [27,41,45,79,81-88].

Dey and Das [89], in a sentiment analysis study based on an approach proposing a modified TF-IDF term weighting method, achieved performance in the range of 62.1% to 89.2% on different datasets. Yoo and Nam [90] conducted a sentiment analysis study using machine learning algorithms and an electronic dictionary in Korean. In this study, they achieved a performance in the range of 76-80% with a hybrid approach on datasets of restaurants, computers, cinema, travel and clothing. Erşahin et al. [91] obtained performances of 73%, 86.32%, and 91.96% on three different datasets consisting of tweets about hotels and cinemas, using three different classification algorithms (SVM, NB, and J48) and the dictionary-based approach in their proposed hybrid models. There are comparative studies in the literature on dictionary-based, ML, and hybrid approaches used for sentiment analysis. In one study compiling 68 analyses, the highest performances

were found to be 88.85% for dictionary-based studies, 98.29% for ML studies, and 91.96% for hybrid studies [62]. In another study, the performances of recent works using DL, ML, and hybrid approaches were reported to range from 74% to 91% [92]. Mahmood et al. [93] obtained performances of 86% and 90% in their hybrid study using Naïve Bayes and SVM as machine learning algorithms and Wordnet as the general dictionary.

The representation of emotion is considered one of the fundamental challenges in sentiment analysis, and it is noted that this area is still in its infancy [13]. In the classification of texts, weighted terms are used to determine the emotional direction of the text. Selecting features with high distinctiveness from weighted terms contributes to solving the high dimensionality problem in matrices created for analysis, enhancing model performance.

Due to the abundance of jargon meanings in tweets and the composition of very short sentences, feature selection becomes a critical process. The number of words and length of the text are elements that affect a document's score as determined by term weighting methods. It is observed that the score values of words in a dataset consisting of tweets are generally lower than the word scores in other datasets [12]. The high word count in customer reviews is used as a factor that increases the reliability of the review, assuming that a higher word count implies more information about the product [5]. In this regard, X differs from reviews containing evaluations directly related to a product or service. The sparsity of words with jargon meanings and short sentences in tweets can result in a lack of distinctive features in the text. In the literature on feature selection, basic techniques such as count-based methods such as bag-of-words, simple statistical values, term frequency (TF), inverse document frequency (IDF), co-occurrence of terms, n-gram statistics are widely used; PATricia (PAT) tree, SWN word groups, graph-based methods, the popular deep learning technique Bidirectional Encoder Representations from Transformers (BERT) architecture, and BERTweet built on top of it [13, 88, 94-99]. The count vector (CV), based on word frequency within sentences, allows for successful feature selection. Additionally, statistical methods such as the term frequency-information gain method (TF-IGM), which is suitable for multi-classification and considers class frequencies of terms, and the term frequency-inverse document frequency-inverse corpus frequency (TF-IDF-ICF), as well as the term frequency-inverse document frequency-inverse cluster size document frequency (TF-IDF-ICSDF), have been successfully employed for feature selection [100].

Alexandrovna et al. [101] stated that performance could be enhanced through careful and efficient feature weighting, and they achieved an improvement of 4%-5% in accuracy using a methodology that reduces the feature dimension. Bandhakavi [102] utilized a domain-specific dictionary and the unigram mixture model (UMM) to identify terms that best represent the text. Sarayna [103] employed the TF-IDF method on a dataset consisting of tweets, while Kaur [104] utilized the TF-IDF method in conjunction with n-grams. Alshehri and Algarni [105] utilized term frequency (TF) and term discrimination ability (TDA), which groups selected features based on their distinctiveness and weights them according to their contribution to each group. Sharma and Kumar [106] employed a multi-feature-based concept ranking algorithm that uses statistical, semantic, and scientifically named entity properties of terms.

Although there are few sentiment analysis studies in the food industry, it is observed that models and methods have been developed with satisfactory performance. Additionally, it is emphasized that there is a continued need for especially domain-specific studies in the field of sentiment analysis [22-24].

Hingle et al. [107] explored eating habits from X data related to the food industry, while Park et al. [108] investigated perceptions of Chinese, Japanese, Korean, and Thai restaurants. Mishra and Singh [109] focused on waste categories in the meat supply chain, and Singh et al. [110] examined dissatisfaction with beef products. El-Khchine et al. [111] conducted a study on the main areas of interest related to chicken products and proposed a model

Zahoor et al. [112] conducted studies on sentiment analysis and categorization of reviews about restaurants in Karachi, focusing on taste, ambiance, service, and value. Alamoudi and Alghamdi [113] performed sentiment classifications based on food, service, ambiance, and price as target features. In a study by Liapakis et al. [114], they analyzed customer reviews for the food and beverage industry for a one-month period in 2018. For the analysis, they identified five features: food quality, customer service, company image, price, and product quantity.

In their literature review examining sentiment analysis studies in the fast food sector, Adak et al. [3] noted that most studies in this field commonly employ lexicon-based and ML methods. They highlighted a limited number of studies applying DL techniques and mentioned that DL techniques exhibit better performance in other sectors. The authors also mentioned that 77% of models in this sector need help in interpretation because of their nature. Ahmed

et al. [115] developed a model that considers implicit meanings to improve classification accuracy. The model, which utilizes domain-specific meanings and target features together, achieved a success rate of 89% when applied to a dataset related to restaurants. Aktaş et al. [116] achieved a performance of 86% in their ML study focused on the food and beverage sector.

This study developed classification models based on the methods and tools used for sentiment analysis and text classification. Methods for feature selection and dimensionality reduction were proposed to improve the performance of text classification models. In this context, lexicons called basic and boosted for domain-specific sentiment and product-service systems were prepared. For all tweets, the polarity scores of words in the lexicons were calculated according to three term weighting methods. These scores were placed in document-term matrix (DTM) cells, and the classes of tweets were determined using DTM with six classification algorithms. The classes determined by the algorithms and lexicons were compared with the actual classes of text to measure performance.

The proposed method of the study that reduces the feature dimension has been employed to transform the basic lexicons into the boosted lexicons. The transformation process that enables the aggregation of sentiment scores of words appearing as separate variables in the dataset and the lexicons is detailed in Section 3. The applied feature reduction method has resulted in an improvement in model performance.

3. METHOD AND MODEL PROPOSAL

Figure 1 depicts the stages of the strategy, which was developed by combining a lexicon-based approach with ML algorithms for sentiment and text classification.

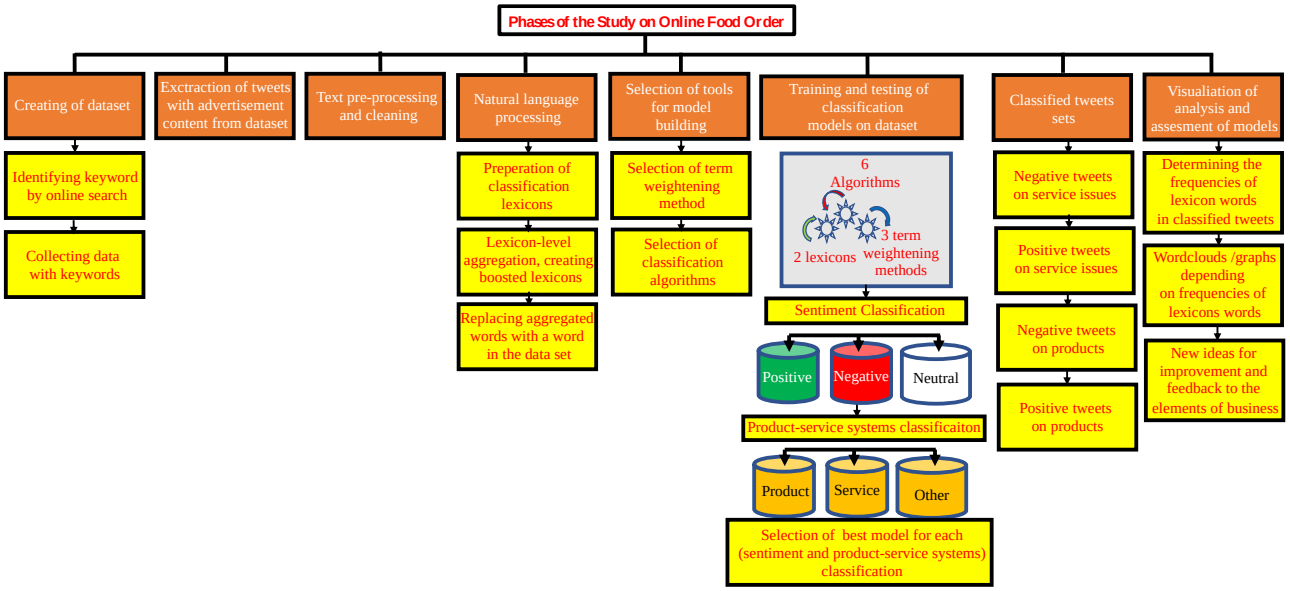


Figure 1. Phases of the study for classifications of tweets about online food ordering

The phases followed are described below.

3.1. The dataset and preprocessing

The data for the analysis was extracted from X tweets on food ordering using the paid Twitonomy application with the Turkish keywords yemek sipariş (food order), yemeksepeti sipariş, döner sipariş (döner kebab order), lahmacun sipariş (thin Turkish pizza order), hamburger sipariş (hamburger order), and pide sipariş (pita order). The tweets are from January 1 to December 31, 2020. The collected tweets are public. Using a specific advertising lexicon, advertising-related tweets were filtered out of the collected dataset. Duplicate tweets, as well as tweets consisting of two or fewer words, were removed. In the dataset, punctuation marks, shapes, and symbols were removed from the tweet texts as part of text preprocessing.

Additionally, the stop words from a list of 175 words prepared for the study were cleaned from the tweet texts (Annex-1). Finally, since the keywords used to collect tweets from X are present in all or most of the tweets in the dataset, which reduces their distinctiveness for the classifications, they were removed from the dataset. The numerical values, such as 8/10, 9/10, etc., found in tweet texts have yet to be cleaned from the text, considering X users' jargon and language usage in the domain. Instead, they have been included in the lexicons as variables (features).

In text classification, the performance measurement of the created models is determined by comparing the actual class of the text (the class label given by the evaluator) with the classification made by the model. The dataset is randomly divided into two groups: 21,039 tweets and 14,389 tweets. The first is used for model creation, training, and testing, and the classification model with the highest performance is identified. The results are presented using the model on the second dataset to show what kind of information can be accessed about the enterprises' fields of activity in the sector and to which areas they can direct their improvement and development activities.

Six evaluators labeled the dataset for the purpose of measuring the models' performance. The evaluators labeled the datasets as product, service, and other for target aspects and positive, negative, and neutral for sentiment classification. Manual labeling of tweets is time-consuming work and requires understanding the purpose of the analysis to do it correctly [117]. The evaluation of a tweet based on subjective evaluation by people with different ways of thinking [3, 8] can produce different labeling results. Due to the difficulties in interpreting sentences in the food sector and the vagueness and ambiguity of colloquial language, evaluators were informed about the labeling process verbally and through a briefing note. The briefing notes include the purpose of the study, the dataset characteristics, the target features of the classification, and Table 1[1]. Table 1 lists the subheadings extracted from food industry classification studies [107-114].

Tablo 1. Class labeling subheadings for product-service systems

Product and Product Quality	Service and Service Quality		
Taste and flavor	Terms of service	Attention, interest and helpfulness of the personnel	Advertising comments about the company/business
Healthy alternatives	Personnel and working conditions of them	Cleanness, hygiene	Online service
Product (menu) and product variety	Consistency	Industry-related advertisements	Discounts
Freshness	Packet	Service speed and duration, weather conditions (courier working conditions)	Promotion applications
Food safety	Price	Courier	Validity of meal cards
Recommended temperature of product	portion adequacy	Customer service staff	Speed of response to complaints, customer service support
Cooking aroma, food smell	foreign body presence in the food	Presentation of food	---
All topics not included in the Table have been labeled as “Other” category.			

The dataset of the study does not have an extreme imbalance [118]. Therefore, no balancing operation has been applied between the classes.To build the model, a first-group dataset consisting of 21,039 tweets was utilized, and classification lexicons and algorithms prepared within the scope of the study were employed for ternary (positive-negative-neutral and product-service-other) and binary (positive-negative and product-service) classifications. For the ternary classification model, all 21,039 tweets were used. In contrast, tweets with the "neutral" and "other" class labels were excluded from the binary classification model, leaving the remaining tweets for analysis.

3.2. The domain-based classification lexicons

Within the scope of the study, two domain-specific lexicons were prepared one for sentiment and the other for product-service systems. These lexicons are named the "basic sentiment lexicon" and the "basic product-service systems lexicon." As an example, he rankings of the first 20 words in the basic lexicon based on their frequencies in the dataset, along with class frequencies, are provided in Table 2.

Despite the progress in language models in recent years and the success of machine learning algorithms and dictionary-based models, these methodologies fail to capture the meanings of words accurately, and these meanings vary depending on the domain they belong to. In particular, the hybrid use of domain-specific dictionaries with one of these methodologies may provide a solution to this problem [11, 13].

A hybrid method of seed word list and corpus-based approach was employed in preparing the lexicons in the following steps: (i) A seed word list was created for the

[1] The details of evaluator briefing notes and labelling the tweet texts by evaluators are in the doctoral thesis which is in the ph.D. thesis of Y. Güneş which is supervised by M. Arıkan, “Twitter (X) Analytics for the Service Sector: An Application on Ordering Meal to Home and Offices”, PhD Thesis, Gazi University, Graduate School of Natural and Applied Sciences, 2024.

domain by examining various social media platforms and business web pages in the industry related to food orders (Annex-2). (ii) The seed word list was expanded with new words using synonym and antonym dictionaries. (iii) The jargon (such as biker, courier, basket maker), slang words, and commonly misused and misspelled words were added to the expanded word list, and the lexicons were formed. (iv) After determining the frequencies of lexicons' words in the dataset, words with zero frequency and the top two words ("food" and "order") with the highest frequency were excluded from the lexicons. However, the words "order" and "food" in a word group such as "order note" continue to be included in the lexicon. As a result, a basic sentiment lexicon consisting of 769 words and a basic product-service systems lexicon consisting of 684 words were obtained [17, 119-124].

Aggregating the sentiment scores of words at different levels is a general method used in text classification problems. In this study, the aggregation method was used in the basic dictionaries to transform them into new classification dictionaries called the boosted dictionary, as described below. At this stage, MS Excel-365's synonyms and antonyms dictionary are deployed, and the aggregation process is performed by evaluating the dictionary information and context-semantic information together [125]. The jargon, slang, or low-value words that may be ineffective in classifications have increased the impact of the calculations through this aggregation at the lexicon level. (i) The words expressed in speech or incorrectly transcribed in the dataset have been grouped based on the correct spelling of the word. (ii) Synonyms, close meanings, or words that are considered to be used in the same sense in the text are groped together as Turkish words alan (field), bölge (region), etraf (around), konum (location), civar (vicinity, nearby), muhit (surroundings, environment), sokak (street), and semt (neighborhood). (iii) Words with the same root but different affixes (due to affixes, sound dropping, softening of hard letters, or vice versa) are grouped as a single word group. (iv) A word from each group is selected to represent the word group. The words other than the representative word in the word group are discarded from the basic lexicon. Simultaneously, the representative word is substituted for the other words in the group in the dataset's tweets. Thus, lexicons with fewer words, which are called boosted lexicons in the study, were obtained.

Tablo 2. Basic sentiment and basic product-service systems lexicon words with the frequencies-20 words

Basic Sentiment Lexicon			Basic Product-Service Systems Lexicon		
Lexicon Words	Dataset Frequencies	Class Dataset Frequencies	Lexicon Words	Dataset Frequencies	Class Dataset Frequencies
yok (unavailable)	1840	845	yorum (comment)	4711	1774
güven (trust)	900	11	saat (hour)	2159	1214
arkadaş (mate)	847	155	restoran (restaurant)	1436	1071
gelm (don't come)	846	611	lahmacun (meat filling)	1222	148
öner (suggest)	843	168	kurye (courier)	1148	995
sev (love)	696	264	burger (burger)	1078	196
istiyor (wants)	604	240	online (online)	879	749
iptal (cancel)	540	495	öner (suggest)	843	170
destek (support)	533	185	zaman (time)	827	250
yemek yok (no food)	475	44	telefon (telephone)	816	595
güzel (beautiful)	465	264	döner (döner kebap)	780	148
kara kara düşün (brood over)	432	15	firma (company)	772	545
zorunda kal (to be forced to)	431	65	paket (package)	768	413
verem (can't give)	358	281	hamburger (hamburger)	692	135
çıkar (self interest)	355	234	pizza (pizza)	672	139
isted (asked)	352	62	adres (address)	606	249
gerçek (real)	349	31	dk (minute)	601	321
getirm (not bring)	343	278	servis (service)	575	351
kazan (earn)	332	41	şimdi (now)	570	192
kalm (no left)	275	121	adam (men)	596	293
The red colors shows the negative class words, the others shows positive class words in the basic sentiment lexicon.			The blue colors shows the product class words, the others shows service class words in the basic product-service systems lexicon.		

As a result, the basic sentiment lexicon's word count decreased by 105, resulting in a boosted sentiment lexicon of 664 words (Annex-3 and Annex-4); the basic product-service systems lexicon's word count decreased by 93, resulting in a boosted product-service systems lexicon of 591 words (Annex-5 and Annex-6). One limitation of this dimension reduction method is that it introduces an additional process of grouping words before analysis and replacing the word representing the group in the texts.

The boosted lexicon structure aims to increase the frequency and weighted values of the words as variables and reduce the matrix size by lowering the number of words (the dimension reduction process). While high-frequency words are more successful for classification in

classic feature selection approaches [88], low-frequency words have a negligible influence. The efficacy is increased by combining low-frequency terms with high-frequency words through the boosted lexicon structure.

Table 3 shows an example of how the proposed word representation in feature selection can lead to a rise in the number of times a word is used in term weighting formulas for boosted lexicons.

Tablo 3. Sample of word representation for boosted lexicon

Process order	Sample of Word Representation for Boosted Lexicon		
1.	Words in basic sentiment lexicon	Group of words (synonyms) characterized by a representative word	Representative word in boosted sentiment lexicon for the group of words
	Ödül (award, prize)	ödül, armağan, hediye	hediye
	Armağan (gift)		
	Hediye (present)		

2.	Tweets	Pre-processed Tweets	Replacement process of Tweet-1 for boosted sentiment lexicon
	Tweet-1	"dün gece etmeden aç aç yattığım kendimi *ödül* amaçlı kahvaltı pizza sipariş ettim"	"dün gece etmeden aç aç yattığım kendimi *hediye* amaçlı kahvaltı pizza sipariş ettim"
	Tweet-2	"kardeşim tatil *hediyesi* olarak etmiş aşko benzememeliydin"	"kardeşim tatil *hediyesi* olarak etmiş aşko benzememeliydin" (Because of the term of "hediye" is representing term for group of terms, the term in the Tweet-2 does not need to be replaced.)

3.	Frequency of terms in dataset based on basic sentiment lexicon	Frequency of terms in dataset based on boosted sentiment lexicon
	Ödül (Award, prize):47	---
	Armağan (gift):2	---
	Hediye (present):160	Hediye (present):209

3.3. Term weightening methods and classification algorithms

The formulas for three-term weighting methods can be found in Table 4. These are the count vector (CV), the term frequency (TF), and the term frequency-inverse document frequency-inverse class frequency (TF/IDF/ICF). Generally, tweet text score values are lower than in other datasets [12]. The method applied in the form of aggregating words involves transferring the weighted scores of words from the basic lexicon to the boosted lexicon structure by increasing them. The calculation of the weighted values used for the boosted lexicon structure of the words given in Table 3 is illustrated with an example in Table 4.

The Logistic Regression (LR), K-Nearest Neighbour (K-NN), Non-Linear Supportive Vector Machine (NL-SVM), Multi-Layer Perceptive Classification (MLPC), Gradient Boosting Machine (GBM), and eXtreme Gradient Boosting (XGB) algorithms were utilized for the creation of models.

When training the models, a 10-layer cross-validation method was applied, taking into account the amount of the data set in order to avoid the bias effect in the data set, and hyper-parameter tuning was performed to determine the best performance of the models [126].

Tablo 4. Term weighting formulas and application of them on basic and boosted lexicons' words with a sample

Term Weightening Methods	Words	Basic Lexicon	Boosted Lexicon
CV $W_{CV}(t_i) = TF(t_i, d_j)$	Hediye (present)	160	209
	Armağan (gift)	2	---
	Ödül (award, prize)	47	---
TF $W_{TF}(t_i) = TF(t_i, d_j)/T_j$	Hediye (present)	0.00052	0.00068
	Armağan (gift)	0.0000065	---
	Ödül (award, prize)	0.00015	---
TF-IDF-ICF $W_{TF-IDF-ICF}(t_i) = (TF(t_i, d_j)/T_j) * \{1 + \log(\frac{D}{d(t_i)})\} * \{1 + \log(\frac{C}{c(t_i)})\}$	Hediye (present)	0.0016	0.0022
	Armağan (gift)	0.000039	---
	Ödül (award, prize)	0.000552	---
$TF(t_i, d_j)$: the frequency of term i in document j.			
T_j: the total number of words in the collection/dataset.			
$d(t_i)$: the number of documents in which the term t_i occurs.			
$c(t_i)$: the number of classes in which the term t_i occurs.			
D: number of documents in the collection/dataset. C: the number of classes in the collection/dataset.			
Values for the terms used in the formula above: T_j=305 672; D=21 039; C=3. The values $d(t_i)$ =147 and $c(t_i)$ =3 for "hediye (present)"; $d(t_i)$ =2 and $c(t_i)$ =1 for "armağan (gift)"; $d(t_i)$ =44 and $c(t_i)$=3 for "ödül (award-prize)" are used in the formulas for basic lexicon. The values, $d(t_i)$ =193 and $c(t_i)$ =3 for "hediye (present)" are used in the formulas for boosted lexicon. This calculation is done for the ternary classification			

3.4. The evaluation of model performance

Measurement criteria such as precision, recall, accuracy, and F1 score values can be used for the performance testing of classification models. The calculations of these measurement criteria are based on a confusion matrix comparing the actual class of the data with the classes predicted by the models. The precision value indicates the overall success of the model in classification and is often a useful measure of the performance of datasets with a balanced class distribution. The F1 value, a more robust measurement tool in unbalanced datasets, can provide results by balancing precision and recall values.

However, the F1 calculation also does not consider the class's proportion of observations (samples). Therefore, the weighted average F1 value, which considers the distributional proportions of the classes, is the preferred method for imbalanced datasets. In the study, the weighted

F1 score was employed for performance comparisons, with class ratios used in computations, taking into account the modest imbalance (20% and below) in the positive and product classes in the dataset.

$$\text{(Weighted F1 score)} = \frac{Number\ of\ samples\ in\ the\ class}{Number\ of\ samples\ in\ the\ corpus} * \frac{2 * Precision * Recall}{Precision + Recall}$$

To create models for classifying sentiment and product-service systems, we looked at how well models built with basic and boosted lexicons, three-term weighting methods, six classification algorithms, and binary and ternary classifications worked. MS Excel's data analysis features, Python programming language, and its libraries were used for the application. The results of the analysis are presented in Section 4.

4. RESULTS AND DISCUSSION

For identifying the best classification model, 144 models were created, 72 for product-service system classification and 72 for sentiment classification. These models included ternary and binary classifications, basic and enhanced sentiment dictionaries, three-term weighting approaches, and six classification algorithms.

Considering the moderate imbalances in the dataset, the models established were initially evaluated based on the weighted average F1 scores to determine their performances. Fine-tuning, which allows the model to capture the best values of the parameters and learn better from the dataset, was used to improve the performance of the models. All comparisons between the models in the study were applied after fine-tuning [9].

4.1. Sentiment Classification

The performance of the models run for sentiment classification is shown in Figure 2. The findings related to the emotion models depicted in Figure 2 are as follows: (i) All binary classifications are more successful than ternary classifications. (ii) It has been observed that performance can be enhanced through hyperparameter tuning in the majority of models (performance before hyperparameter tuning is shown in black, and performance after tuning is shown in gray). (iii) Two-class models built with the boosted sentiment lexicon (average shown with a red dashed line) demonstrated, on average, a performance superiority of 1.36% over two-class models built with the basic lexicon (average shown with a green dashed line). This superiority was observed in ternary models at a rate of 0.26%. (iv) The model created using the boosted sentiment lexicon, binary classification, TF-IDF-ICF weighting method, and K-NN algorithm achieved the highest performance for sentiment classification with a rate of 85.217%.

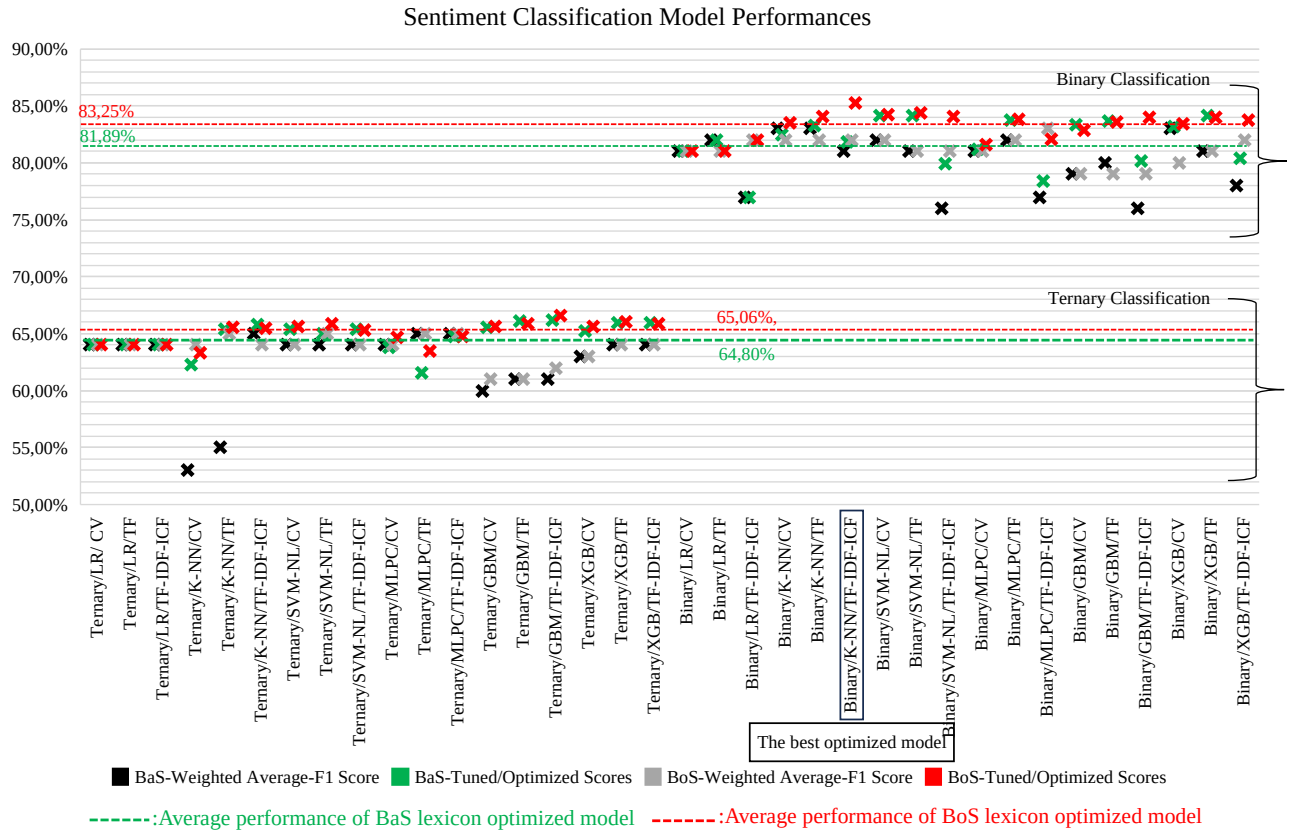


Figure 2. Sentiment classification model tuned performances of basic (BaS) and boosted (BoS) lexicons

The initial and optimized performances, along with the parameters of the proposed sentiment classification model, are presented in Table 5. During the fine-tuning process, various parameters—such as learning rate, maximum depth, number of estimators, minimum sample split, and number of neighbors in different algorithms—were evaluated using a random search approach. The optimal parameter combinations were then determined to enhance the performance of the optimized models [9].

Tablo 5. Proposed model final report for binary sentiment classification

Proposed Model for Binary Sentiment Classification			
Algorithm	Term Weightening Method	Lexicon for Classification	Initial Model
K-NN	TF-IDF-ICF	Boosted Sentiment Lexicon	make-pipeline (StandartScaler( ) KNeigbors Classifier())
Initial Performances			
Training Score	Test Score	10-K-Fold Score	
0.8673	0.8362	0.7667	
Final Performances			
	Precision	Recall	F1 Score
Macro-Avg	0.77	0.68	0.71
Weighted-Avg	0.82	0.84	0.82
Optimized Model Parameters and Performances			
Training Score	Parameters Tested in Initial Model for Optimisation	Final Model After Fitting the Parameters	Optimized Test Score
0.8416	{"n_neighbors": np.arange(1,50}	KNeighbors Classifier(11)	0.8522

4.2. Product-Service Systems Classification

The performance of the models run for product-service classification is shown in Figure 3. The findings related to the product-service systems models depicted in Figure 3 are as follows: (i) All binary classifications are more successful than ternary classifications. (ii) Similar to sentiment classification models, it has been observed that performance in the majority of product-service systems models can be enhanced through hyperparameter tuning. (iii) Two-class models constructed with the boosted product-service systems lexicon (average shown with a red dashed line) demonstrated, on average, a performance superiority of 1.12% over two-class models constructed with the basic lexicon (average shown with a green dashed line). This superiority is observed in ternary models at a rate of 3.95%. (iv) The model created using the boosted product-service systems lexicon, binary classification, CV weighting method, and GBM algorithm achieved the highest performance for product-service systems classification with a rate of 88%.

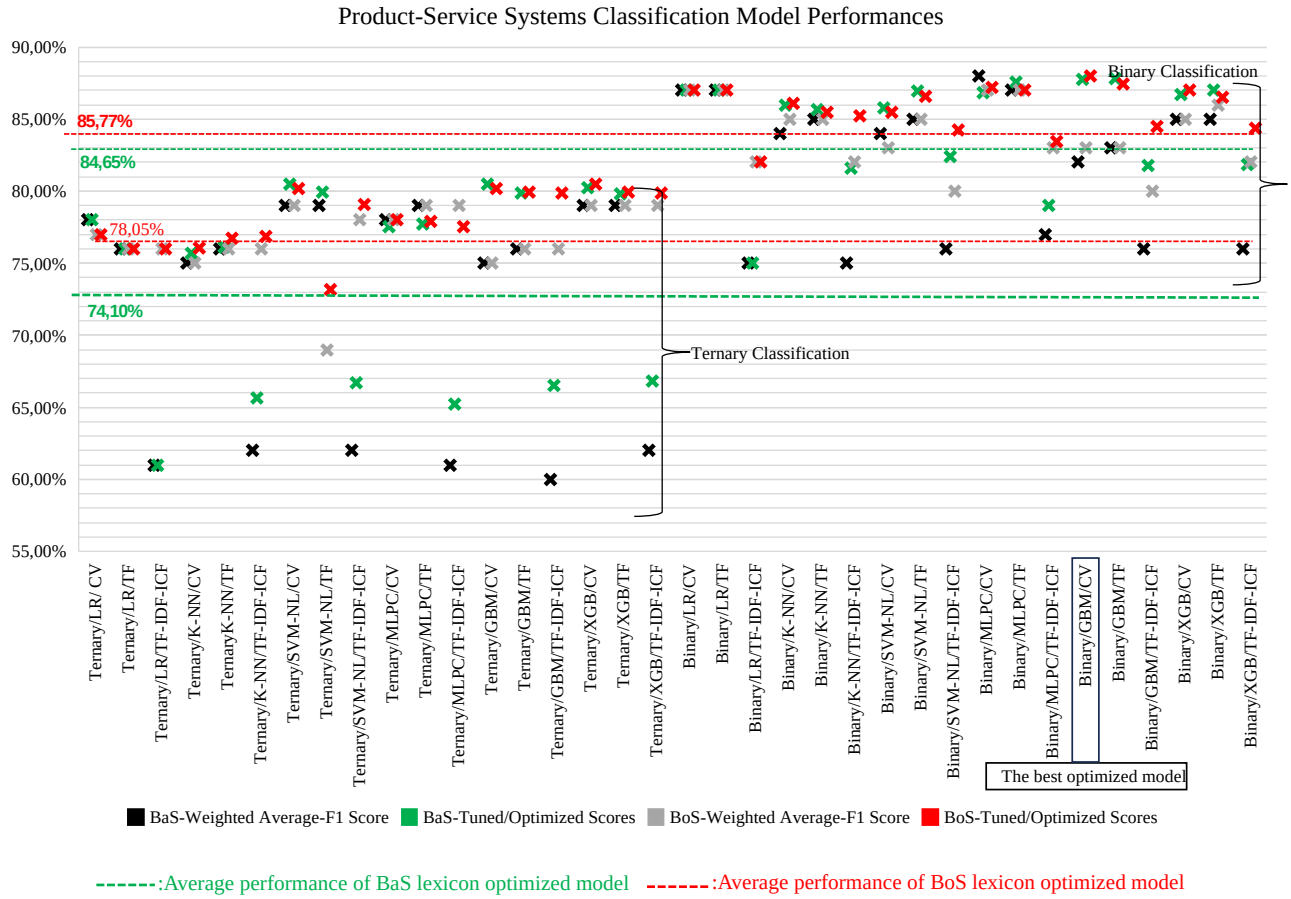


Figure 3. Product-service systems classification model tuned performances of basic (BaS) and boosted (BoS) lexicons

The initial and optimized performances and parameters of the proposed model for sentiment classification is shown in Table 6.

Tablo 6. Proposed model final report for binary product-service systems classification

Proposed Model for Binary Product-Service Systems Classification			
Algorithm	Term Weightening Method	Lexicon for Classification	Initial Model
GBM	CV	Boosted Product-Service Systems Lexicon	make_pipeline (Standart Scaler (), GradientBoostig Classifier())
Initial Performances			
Training Score	Test Score	10-K-Fold Score	
0.8769	0.8521	0.8477	
Final Performances			
	Precision	Recall	F1 Score
Macro-Avg	0.82	0.63	0.67
Weighted-Avg	0.84	0.85	0.82
Optimized Model Parameters and Performances			
Training Score	Parameters Tested in Initial Model for Optimisation	Final Model After Fitting the Parameters	Optimized Test Score
0.8829	{“max_depth”:range(3,5),”n_estimators”:[100,500,1000],”min_samples_split:[2,10]”:}	GradientBoostingClassifier(max_depth=4,n_estimators=1000,min_samples_split=2)	0.8801

The best results from comparing 144 models created through the combination of different algorithms, term weighting methods, binary-ternary classifications, and the alignment of sentiment and product-service systems classifications are shown in Table 7. When the weighting methods and algorithms of the constructed models are examined in terms of average values across all classifications, (i) the best results for each of the three-term weighting methods were obtained with the GBM algorithm, and (ii) the TF method with the highest average of 79.3% was observed.

In addition to the model with the best performance highlighted in italic in Table 7, other binary classification models also exhibit satisfactory performance and can be used as classification models.

Table 7. Optimized/tuned classification results

Sentiment Classification Optimized/Tuned Test Scores			
Number of Classes	Term Weightening Method	Basic Sentiment Lexicon (769 words)	Boosted Sentiment Lexicon (664 words)
Three classes	CV	GBM: 65,53%	GBM:65,65%
	TF	GBM: 66,11%	XGB: 66,01%
	TF-IDF-ICF	GBM: 66,18%	GBM: 66,62%
Two classes	CV	NL-SVM: 84,15%	NL-SVM: 84,19%
	TF	NL-SVM/XGB: 84,15%	NL-SVM: 84,39%
	TF-IDF-ICF	KNN: 81,79%	*KNN: 85,22%*
Product-Service Systems Classification Optimized/Tuned Test Scores			
Number of Classes	Term Weightening Method	Basic Product-Service Systems Lexicon (684 words)	Boosted Product-Service Systems Lexicon (591 words)
Three classes	CV	GBM: 80,52%	XGB: 80,49%
	TF	NL-SVM: 79,92%	XGB: 79,94%
	TF-IDF-ICF	XGB: 66,82%	GBM/XGB: 79,9%
Two classes	CV	GBM: 87,77%	*GBM: 88%*
	TF	GBM: 87,82%	GBM: 87,43%
	TF-IDF-ICF	NL-SVM: 82,4%	KNN: 85,22%

It was seen that both the basic and boosted lexicons made for the study could be used for classifications. The boosted lexicon makes the model perform better than it did with the basic lexicon. In the literature, performance ranges in text classification using dictionary-based, machine learning, and hybrid approaches vary between 62% and 98% [58, 86-88]. Considering the unique challenges posed by tweet texts compared to other types of texts, the achieved performance of 85.22% and 88% in this study is considered superior and competitive compared to many studies in the literature.

4.3. Implementation of the Proposed Classification Models on the Second Group Dataset

After using the first group dataset, the suggested classification models are as follows: for sentiment classification, models made up of the boosted lexicon, binary classification, TF-IDF-ICF weighting method, and K-NN algorithm; and for product-service classification, models made up of the boosted lexicon, binary

classification, CV weighting method, and GBM algorithm. The type of insights that can be derived using the recommended models in relation to the industry is demonstrated with the second group dataset, consisting of 14,389 tweets. For this purpose, the second dataset has been classified using the recommended two models, and the distributions resulting from the classification are shown in Figure 4. Based on the distribution ratios, it is observed that over 80% of the tweets in the second dataset consist of negative opinions.

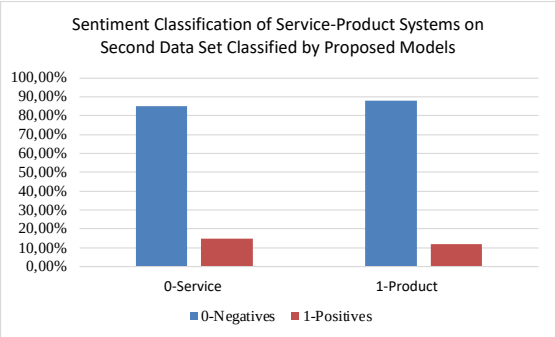


Figure 4. Proportion of second dataset class after classified by optimized models

At the end of the classification process of the second dataset, word frequencies in the tweets belonging to each class were determined, and word clouds were created based on these frequencies. The prominent topics in the word cloud can provide guiding results for businesses in the sector concerning improvement and development.

The prominent topics among users regarding the positive and negative aspects of the service class as a result of the classifications are shown in Figure 5. Accordingly, the most critical complaint topics that need improvement in the service aspect are observed to be restaurants (restoran), couriers (kurye), advertisements (reklam), cancellation processes (iptal), time-duration (zaman-süre), payment-related transactions (ödeme), delivery (teslim), and non-delivered (gelm). It is observed that users' positive opinions about the service are concentrated around winning (kazan), discounts (indirim), loyalty programs (joker), recommendations (öner), and friend (arkadaş), like (sev), plus (artı), want (istiyor), support (destek) topics.

Figure 5. Negative and positive highlights on service issues after classified second dataset

The prominent topics among users regarding the positive and negative aspects of the product class are shown in Figure 6. Accordingly, users express their negative opinions and complaints about the product mostly using words such as pizza (pizza), product (ürün), chicken (tavuk), cold (soğuk), spicy (acı), onion (soğan), salad (salata), missing (eksik), fatty (yağlı), dough (hamur), wrap (dürüm), minced meat (köfte), portion (porsiyon), stomach (mide), crispy (çıtır), awful (rezil). On the positive side, opinions about this matter are concentrated around words such as recommendation (öner), like (sev), chicken (tavuk), cold (soğuk), large size (büyük boy), delicious (lezzetli), instant (anında), product (ürün), sauce (sos), ketchup (ketçap), mayonnaise (mayonez), fried (kızartm), nutrition (besin), garnish (garnitür), vegan (vegan).

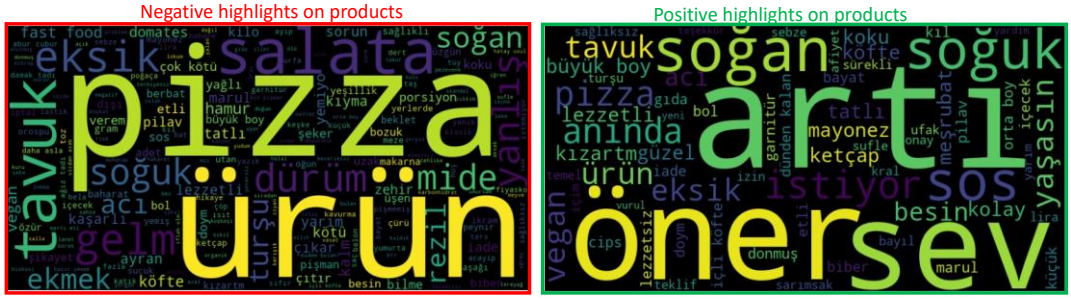


Figure 6. Negative and positive highlights on product after classified second dataset

There are many different models in text classification and sentiment analysis. These models continue to be developed, either on their own or in different combinations.One of the important problems in these models, including advanced language models, is the correct determination of the meaning of the word in the text in which it is used. Failure to correctly determine jargon, implied, and contextual meanings reduces model performance. Another problem of text classification for language models are high dimensionality, and extraction of keyword from the text efficiently [13]. In order to solve this problems, domain-based studies can provide important contributions to the domain. Domain-based lexicon helps to the model in focusing essential and meaningful words [127]. The words in these lexicons will be used only for classification related to the subject, they will not have high dimensionality as in general dictionaries, and words with low effectiveness in lexicons for classification can be excluded from the lexicons to decrease the dimensionalty. The use of a food-specific dictionary and classification algorithms together has achieved a level of success that can compete with state of the art models in this field. Further work in this area could form the basis for future advanced models in sentiment analysis and text classification.

4. CONCLUSION

In this study, an online food ordering classification model has been developed using a lexicon-based approach and classification algorithms in a hybrid method. A total of 144 model comparisons were conducted to form a model for sentiment and product-service system classification. The study will contribute to fill the gap in domain-based text classification and helps to industry to analyze with a robust text classification and sentiment analysis model in food domain. It also will encourage academicians to work on new classification models in other unstudied domains such as the clothing industry, cargo sector, which has potential for development in sentiment classification.

The study's contribution is the proposal of the boosted lexicons for use in sentiment and product-service classifications. The boosted lexicon structure not only yields better results compared to the basic lexicon but also reduces the complexity of the problem due to its smaller size. It has been observed that the applied method improves performance in both sentiment and product-service system classifications. The suggested approach and classification models obtained classification performance of 85% or higher, surpassing several studies on sentiment analysis and text classification found in the existing literature.

Within the scope of the study, four dictionaries were prepared specifically for the food ordering domain, including one basic and one boosted for both sentiment and product-service system classifications. It was observed that the boosted lexicon outperformed the basic lexicon, binary classifications performed better than ternary classifications, and product-service system classifications were better than sentiment classifications. Among the term weighting methods used in the models, TF was found to have the best performance average. Among the algorithms, GBM exhibited the highest performance. The recommended classification models, developed domain lexicons, and sentiment analysis conducted on customer feedback in the context of online food orders enable the measurement of customer satisfaction based on product and service target features. The results provide an opportunity to identify areas needing improvement that can potentially shape the industry.

In the following periods, the domain lexicons developed within the scope of the study can be developed and used in new studies specific to the field of food. The boosted lexicon structure, proposed as a solution to the dimensionality reduction problem, which is a significant issue in text classification problems, can be applied to classifications of other text types with a higher word count in text compared to tweets. Thus, the problem of high dimensionality in text classification issues is addressed, and performance comparisons with other models and methods can be made.

Limitations of the Study

The boosted lexicons created with the dimensionality reduction method have improved the classification performance. However, the recommended method also has some limitations. The method requires additional processes before the analysis operations. Words within word groups in the dataset should be replaced with representative words. The grouping of words for the method has been done considering synonyms and meaning similarities arising from jargon, domain-specific uses, and figurative expressions. Executing this method manually can be a time-consuming process. However, it contributes to producing useful domain lexicons for text classifications, considering the natural usage of language.

REFERENCES

[1] S. Dixon, Number of Twitter Users Worldwide from 2019 to 2024, https://www.statista.com/statistics/303681/twitter-users-world wide/. 25 .04.2023.

[2] Y. Güneş and M. Arıkan, "Exploring Twitter dataset content by descriptive analysis: An application on online food ordering", *International Journal of Informatics Technologies (JIT)*, 16(2), 119-133, 2023.

[3] Adak A, Pradhan B and Shukla N. "Sentiment analysis of customer reviews of food delivery services using deep learning and explainable artificial intelligence", Systematic review. Foods, 11(10), 1500, 2022.

[4] L.R. Krosuri and R.S. Aravapalli, "Feature level fine grained sentiment analysis using boosted long short-term memory with improvised local search whale optimization", PeerJ Computer Science, (9)e, 1336, 2023.

[5] M.V. Gopalachari, S. Gupta and S. Rakesh, et al, "Aspect-based sentiment analysis on multi-domain reviews through word embedding", *Journal of Intelligent Systems*, 32(1), 2023.

[6] S. Atan and Y. Çınar, "Relations between financial news and market capitalizations of companies in BIST30 Index: Text mining and sentiment analysis methods", *Ankara University Journal of Social Sciences*, 74(1), 1-34, 2019.

[7] T. Shaik, X. Tao and Dann C, et al, "Sentiment analysis and opinion mining on educational data: A survey", *Natural Language Processing Journal*, 2, 1-11, 2023.

[8] W. Medhat, A. Hassan and H. Korashy, "Sentiment analysis algorithms and applications: A survey", *Ain Shams Engineering Journal*, (5), 1093–1113, 2014.

[9] G. K. Shinde. V. N. Lokhande, R. T. Kalyane, V. B. Gore and U. M. Raut. "Sentiment Analysis Using Hybrid Approach". *International Journal for Research in Applied Science and Engineering Technology*, (9), 282-285, 2021.

[10] R. Vatambeti, S.V. Mantena and K.V.D. Kiran, et al, "Twitter sentiment analysis on online food services based on elephant herd optimization with hybrid deep learning technique", Cluster Computing, 2023.

[11] Trust P, Minghim R. "A Study on Text Classification in the Age of Large Language Models". Machine Learning and Knowledge Extraction. 6(4), 2688-2721, 2024.

[12] DMEDM Hussein, "A survey on sentiment analysis challenges", *Journal of King Saud University, Engineering Sciences*, (30), 330-338, 2018.

[13] M. Bordoloi and SK Biswas, "Sentiment analysis: A survey on design framework, applications and future scopes", Artificial Intelligence Review, 2023.

[14] F. Sağlam, Automated sentiment lexicon generation and sentiment analysis of news. PhD Thesis, Hacettepe University, Computer Engineering, 2019.

[15] T. Keller, Mining the internet of things: Detection of False-Positive RFID Tag Reads using low-level reader data, PhD Thesis, University of St. Gallen, 2011.

[16] V. Singh, P. Rajesh and U. Ashraf et al, "Sentiment analysis of textual reviews; Evaluating machine learning, unsupervised and SentiWordNet approaches**", Institute of Electrical and Electronics Engineers 5th International Conference on Knowledge and Smart Technology-KST**, Chonburi, Thailand, 2013.

[17] O. Appel, F. Chiclana and J. Carter J et al, "IOWA & Cross-ratio Uninorm Operators as Aggregation Tools in Sentiment Analysis and Ensemble Methods.", **Institute of Electrical and Electronics Engineers International Conference on Fuzzy Systems, Naples**, Italy, 2017.

[18] M.A. Mirtalaie, O.K. Hussain, "Sentiment aggregation of targeted features by capturing their dependencies: Making sense from customer reviews", *International Journal of Information Management*, (53), 2020.

[19] M.E. Basiri, A. Kabiri and M. Abdar et al, "The effect of aggregation methods on sentiment classification in Persian reviews", Enterprise Information Systems, (14), 1394-142, 2020.

[20] Y. Li, Q. Pan and T. Yang et al, "Learning word representations for sentiment analysis", Cognitive Computation, 9(6), 843–851, 2017.

[21] H. Benghuzzi, M.M. Elsheh, "An investigation of keywords extraction from textual documents using word2vec and decision tree", *International Journal of Computer Science and Information Security*, 18(5), 13-18, 2020.

[22] S. Suneetha and V. Row, "Aspect-based sentiment analysis: A comprehensive survey of techniques and applications", *Journal of Data Acqusition and Processing-JCST*, 38 (3), 177-203, 2023.

[23] S. Ounacer, D. Mhamdi and S. Ardchir et al. "Customer sentiment analysis in hotel reviews through natural language processing techniques", *International Journal of Advanced Computer Science and Applications*, 14(1), 569-579, 2023.

[24] K. Du, F. Xing and E. Cambria, "Incorporating multiple knowledge sources for targeted aspect-based financial sentiment analysis", ACM Transactions on Management Information Systems, 4(3), 1-24. 2023.

[25] A.P. Rodrigues, R. Fernandes and A. Shetty et al, "Real-time Twitter spam detection and sentiment analysis using machine learning and deep learning techniques", Computational Intelligence and Neuroscience, 5211949, 14 pages, 2022.

[26] T. O'Keefe and I. Koprinska, "Feature selection and weighting methods in sentiment analysis", ADCS'14: Proceedings of the 14th Australasian Document Computing Symposium, Sydney, Australia, January, 2009.

[27] A. Giachanou F. Crestani, "Like it or not: A survey of Twitter sentiment analysis methods", ACM Computing Surveys (CSUR), 49(2), 1-41, 2016.

[28] S. Taj, A. F. Meghji and B. B. Shaikh, "Sentiment Analysis of News Articles: A Lexicon based Approach", **2nd International Conference on Computing Mathematics and Engineering Technologies (ICOMET),** United States, 30-31 Jan 2019.

[29] K. Tutar, M. O. Ünalır and L. Toker, "Development of a framework for ontology-based sentiment analysis on social media", *Pamukkale University Journal of Engineering Sciences*, 21(5), 194-202, 2015.

[30] G.A. Miller, "Beckwith R and Fellbaum C, et al. Introduction to WordNet: An online lexical database", *International Journal of Lexicography*, 3(4), 235-244, 1990.

[31] S. Blair-Goldensohn, K. Hannan and R. McDonald et al, "Building a sentiment summarizer for local service reviews", NLP Challenges in the Information Explosion Era (NLPIX), Beijing, China, April 22, 2008.

[32] N. Oliveira, P. Cortez and N. Areal, "Stock market sentiment lexicon acquisition using microblogging data and statistical measures", Decision Support Systems, 85(6), 2-73, 2016.

[33] S. Baccianella, A. Esuli and F. Sebastiani, "Sentiwordnet 3.0: An enhanced lexical resource for sentiment analysis and opinion mining", **The Seventh International Conference on Language Resources and Evaluation**, Valletta, Malta, 2010.

[34] G. Vural, B. B. Cambazoglu and P. Senkul et al, "A Framework for Sentiment Analysis in Turkish: Application to Polarity Detection of Movie Reviews in Turkish", International Symposium on Computer and Information Sciences, London, UK, 29 October 2012.

[35] C. Türkmenoğlu and A.C. Tantuğ, "Sentiment analysis in Turkish media", **International Conference on Machine Learning (ICML)**, Beijing, China, June 2014.

[36] K.S. Vishnu, T. Apoorva and D. Gupta, "Learning domain-specific and domain independent opinion oriented lexicons using multiple domain knowledge", **Seventh International Conference on Contemporary Computing**, Noida, India, 2014.

[37] S. Akgül, C. Ertano and B. Diri, "Sentiment analysis with Twitter", *Pamukkale University Journal of Engineering Sciences*, 22(2), 106-110, 2016.

[38] F.S. Çetin, G. Eryiğit, "Investigation of aspect based Turkish sentiment analysis subtasks – Identification of aspect term, aspect category and sentiment polarity", *Journal of Information Technologies*, 11(1), 43-56, 2018.

[39] E. Bilgiç, A. Koçak. "Sentiment and opinion mining: Recent issues and an application for the third airport", *Bingöl University Journal of Social Sciences Institute*, 9(17), 327-338, 2019.

[40] G. Arslantürk, "Anti-immigrant attitudes in social media: An examination within the frame of integrated threat theory", Psychology Researches, 1(1), 06-16, 2021.

[41] L.S. Zaabar, M.R. Yaakub and M.I. Abu Latiffi, "Combination of lexicon based and machine learning techniques in the development of political tweet sentiment analysis model", *International Journal of Synergy in Engineering and Technology (IJSET)*, 3(2), 72-83, 2022.

[42] A. Dey, M. Jenamani and J.J. Thakkar, "Senti-N-Gram: An n-gram lexicon for sentiment analysis", Expert Systems with Applications, 2018.

[43] H.S. Hota, D.K. Sharma and N. Verma, "Lexicon-based sentiment analysis using Twitter data: a case of COVID-19 outbreak in India and abroad", Data Science for COVID-19, 275–95, 2021.

[44] B. Muppuru, S. Vamsi and J. Jayashree et al, "Real time sentimental analysis of chat data using deep learning", *Journal of Data Acquisition and Processing*, 38 (2), 4545-4556, 2023.

[45] B.S. Ainapure, R.N. Pise and P. Reddy, et al, "Sentiment analysis of COVID-19 Tweets ssing deep learning and lexicon-based approaches" Sustainability, 15, 2573, 2023.

[46] M. Hu, B. Liu, "Mining and Summarizing Customer Reviews", 10th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, Washington, USA, 2004.

[47] R. Dehkharghani, Y. Saygin and B.A.Yanikoglu, et al. "SentiTurkNet: a Turkish polarity lexicon for sentiment analysis", Language Resources and Evaluation, 50(3), 667-685, 2016.

[48] T. Sezer, What is TS Corpus Project, https://tscorpus.com. 05.04.2023.

[49] Turkish National Corpus (TNC), https://www.tnc.org.tr/tr/. 25.04.2023.

[50] Spoken Turkish Corpus, https://std.metu.edu.tr. 25.04.2023.

[51] U. Eroğul, Sentiment analysis in Turkish, PhD Thesis, Middle East Technical University, 2009.

[52] M. Kaya, G. Fidan and I.H. Toroslu, "Sentiment analysis of Turkish political news", In Proceedings of WI-IAT-12 IEEE/WIC/ACM International Joint Conferences on Web Intelligence and Intelligent Agent Technology, Macau, China, 2012.

[53] E. Akbaş, Aspect based opinion mining on Turkish tweets, PhD Thesis, Bilkent University, 2012.

[54] M. Çetin, M.F. Amasyalı, "Supervised and traditional term weighting Methods for sentiment analysis" **Signal Processing and Communications Applications Conference (SIU)-21st**, (978):1-4, 2013.

[55] Ç. Aytekin, "An opinion mining task in Turkish language a model for assigning opinions in Turkish blogs to the polarities", *Journalism and Mass Communication*, 3(3), 179-198, 2013.

[56] C. Özsert and A. Özgür, "Word polarity detection using a multilingual approach", **Computational Linguistics and Intelligent Text Processing Conference Paper**, 75-82, 2013.

[57] M. Meral and B. Diri, "Sentiment analysis on Twitter", IEEE 22nd **Signal Processing and Communications Applications Conference (SIU)**, Trabzon. Turkey, 2014.

[58] M. Baykara and U. Göktürk, "Classification of social media shares using sentiment analysis", UBMK-17, **2nd International Conference on Computer Science and Engineering (UBMK),** Antalya, Turkey, 2017.

[59] Ö. Demir, A.I. Baban Chawai and B. Doğan, "Sentiment analysis using dictionary based approach in Turkish texts", International Periodical of Recent Technologies in Applied Sciences, 1(2), 58-66, 2019.

[60] R. Ghasemi, S.A. Ashrafi Asli and S. Momtazi, "Deep Persian sentiment analysis: Cross-lingual training for low-resource languages", *Journal of Information Science*, 48-1, 2020.

[61] R. Yang, "Machine learning and deep learning for sentiment analysis over students' reviews: An overview study", Preprints.org, 2021.

[62] S. Kılıçer, R. Samlı, "Review of sentiment analysis studies on texts in different languages", *Dicle University Engineering Faculty Journal of Engineering*, 13(3), 2022.

[63] D. Faggella, "What is artificial intelligence, an informed definition", https://emerj.com/ai-glossary-terms/what-is-artificial-intelligence-an-informed-definition/.28.04.2023.

[64] B. Mahesh, "Machine learning algorithms-A review", International Journal of Science and Research (IJSR), 9(1), 381-386, 2018.

[65] H. Akpınar, 2000, "Knowledge discovery and data mining in databases", *Istanbul University Journal of the School of Business (IUJSB)*, 29(1), 1-22, 2000.

[66] M. Sağlam, "Key themes in brand reputation research: A bibliometric analysis with VOSviewer software", *Research Journal of Business and Management*, 9(1), 1-12, 2022.

[67] Ü.H. Atasever, "The use of boosting, support vector machines, random forest and regression tree methods in satellite images classification, PhD Thesis, Erciyes University, 2011.

[68] E. Akçetin, H. Turgut, "Data analyzing and data mining applications in incomplete data in the business database", *Route Educational and Social Science Journal*, 2(5), 181-199, 2015.

[69] H. Turgut, An application for the diagnosis of alzheimer's disease using data mining processing, PhD Thesis, Süleyman Demirel University, 2012.

[70] A.C. Tantuğ, "Text classification", *TBV Journal of Computer Science and Engineering*, 5(2), 2012.

[71] M. Muhammed, F. Rustam and F. Alasim, et al, "What people think about fast food: Opinions analysis and LDA modeling on fast food restaurants using unstructured tweets", PeerJ Computer Science, 2023.

[72] F. Benrouba, B. Rachid, "Emotional sentiment analysis of social media content for mental health safety", Research Square, 2023.

[73] H.J. Alantari, I.S. Currim and Y. Deng, et al, "An empirical comparison of machine learning methods for text-based sentiment analysis of online consumer reviews", *International Journal of Research in Marketing*, 39 (1), 1-19, 2021.

[74] P. Monika, C. Kulkarni and N. Harish Kumar, et al, "Machine learning approaches for sentiment analysis: A survey", *International Journal of Health Sciences*, 6(S4), 2022.

[75] S. Saifullah, R. Dreżewski and F.A. Dwiyanto, et al, "Sentiment analysis using machine learning approach based on feature extraction for anxiety detection", Computational Science-ICCS, (6), 2023.

[76] K. Gulati, S.S. Kumar and R.S. Kumar Boddu, et al, "Comparative analysis of machine learning-based classification models using sentiment classification of tweets related to COVID-19 pandemic", Materials Today: Proceedings, 51(1):38-41, 2021.

[77] Y. Chen, Y. Rao and S. Chen, et al., "Semi-supervised sentiment classification and emotion distribution learning across domains", ACM Transactions on Knowledge Discovery from Data, 17 (5), 1-30, 2023.

[78] M. Shahzad, C. Freeman and M. Rahimi, et al, "Predicting Facebook sentiments towards research", *Natural Language Processing Journal*, 3 (100010), 2023.

[79] Y. Wang, J. Guo and C. Yuan, et al, "Sentiment analysis of Twitter data", Applied Sciences, (12) 22, 2022.

[80] Y. Al Amrani, M. Lazaar and K.E. El Kadiri, "Random forest and support vector machine based hybrid approach to sentiment analysis", Procedia Computer Science, (127), 511-520, 2018.

[81] M. Thelwall, K. Buckley and G. Paltoglou, et al, "Sentiment strength detection in short informal text*", Journal of the American Society for Information Science and Technology*, (61)12, 2544-2558, 2010.

[82] X. Bai, "Predicting consumer sentiments from online text", Decision Support Systems, (50)4, 732-742, 2011.

[83] C. Kaur, M.S.A. Boush and S.M. Hassen, et al, "Incorporating sentimental analysis into development of a hybrid classification model", *International Journal of Health Sciences (IJHS)*, 6 (S1), 1709–1720, 2022.

[84] D. Tiwari, B. Nagpal and B.S. Bhati, et al, "A systematic review of social network sentiment analysis with comparative study of ensemble-based techniques", Artificial Intelligence Review, 2023.

[85] S. Khaleghparast, M. Maleki and G. Hajianfar, et al, "Development of a patients' satisfaction analysis system using machine learning and lexicon-based methods", BMC Health Services Research, (23), 280., 2023.

[86] M. Wankhade, A.C.S. Rao and C. Kulkarni, "A survey on sentiment analysis methods, applications, and challenges", Artificial Intelligence Review, (55), 5731–5780. 2022.

[87] M. Kristina, M. Marian and H. Miroslava, "Classification of special web reviewers based on various regression methods", Acta Polytechnica Hungarica, (17), 229-248, 2020.

[88] J.R. Chang, H.Y. Liang and L.S. Chen, "Novel feature selection approaches for improving the performance of sentiment classification", *Journal of Ambient Intelligence and Humanized Computing*, 2020.

[89] R.K. Dey, A.K. Das, "Modified term frequency-inverse document frequency based deep hybrid framework for sentiment analysis", Multimedia Tools Applications, (82), 32967–32990, 2023.

[90] G. Yoo and J. Nam, "A Hybrid Approach to Sentiment Analysis Enhanced by Sentiment Lexicons and Polarity Shifting Devices", The 13th Workshop on Asian Language Resources, Kiyoaki Shirai, Miyazaki, Japan, May 2018.

[91] B. Erşahin, Ö. Aktaş and D. Kılınç, et al. "A hybrid sentiment analysis method for Turkish", *Turkish Journal of Electrical Engineering and Computer Sciences*, 27(3), 1780 -1793, 2019.

[92] Y. Madani, M. Erritali and B. Bouikhalene, "A new sentiment analysis method to detect and analysis sentiments of Covid-19 moroccan tweets using a recommender approach", Multimedia Tools and Applications, 82, 27819–27838, 2023.

[93] A. T. Mahmood, S. S. Kamaruddin, Raed Kamil Naser, "A Combination of Lexicon and Machine Learning Approaches for Sentiment Analysis on Facebook", *Journal of System and Management Sciences*, 10(3), 140-150, 2020.

[94] S.K. Bharti, K.S. Babu, "Automatic keyword extraction for text summarization: A survey", arXiv preprint, 2017.

[95] P-I. Chen, S-J. Lin, "Automatic keyword prediction using google similarity distance", Expert Syst Appl; 37(3), 1928-1938, 2010.

[96] M.R.R. Rana, S.U. Rehman and A. Nawaz, et al, "Aspect-based sentiment analysis for social multimedia: A hybrid computational framework", Computer Systems and Engineering, 46(2), 2415-2428, 2023.

[97] T. Sun, L. Jing and Y. Wei, et al, "Dual consistency-enhanced semi-supervised sentiment analysis towards COVID-19 tweets", IEEE Transactions on Knowledge and Data Engineering, 1-13, 2023.

[98] S. Barreto, R. Moura and J. Carvalho, et al, "Sentiment analysis in tweets: an assessment study from classical to modern word representation models", Data Mining and Knowledge Discovery, (37), 318–380, 2023.

[99] D. Yulia, E.A. Bilashova, "Learning analytics of MOOCs based on natural language processing: CEUR", 3017(15), 187-197, 2021.

[100] T. Doğan, A.K. Uysal, "Comparing the Performances of Term Weighting Methods on Medical Document Classification", 6th International Symposium on Innovative Technologies in Engineering and Science (ISITES201), Antalya, Türkiye, 2018.

[101] M.M. Alexandrovna, A.L.M. Ali Mohsin, "Tweet sentiment analysis with CNN and XG-BOOST", Sustainability, 21(4), 1-9, 2023.

[102] A.S. Bandhakavi, W. Nirmalie and P. Deepak, "Lexicon based feature extraction for emotion text classification", Pattern Recognition Letters, 2016.

[103] S. Saranya, G. Usha, "A Machine learning-based technique with IntelligentWordNet lemmatize for Twitter sentiment analysis", Intelligent Automation and Soft Computing, 36(1), 339- 352, 2023.

[104] G. Kaur, A. Sharma, "A deep learning-based model using hybrid feature extraction approach for consumer sentiment analysis", *Journal of Big Data*, 10 (5), 2023.

[105] A. Alshehri and A. Algarni, "TF-TDA: A novel supervised term weighting scheme for sentiment analysis", Electronics, 12(7):1632, 2023.

[106] A. Sharma, S. Kumar, "Machine learning and ontology-based novel semantic document indexing for information retrieval", Computers and Industrial Engineering, 176, 108940, 2023.

[107] M. Hingle, D. Yoon and J. Fowler, et al, "Collection and visuliation of dietary behavior and reasons for eating using Twitter", *Journal of Medical Internet Research*, 15(6), e125, 2013.

[108] B.S. Park, J. Jang and C. Ok, "Analyzing Twitter to explore perception of Asian restaurants", *Journal of Hospitality and Tourism Technology*, 7(4), 405-422, 2016.

[109] N. Mishra and A.A. Singh, "Use of Twitter data for waste minimisation in beef supply chain", Annual Operations Research, (270), 337-359, 2018.

[110] A. Singh, N. Shukla and N. Mishra, "Social media data analytics to improve supply chain management in food industry", Transportion Research Part E: Logistics and Transportation Review, 114(C), 398-415, 2018.

[111] R. El-Khchine, A. Amar and Z.E. Guennoun, et al, "Machine learning for supply chain's big data: State of the art and application to social network's data", **MATEC Web of Conferences**, 2018.

[112] K. Zahoor, N.Z. Bawany and S. Hamid, "Sentiment analysis and classification of restaurant reviews using machine learning", **21st International Arab Conference on Information Technology (ACIT)**, 1-6 Giza, Egypt, 2020.

[113] E.S. Alamoudi, N.S. Alghamdi, "Sentiment classification and aspect-based sentiment analysis on yelp reviews using deep learning and word embeddings", *Journal of Decision Systems*, 30(2-3), 259-281, 2021.

[114] A. Liapakis, T. Tsiligiridis and C. Yialouris, et al, "A Sentiment lexicon-based analysis for food and beverage industry reviews, the Greek Language paradigm", Natural Language Engineering, (9), 21-42, 2020.

[115] K. Ahmed, M.I. Nadeem and Z. Zheng, et al, "Breaking down linguistic complexities: A structured approach to aspect-based sentiment analysis*", Journal of King Saud University - Computer and Information Sciences*, 35 (8), 1-21, 2023.

[116] Ö. Aktaş, B. Coskuner and I. Soner, "Turkish Sentiment Analysis Using Machine Learning Methods: Application on Online Food Order Site Reviews", *Journal of Artificial Intelligence and Data Science*, 1(1), 1-10, 2021.

[117] Z. Liu, C. Yang and S. Rüdian, et al, "Temporal emotion- aspect modeling for discovering what students are concerned about in online course forums", Interactive Learning Environment, 27 (5–6), 598–627, 2019.

[118] Imbalanced Data. https://developers.google.com/machine-learning/data-prep/construct/sampling-splitting/imbalanced-data. 25. 04. 2023.

[119] V. A. Pitogo and C. D. L. Ramos, "Social media enabled e-participation: a lexicon-based sentiment analysis using unsupervised machine learning", **In: Proceedings of the 13th International Conference on Theory and Practice of Electronic Governance, ACM,** 518–528, 2020.

[120] R. Purbasari, E. Munajat and F. Fauzan, "Digital innovation ecosystem on digital entrepreneur: Social network analysis approach", *International Journal of E-Enterpreneurship and Innovation*, 13(1), 21, 2023.

[121] Y. Kim, T. Y. Choi, T. Yan and K. Dooley, "Structural investigation of supply networks: A social network analysis approach", *Journal of Operations Management*, 29, 194-211, 2011.

[122] Zargan Translation Dictionary. https://www.zargan.com/tr/q/demeaning-ceviri-nedir, 2021.

[123] Tuareg Translation Dictionary. https://www.dictionary.com/browse/tuareg, 2021.

[124] F. Å. Nielsen, "A new ANEW: Evaluation of a word list for sentiment analysis in microblogs", Proceedings of the ESWC2011 Workshop on 'Making Sense of Microposts': Big things come in small packages 718 in CEUR Workshop Proceedings 93-98. 2011.

[125] Excel-365 Synonyms and Antonyms Dictionary. Microsoft Office-365, 2021.

[126] O. C. Atalaya, J. A. Tuesta, D. B. Mares, A. G. Pacheco, O. M. León, M. Q. Silvestre, G. T. Quispe and R. S. Bazalar, "K-Fold Cross-Validation through Identification of the Opinion Classification Algorithm for the Satisfaction of University Students", *International Journal of Online and Biomedical Engineering (iJOE)*. 19(11), 140-158, 2023.

[127] A. Srivastava. "A Review on Sentiment Analysis of Twitter Data Using Machine Learning Techniques". *International Journal of Engineering and Management Research*, 14(1), 2024.

ANNEXES

Annex-1- Turkish Stopword List

Turkish Stopwords Prepared fort the Study
acaba, akabinde, akebinde, altı, altına, altında, altta, ama, ancak, aralarında, arasında, arasından, arkada, arkasında, artık, asla, aslında, aşağı, aşağıdan, aşağısı, aşağıya, aynen, aynı, ayrıca, az, azıcık, bana, başka, bazen, bazı, bazıları, bazısı, belki, ben, beni, benim, benzer, benzeri, beş, bi, bide, bile, binaenaleyh, bir, bir defa, bir hayli, bir kere, bir kere daha, bir kerecik, bir kimse, bir miktar, bir şey, bir şeyi, bir takım, bir vakitler, bir zamanlar, biraz, biraz önce, birbirinden, birçoğu, birçok, birçokları, birde, biri, birisi, birkaç, birkaçı, birşey, birşeyi, bişey, bişi, biz, bizatihi, bize, bizi, bizim, bizimki, bizzat, bizzat ben, bizzat kendileri, bizzat kendimiz, bizzat kendisi, boyunca, böyle, böylece, böylelikle, böylesine, bu, bu gibi, bu kadar, bu noktada, bu suretle, bu şekilde, bu türlü, bugünlerde, buna, buna benzer, bundan, bundan başka, bunlar, bunu, bunun, bunun gibi, bununla birlikte, burada, buraya, bütün, civarında, çevresinde, çoğu, çoğuna, çoğunu, çok, çok az, çünkü, da, daha, daha çok, daha evvel, daha fazla, daha önce, daha ziyade, dahi, de, dedi, dedik, dediler, dedim, dedin, dediniz, değin, demek, demek ki, demi, dışarda, dışarı, dışarıda, dışarıya, diğer, diğeri, diğerleri, dimi, diye, diyor, dokuz, dolayı, dört, ediyor, eğer, ek olarak, elbette, en, epeyce, eski, eskiden, esnasında, etmek, etrafında, evvelce, evvelki, fakat, falan, felan, filan, gene, gibi, hala, halbuki, halinde, hangi, hangisi, hangisini, hani, hatta, hayır, hem, hemen, hemen sonra, henüz, hep, hepsi, hepsine, hepsini, her, her biri, her ikisi, her ikisini, her ne, her ne kadar, her tarafa, herbiri, herhangi bir, herhangibir, herkes, herkese, herkesi, hiç, hiç birine, hiç birini, hiç kimse, hiçbirine, hiçbirini, içerisinde, içerisine, içi, için, içinde, içinden, içine, iken, iki, ikisi, ikisini, ilaveten, ile, ileri, ileride, ileriye, ilk, ise, ise de, işbu, işte, itibarıyla, itibarıyle, itibariyla, itibariyle, iyi, kaç, kadar, karşı, kendi, kendi kendine, kendi kendini, kendi kendinize, kendi kendisine, kendi kendisini, kendilerinde, kendilerine, kendilerini, kendiliğinden, kendim, kendin, kendine, kendini, kendinin, kendiniz, kendinizde, kendinize, kendisi, kendisinin, keza, ki, kim, kime, kimi, kimin, kimisi, kimse, lakin, madem, mı, mıdır, mısın, mısınız, mıydı, mıyım, mi, midir, misin, misiniz, miydi, miyim, mu, mudur, musun, musunuz, muydu, muyum, mü, müddetince, müdür, müsün, müsünüz, müydü, müyüm, nasıl, ne, ne kadar, ne sebeple, ne vakit, ne zaman, neden, nedeniyle, nedir, nerede, nereden, neredeyse, nereye, nesi, netice olarak, neyse, niçin, niye, o anda, o halde, o hususta, o kadar, o noktada, o türlü, o vakit, o yer, o yere, o zaman, o zamanın, o zamanki, oldukça, olmak, olmakla beraber, olur olmaz, on, ona, ondan, ondan sonra, onlar, onlara, onlardan, onları, onların, onlarınki, onu, onun, onunki, ora, orada, oradaki, orası, orasında, oraya, oysa, oysaki, öbür, öbürü, ön, önce, önceden, önceki, önünde, ötede, öteki, öteye, ötürü, öyle, öyle ise, öylesine, özellikle, pek çok, rağmen, sadece, sana, sanki, sebebiyle, sebep, sekiz, sen, senden, seni, senin, seninki, sırf, siz, sizden, size, sizi, sizin, son, son derece, sonra, sonuç olarak, sözcük, süresince, şahıs, şahsı, şayet, şey, şeyden, şeye, şeyi, şeyler, şimdi, şöyle, şu, şu anda, şu halde, şu kadar, şu sırada, şuna, şunda, şundan, şunlar, şunu, şunun, şurada, şuraya, ta kendisi, ta ki, takdirde, takriben, tam, tamamen, tamamı, tamı tamına, tastamam, tekrar, tıpkı, tıpkısı, tüm, tümü, üç, üstelik, üstü, üstünde, üstüne, üzere, üzerinde, vaktiyle, var, vasıtasıyla, vb, ve, veya, veyahut, vs, ya, ya da, yada, yahu, yahut, yakınında, yaklaşık, yalnız, yanında, yani, yalnız, yapar, yapıyor, yapmak, yeniden, yerine, yıl, yine, yoksa, yukarı, yukarısı, yukarısında, yukarıya, yüzünden, zarfında, zaten, zira, ziyade

Annex-2- Seed Word List

Seed Words for Preparing the Classification Lexicons
abartma, abla, acı, acılı, adalet, adres, alakasız, açık paket, aç kaldık, açım, adet, adres, ağbi, ağız tadı, aksaklık, akşam, alakasız, alış veriş, alışveriş, amatör, anne, aracılık, artık yeter, asla, aşırı, aşırı yağlı, aynı hata, ayran, az, az yağlı.
baba, bafra pidesi, bahane, bakmak, bayat, bayat ürün, bekle, berbat, besleyici, bıçak, bıkmak, bilet, bimutlulukgetir, bir daha asla, boş, boş köfte, bozuk, bozuk salata, bölge, bulantı, burger, buz, buz gibi, büfe, büyük boy, büzüşmüş.
cafe, canım çekti, canlı yardım, cevap.
çatal, çeşit azlığı, çiğ köfte, çocuk, çok, çok az, çok iyi, çok kızarmış, çok kötü, çöp, çözüm.
dağılmış, dakika, değer, deli olmak, deneyim, dışarıdan sipariş, dikkatsiz, dip, diş kovuğu, diyet, doğru bir nokta, domates, double, doymadım, doymayan, döner, duble, dükkan, dürüm, düzeltme.
eğitim, eğitimsiz personel, ekmek, ekmek arası, eksik, eksik geldi, eksik ürün, en beğendiğim, en güzel, en güzel yanı, en kötü, en kötü yemek, et et döner, etraf, extra, ekstra, ev, evde yemek, ev yemeği.
fark, fast food, fast food zinciri, fazla fiyat, fındık lahmacun, fıstık lahmacun, firma, fiyasko, fiyat, fiyat politikası.
gaflet, gıda, gram, gece, gece yarısı, gecikme, geç gelen, gönderi, gönderme, güzel.
haber ver, hak, hak etmiyor, hamur, hamur gibi, hamburger, hata, hayal kırıklığı, hazır yemek, hediye, helal etmiyorum, helal olsun, hesap, hızlı dönüş, hızlı yemek, hijyen, hizmet, hizmet kalitesi, hizmet sıfır, hizmet verme, homeburger.
ıslak, ıslak hamburger.
iade, iade talebi, içecek, içi boş, iftar, iletişim, iletişim sorunu, ilgisiz, indirim, indirim kodu, insaf, insan sağlığı, internet sitesi, internetten yemek siparişi, ishal, itham, itibar, iptal, işte yemek, İtalyanpizza, İtalyan lezzetleri, iyi fikir.
joker, joker indirim, kaba, kampanya, kardeş, karadeniz pidesi, karışık pide, kart, kasa, kaşar, kaşarlı, kaşık, kavurmalı, kavurmalı kaşarlı, kayış gibi, kazanma, ketçap, keyif, kıkırdak, kıl, kıral gibi, kıralsın, kıymalı, kıymalı pide, kızarmış, kızartma, king burger, koku, kola, köfte, köpek gibi açım, kötü, kötü puan, kredi kartı, kurumsal, kurye, kusma, kuşbaşılı, kuşbaşılı kaşarlı, kuş tüyü, küçük, küçük boy,.
lahmacun, latte, lanet, lanet olsun, lanet ediyorum, leziz, lezzetli, lezzetsiz, limit
mağdur, mağduriyet, mahal, mail adresi, malzeme, malzeme eksikliği, manipülasyon, mayhoş, mayonez, memnuniyet, memnuniyetsizlik, menü, mönü, merkez, mesaj, mide, minimum, minimum tutar, mis, mis gibi, mobil uygulama, muamma, muhatap, multinet, mutfak, mükemmel, mükemmel ürün, müşteri, müşteri hizmetleri, müşteri memnuniyeti, müşteri memnuniyetsizliği.
nakit, ne yesem, niyet, numara.
objektif, olumlu, olumlu yorum, olumsuz, olumsuz yorum, onay, online, online ödeme, online sipariş, orta boy, otomatik ulaşmak, oyala.
ödeme, ödeme yöntemi, öğle, öğrenci, öğün, öncelik, öneri, özen, özensiz, özür.
paket, paketleme, paket servisi, paket sipariş, para, para iadesi, patates, personel, personel ilgisizliği, peynirli, peynirli pide, pide, pilav üstü döner, pişmanlık, pipet, pişmiş, pizza, pizzasipariş, pizza siparişi, poşet, problem, promosyon, puan.
resmi tatil, restoran, restoran zinciri, restorant, restorant zinciri, restaurant, rezalet.
saat, saç, saçmalık, sağlıklı, sağlıksız, salata, Samsun pidesi, Samsun pidecisi, sanki, sayfa, saygısızlık, servis, servis elemanı, servis sıfır, set card, setcard, setkart, severek, seviliyorsun, seviliosun, seviyorum, sıkıntı, sinek, sipariş, sipariş hattı, sipariş iptali, sipariş notu, sipariş onayı, sipariş öncesi, sipariş sonrası, sistem, sodexo, soğuk, soğumuş, son sipariş, sonuç, sorumlu, sorumsuzluk, sorun, sos, sübjektif, suç, suçlama, sufle, süre, şımartmak, şikayet, şube, şüpheli.
taahhüt, takip, talep, talep etmek, tam zamanında, tat, tavuk, tavuk burger, tavuk döner, tavuk hamburger, tecrübe, tehlike, telafi, telefon, temiz, terleme, teslim, teslim etmek, teslimat, teslimat süreci, teslimat süresi, teşekkür, ticket, tövbe, trend, tutar, tuz, Türk mutfağı, tüy.
uğraşıyorum, ulaşamamak, umursamaz, umursuz, unutmuş, urfa, uygulama, uygun lokasyon, uzak.
ücret, ücret iadesi, ürün, üstü boş, üşenmek, üye.
vıcık, vıcık vıcık, vurdum duymaz.
Web sitesi
yağlı, yalan, yanlış, yalnış, yanlış sipariş, yanlış sipariş, yanmış, yapışmış, yaptırım, yardım, yarım, yaşasın, yazık, yemek, yemek arası, yemeksepeti, yemek sepeti, yemek yapma, yemek yok, yeter artık, yetersiz, yetki, yetkili, yettim, yoğunluk, yol, yorum, yumurtalı, yüzde.
zaman, zamanında, zehirlendim, zehirlenme, zevk veren.

Annex-3- Grouped Words for Boosted Sentiment Lexicon

Grouped Words (Aggregation of Words) for Boosted Sentiment Lexicon
{"acele":["acele", "acil"], "ağzımın su":["ağzımın su", "azımın su"], "aksiyon":["aksiyon","ekşın"], "aralıksız":["aralıksız", "durmadan"], "hediye":["armağan", "ödül", "hediye"], " aşık":["asko", "aşko", "aşkito", " aşık"], "beğeni":["beğeni","beğend"], "canı çek":["canım çek", "canı çek"], "çık hayat":["çık aklı", "çık hayat"], "eline sağlık":["eline sağlık", "ellerinize sağlık"]," esas ":[" elit"," esas "], "gönül":["gönül", "gönlü"], "hakikaten":["hakikaten","hakkaten"], "hapır hupur": ["hapır hupur", "hapur hupur"], "hayırlısı":["hayır ol","hayırlısı"], "hastası":["hastası", "hasta ol"], "hemen":["hemen", "hızla", "hızlı", "ivedi", "çabuk"], "inşallah":["inşallah", "inş "], " istiyor": [" istiyor"," ister"," istey"," istemiş"," isted"], "iyilik":["iyice", "iyilik"], "kalp": ["kalp", "kalb"], "kardeşim":["kardeşim", "kardeşlerim"], "keyif":["keyf","keyif"], "lezzet":["leziz", "lezzetli", "nefis", "lezzet"], "müsaade":["müsade","müsaade"], "otomatik":["otomasyon", "otomatik"], "özenli":["özendir", "özenli"], "saygılı":["saygılı", "saygın"], "şükür":["şükür", "şükran"], "taktir":["taktir", "takdir"],"uygun":["uygun"," uyum"],"usta":["usta"," piri "], "teşvik":["teşfik","teşvik"], "yakışıklı":["yakışıklı", "yakışır"], "abuk sabuk":["abidik gubidik", "abuk sabuk", "antin kuntin"], "acayip":["acayip", "absürt"],"acemi":["acemi","toy"], "açlık": ["aç kald", " açım"," açız", "açlıktan bayıl"], "ağlıcam":["ağlıcam","ağlicam"], "sövmek":[" amk","amq","aq"," mk ", "skm","skt"," söv"], "aşırı yağlı":["aşırı yağ", "çok yağlı"], "bağlanmak": ["bağlanam","bağlanm"], "bomboş":["bomboş"," boş "], "yasak":["banla", "yasak"], "bela": ["belanı versin", "belanızı versin"], "bıkmak":["bıkkın", "bıktım"],"boşuna":["boş yere", "boşuna"], "bozuk":["bozuk", "bozulmuş"], "donmuş":[" buz ","donmuş"],"pişmemiş":[" çiğ ", "pişmemiş"], "çökmüş":["çökmüş", "çöktü"], "dağınık":["dağınık", "dağıl"], "dağ başı":["dağ baş", "dağbaş","dağın baş"], "deli ol":["deli ol", "delir"], "duygu sömür":["duyar kas", "duygu sömür"], "gecik":["gecik", "geç gelen"], "gergin":["gergin", "geril"], "getirmek": ["getirem", "getirm"], " uzak": [" ırak "," uzak"], " istemiyor":[" isteme"," istemiyor"],"kafayı ye": ["kafayı ye", "kafayı yi"], "kahrolsun":["kahretsin","kahrolsun"], "kaldık":["kaldık","kaldım"], "kirli": ["kirlenmiş","kirli"], "lanet":["lanet","nalet"],"negatif":["negatif"," eksi "], "orospu":["or*spu çocuğu","or*spuçocuğu","orospu"], "rezil":["rezalet", "rezil"], "soğuk":["soğuk", "soğumuş"], "trip":["trib","trip"], "ulaşam":["ulaşam","ulaşmam", "ulaşm"], "umrunda değil":["umrumda değil", "umrunda değil"], "üzgün":["üzgün","üzül", "üzüyo", "üzücü"], "yanık":["yanık", "yanmış"], "yeter artık":["yeter artık", "yeter yahu"], "yetersiz":["yetersiz", "yok"], "zarar ":["zarar ","zararlı"], "zoraki":["zor bela", "zoraki", "zorla", "zorunda bırak", "zorunda kal"]}

Annex-4- Boosted Sentiment Lexicon

Boosted Sentiment Lexicon Words
["10 puan","number one", "10/10", "6/10", "8/10", "acar", "acele", "adalet", "aferin", "afiyet", "ağzımın su", "ak pak", "aksiyon", "alakalı", " alfa", "anında", "anlamlı", "aralıksız", "arkadaş", " artı", "arzu"," asıl "," aşık", "aşer"," aşkın", "avantajlı", "bağımlı", "bahşiş", "başarılı", "başlıca", "bayıl", "bayram", " baz ", "bebeğim", "beğeni", "beklenti", "beyefendi", "bilerek", "bilgili", "bilinçli", "boğazımdan geçm"," bol ","bravo", "buruk", "canı çek", "cansın", "centilmen", "ciddi", "çare", "çekiliş", "çeşni", "çık hayat", "çılgın", "çok iyi", "çözüm", "daima", "dayanışma", "değerli", "demlenmiş", "dengeli", "derle", "derman", "destek", "devamlı", "dikkatli", "dilek", "doğal", "doğru","dolu","doyam","dua", "duyarlı", "dürüst", "düşünceli", "düzelt", "düzenli", "efendi", "eğitim", "eksiksiz", "ekstra", "ev yemeği", "eline sağlık"," esas ", " esen "," eşit"," etkili", "faydalı", "fazilet", "gayet", "gerçek ","gerekli"," göm", "gönül", "görev", "görgülü", "gurur", "güven", "güzel", "hakikaten ", "hakkıyla", "halis", "hapır hupur", "hasret kal", "hassas", "hastası", "havalı", "hayati", "hayırlısı", "hayran", "hediye", "helal", "hemen", "heves", "hijyen"," hoş ","hukuk", "huzur", "içim", "ikram", "ilave"," ilgili", "iltifat", "incelik", "indirim", "insan sev", "insani", "inşallah"," istiyor", "işbirli", "iştahlı", "itibar", "iyi fikir", "iyi puan", "iyi yemek", "iyilik"," izin", "jest", "joker", "kabul", "kalp", "kalıcı", "kaliteli", "kanka", "kardeşim", "karlı", "kazan", "kefil", "kesintisiz", "keyif", "kızarmış", "kibar", "kolay", "kral", "kurban", "latif", "layık", "lokum", "makbul", "mantıklı", "medeni", "memnuniyet", "merhamet", "mesut", "meşhur", "minimum tutar", "minnet", " mis ", "motivasyon", "muhteşem", "mutlu", "mükemmel", "müsaade", "müsait ","nazik", "nitelikli", "nizam", "olağanüstü", "olumlu", "onay", "onur", "optimum", "otomatik", "öncelik", "önem", "öner", "özel", "özenli", "patla", "pişkin", "pişmiş"," pls ", "pozitif", "prestij", "profesyonel", "promosyon", "rahat", "rica", "rüzgar", "safa", "sağlam", "sağlıklı", "sakin", "salim", "saygılı", "sefa"," seri", "sev", "sıcacık", "sipariş onay","sistem","sorumlu","stalk","sürekli", "sürpriz", "şahane", "şerefine", "şevk", "şımart", "şükür", "taahhüt", "tadı güzel", "takip", "taktir", "talep", "tarafsız", "tatlı", "tavla", "taze", "tecrübe", "teklif", "telafi", "temel", "temiz", "terbiyeli", "teşekkür", "teşvik"," tez ","titiz", "tolerans", "toparla", "toplu", "tutul", "ucuz", "umut", "usta", "uyanık"," uygun", "ücretsiz", "üstlen", "üstün", "vaktinde", "verimli", "vurul", "yakın", "yakışıklı", "yarar", "yardım", "yaşasın", "yemek video", "yeni ","yerinde", "yeterli", "yoğunlaş", "yöntem", "yüce", "zahmet olmazsa", "zevkli", "abart", "abes", "abuk sabuk", "acayip", "acemi"," acı ","acitasyon", " açık ", "açlık", " adi ", " ağır","ağlıcam", "ahlaksız", " aksak"," aksi ","alakasız", "alt tarafı", "sövmek", "andaval", "anksiyete", "anlamsız", "arıza", "asılsız", "aşağı", "aşınmış", "aşırı yağlı", "ayar ol"," ayıp"," ayı ","azal", "azar ye", " azarla", "bağlanmak", "bahane", "balık hafıza", "balon", "yasak", "basit", "başarısız"," bat", "bayağı", "bayat", "baygın", "beceriksiz", "beddua", "beğenm", "beklet", "bela", "bencil", "bitmiştir", "berbat", "bereket", "beter", "beyhude", "beyinsiz", "bez", "bıkmak", "bilgisiz", "bilinçsiz"," bilme", "bin pişman", "daha asla", "bitkin", "blokl", " bok ", "bomboş", "boşuna", "botla", "boykot", "bozuk", "böcek"," bulan", "donmuş", "büyütme", "cahil", "camış","cansız", "cenabet", "cennet", "cereme", "cesaretim yok", "ceza"," cılk", "crash","çakal","çamur","çekin","çelişki","çıkar","çıkış","çıldır","pişmemiş","çirkin", "çok kötü", "çökmüş"," çöp ","çürü", "dağ başı", "dağınık", "dalgın", "dandik"," dar ", "dayanma sınır", "dedikodu", "değersiz", "deli ol", "dengesiz", "dert"," dışı", "diken üst", "doym", "dökülmüş", "dönek", "dram", "duygu sömür", "duyarsız", "düğüm", "düşman", "düşük", "düşüncesiz", "düzenbaz", "düzensiz", "edepsiz"," eksik ", "eleştiri", "enayi"," engel", "erimiş"," ertele"," eşek ","eyvah"," ezik","fake","fani", "fasa fiso", "faydasız", "fazla", "felaket", "feleğim şaş","fena","fırsatç","fiyasko", "gaflet", "gazabına uğra", "geber", "gecik", "geçici", "gelm", "gereksiz", "gergin", "getirmek", "gıcık", "gına gel", "görgüsüz", "gözü karart", "gudubet", "hadsiz", "hak etmiyor", "hakaret", "haksız", "halt", "hamur", "haram", "hatalı", "hayal kırık", "haysiyetsiz", "hayvan", "hazetm","hikaye"," uzak", "ıslanm","ıssız","iade","ibne","iflah olm", "iflas", "iğren", "ihmal", "ihtar", " ilgisiz"," illet", "insaf", "iptal", "israf"," istemiyor", "istismar", "isyan", "işkence", "işsiz", "iştahsız", " itici", "iyi değil", "kaba ","kafayı ye", "kahrolsun", "kalas", "kaldık", " kalm","kan emici", "kanser", "kapısının önü", "kara kara düşün","kara liste", "karışık", "kaybet", "kayış", "kazık", "kendimi tut", "keşke"," kını", "kırıl", "kıtlık", "kifayetsiz", "kilo ", "kirli", "kitle", "kokan", "kokm", "kopya", "korkunç", "köle", "kömür", "kötü ","köylü"," kuru ", "kusm", "kusur", "küçük", "küflü", "küstah"," laf ", "lakayıt","lanet", "lastik", "leke"," leş ", "lezzetsiz", "lüzumsuz", "mağdur", "mahsur", "mantıksız", "manyak", "maraz doğ", "mesele", "midem bulan", "mikrop","minik","muhatap","mutsuz","nefret","negatif","olacaksa ol", "odun"," oha ", "olumsuz", "orospu", "ortadan kaldır", "oyala", "öküz"," ölü ","özensiz", "özür", "pahalı", "panik", "paranoyak", "pezevenk"," pis ", "pişman", "problem", "psikopat", "rasgele", "reddet", "rezil", "risk", "rötar", "ruh hasta", "saçma"," saç ", "sağlıksız", "sahte", "sakat", "sakınca", "salak", "salla", "saman", "sası", "savsak", "saygısız", "sebepsiz", "serseri", "sert", "ses seda yok", "sıfır", "sıkıl", "sıkıntı", "sıradan", "sızlan", "sinek", "sipariş hata", "sipariş iptal", "sitem", "sivrisinek", "skandal", "soğukta", "soğuk", "sorm", "sorumsuz", "sorun", "sömür", "sözde"," suç", "sürün", "şaka", "şans gülm", "şerefsiz", "şeytan", "şımarık", "şikayet", "şişir", "şişko patates", "şopar", "şüpheli", "taciz", "tadı yok", "takat", "talan et", "talihsiz", "tasa ", "taş ","tatava yap", "tatsız", "tehlike", "tekel", "telaş", "tembel", "terbiyesiz", "ters", "tırs", "toz", "trip", "tuzsuz","tüketme","tükür","tüy","uçurum","ufak","uğraş","ukala"," ulan", "ulaşam", "umrunda değil", "umursam", "unutmuş", "unutul"," usan"," utan", "uyar", "uydur", "ürkütücü"," üşen", "üzgün", "vahim", "vasat", "vazgeç"," verem", "verimsiz", "vıcık","vicdan azab","vicdansız", "vizyonsuz", "yağı don", "yağsız", "yalaka", "yalan", "yanlış", "yamuk", "yanık", "yanıl", "yapışmış", "yaram", "yaratık", "yaşlan", "yavaş", "yavşak", "yazık", "yemek yok","yemiyo","yerlerde", "yeter artık", "yetersiz", "yorgun", "zarar ","zehir", "zevksiz", "zıkkım", "ziyan", "zoraki"]

Annex-5- Grouped Words for Boosted Product-Service Systems Lexicon

Grouped Words (Aggregation of Words) for Boosted Product-Service Systems Lexicon
{"adet":["adet", "tane"],"çok yağlı":["aşırı yağ", "çok yağlı"],"bozuk":["bozuk", "bozulmuş"], "donmuş":[" buz ","donmuş"], "canı çek":["canım çek", "canı çek"],"pişmemiş":[" çiğ ", "pişmemiş"], "damak tadı":["damak tadı", "damak zevk"], "duble":["double","duble"], "doyurucu":["doyum", "doyurucu"], "ekşimiş":["ekşim"," ekşi"], "ev yapımı":["el yapım", "ev yapım ", "ev yemeği"],"fast food":["fast food","fastfood"],"lezzetli":["leziz", "lezzetli", "nefis"], "soğuk":["soğuk", "soğumuş"], "vejetaryan":["vejetaryan", "vejeteryan"], "lira":["₺","lira", "nakit", " para ", " tl"], "acemi":["acemi", "toy"], "konum":[" alan ","adres", "bölge", "civar"," etraf", "konum", "mahal", "mevki", "muhit", "semt", "sokak"], "alışveriş":["alış veriş", "alışveriş"], "alt sınır":["alt limit", "alt sınır"], "anasayfa":["anasayfa","ana sayfa"], "aplikasyon": ["aplikasyon", " app","mobil uygulama"], "boşuna":["boş yere", "boşuna"], "callcenter": ["callcenter", "call center","canlı destek", "canlı yardım"], "çökmüş":["çökmüş", "çöktü"], "dağın baş":["dağ baş", "dağbaş","dağın baş"], "dağınık":["dağınık", "dağıtıcı", "dağıtım ağı"], "dakika": ["dakika","dk","saat"], "davranış":["davranış","davranm"], "debit":["debit","setcard","sodexo", "multinet"], "dışardan sipariş":["dışardan sipariş", "dışarıdan sipariş"], "entegrasyon": ["entegrasyon","entegre"], "fiyat":["fiat","fiyat"], "gecik":["gecik", "geç gelen"], "gel al":["gel al", "gel-al"], "gelen abi":["gelen abi", "gelen arkadaş"], "gönderi": ["gönderi","gönderm"], "hemen": ["hemen", "hızlı"], "hes cod": ["hes cod", "hes kod"], " uzak":[" ırak ", " uzak"], "iletm":[" ileti", "iletm"], "kapalı":["kapalı", "kapanmış", "kapanış", "kapatmış", "kapatmak"], "kara liste":["kara liste","karaliste"], "kokorecci":["kokoreççi", "kokorecci"], "kurye":["kuriye", "kurye"], "posta":["mail", "posta"], "min tutar":["min paket tutar", "min sipariş tutar", "minimum sipariş tutar", "minimum tutar"], "motokurye":["motokurye", "motosiklet", "motor"],"nerde kal":["nerde kal", "nerden gel"], "otomasyon":["otomasyon", "otomatik"], "özenli":["özenli", "özen göster"], "paket servis":["paket servis", "paket sipariş"], "rasgele": ["rasgele", "rastgele"], "saygılı":["saygılı", "saygın"], "sipariş iptal":["sipariş hata", "sipariş iptal"], "soğuk hava": ["soğukta", "soğuk hava"], "telefon":["telefon","mesaj","sms"], "yemekçi": ["sepetçi", "yemekçi"],"trip":["trib","trip"], "ulaşmam":["ulaşam", "ulaşmam", "ulaşm", "ulaşım"]}

Annex-6- Boosted Product-Service Systems Lexicon

Boosted Product-Service Systems Lexicon Words
["abur cubur", " acı ", "adet", "ağız tadı", "altın günü yiyecek", "ana yemek", "aperatif", "çok yağlı", "ayran", "azıcık", "baharat", " bal ", "bayat", "besin", "biber", "bisküvi", " bol ", "bozuk", "böcek", "cips", "börek", "burger", "buruk", "donmuş", "büyük boy", "canı çek", "çeşni", "çevirm", "çıtır", "pişmemiş", "çoban", " çöp ", "çörek", "dağıl", "damak tadı", "dolma", "domates", "duble", "doym", "doyurucu", "döner", "dünden kalan", "dürüm"," ekle ","ekşimiş", " ekmek ", " eksik ", "ev yapımı", " erzak", "gıda", "eşantiyon", " etli", " etsiz", "fast food", "futuristik", "garnitür", "gevrek", "gluten", "gram", "gurme", "hafif yemek", "hamburger", "hamur", "haram", "hastası", "hatay usul", "havyar", "hazır yemek", "helal"," hoş "," ısıt", "ıspanak", "iade", "içecek"," içki ", "içli köfte", "ikram", "kadınbudu", "kalın hamur", "kayış", "karbonhidrat", "karışık", "kaşarlı", "katık", "kavurma", "köpek yem", "kestane", "ketçap", " kıl ", "kıtır"," kıyma", "kızarmış", "kızartm","kilo ", "kişi başı", "klasik", "kokan", "kokm", "koku", "konserve", "köfte", "kömür", "kurt", " kuru ", "kusm", "kuşbaşı", "küçük", "küflü", "lahmacun","lastik", "latif","latte", "lüfer", "lezzetli", "lezzetsiz", "lokma", "lokum", "makarna", "martı eti", "marul", " maya", "mayonez", "meyve", "meşrubat", "meze ","mide", "minik"," mis ","nane", "organik", "orta boy", "öğün", "ölçü", "patates", "patlıcan", "peynir", "pide", "pilav", "pişmiş", "pizza", "poğaça", "porsiyon", "pörsüm", "reçel", " saç ", "sağlıklı", "sağlıksız", "salata", "salça", "saman", "sarımsak", "sası", "sebze", "sıcacık", "sinek", "soğan", "soğuk", "son kullanma tarih"," sos ", "sucuk", "sufle", "sulu", "şalgam", "şeker", "şerbet", "taraftar", "taş ", "tatlı", "tatsız", "tavuk", "taze", "tereyağ", "turşu", "tuzlu", "tuzsuz", "ufak", "urfa", "ürün", "vegan", "vejetaryan", "vıcık", "yağı don", "yağlı", "yağsız", "yanık", "yanmış", "yapışmış", "yarım", "yemiş", "yeşillik", "yöresel", "yudum", "yumurta", "zehir","zevkli", "lira","3d secure", "abone", "acele", "acemi"," açık "," açıl", "adalet"," adam", "adisyon"," adi ", "ahlaksız", "ak pak", " aksi ", "aktarm", "alakalı", "alakasız", "konum", "alçal", " algı", "alışveriş", "alt sınır", "altyap", "anasayfa", "anında", "aplikasyon", "anlamsız","anlaşılm", "anlayış", "aracı", " asıl ", "asistan", "aşağı"," ayı "," azarla", "bağlantı", "bahşiş", "bakım", "basit", "başlıca", "bayağı", " baz ", "beceriksiz", "bedel", "beklet", "belgeli", "bencil", "beyhude", "beyinsiz", "bıçak", "bildiri", "bilet", "bilgisiz"," anda ", "blokl", "boşuna", "bölüm", "buton", "cahil", "callcenter", "cevap", "crash", "cüzdan", "çaba", "ivedi", "çağrı", "çakal", "çalışan", "çalışm", "çatal", "çelişki", "çevre", "çıkış", "çiçek", "çirkin", "çökmüş", "dağın baş", "dağınık", "dakika", "dalga", "dalgın", "danışma", "davranış", "debit", "değerli", "değersiz", "deney", "denk", "dert", "destek", "devre", "dezenfekte", "dış kapı", "dışardan sipariş", "dikkate alm", "diyet", "doğru", "dönem", "duyarsız", "dükkan", "dürüst", "düşük", "düşüncesiz", "düşünceli", "düzensiz", "edepsiz", " ederi"," eğitim", "eksiği", "eleman", "emekçi", "entegrasyon", " esas ", " esnaf "," eşek ", " eşit", "fatura", "faydasız", "fazilet", "fazla", "fiyat", "filtre", "firma", " fiş", "gamsız", "gayret", "gece yarısı", "gecik", "gel al", "gelen abi", "gelir", "gerçek ", "gereksiz", "getiren", "getirtm", "gönderi", "görev", "görgülü", "görgüsüz", "görmem", "görüş", "götürm", "haberleş", "hack", "hain", "hakaret", "hata ver", "havalı", "hayvan", "hediye", "hemen", "hes cod", "hesap", "hizmet", "hukuk", " uzak", "ısmarlam", " icra", "ihmal", "ihtar", "iletm", " ilgili", "ilgisiz", " ilişki", " ilkel", "indirim", "influencer", "insan sağlığı", "internet site", "internetten yemek siparişi", " işçi", "işlem", "işletme", "işyeri", " izin", "joker", "kaba ", "kaide", "kalas", "kampanya", "kapalı", "kapıda öde", "kara liste", "kargocu", "karşılı", "kasiyer", "kaşık", "katır", "kaytar", "kazan", "kazık", "kdv", "kebapçı", "kısır", "kıymetli", "kifayetsiz", "kokorecci", "komisyon", "konsept", "kota", "kart", "kullanıcı dost", "kupon", "kural", "kurye", "kurumsal", "küçümseme", "küstah", "lakayıt", "legal", "posta", "maliye", "malzeme", "mekan", "memur", "mendil", "menü", "merkez", "mesafe", "mesele", "meslek", "mevsim", "mezun", "miktar", "min tutar", "misli", "motokurye", "muamele", "muhatap", "mukabil", "mutfak", "mücadele", "müdavim", "müddet", "müşteri", "naçiz", "nazik", "negatif", "nerde kal", "nizam", "nöbet", "numara", "odun", "online", "operatör", "ortak", "otomasyon", "ödeme", "ödev", "öküz", "öner", "örgüt", "özenli", "özensiz", "pahalı", "paket servis", "paket", " pay ", "peçete", "perhiz", "personel", "pipet", "platform", "portör", "pos cihazı", "poşet", "pratik", "promosyon", "prosedür", "puan", "range", "rasgele", "rejim", "reklam", "restoran", "rozet", "rötar", "ruhsat", "rut dışı", "sağanak", "salah", "sapa", "satış", "savsak", "saygılı", "saygısız", "yemekçi", " seri", "server", "servis", "sezon", "sıradan", "sırılsıklam", "sipariş hattı", "sipariş iptal", "sipariş not", "sipariş onay", "sipariş önce", "sipariş sonra", "sistem", "sitem", "soğuk hava", "sorumlu", "sorumsuz", "sorun", "stil", "story", "sunucu", " süre", "şimdi", "şirket", "şube", "taciz", "tahsil", "tahsis", "taksit", "tarih", "tarz", "taşı", "tatbik", "tayfa", "tecrübe", "tehlike", "tek kişi", "teklif", "teknoloji", "telefon", "temas", "temel", "temiz", "temsilci", "terbiyesiz", "ters", "teslim", " tez ","ticket", "titiz", "token", "tolerans", "toplam", "trafik", "trip", "tutar", "tutum", "tüketici", "tükür", "türe", "tüy", "uğraş", "ukala", "ulaşmam", "umursam", "unutmuş", "usta"," usul", "uyar", "uygula", " uyum", "ücret ", "üyeli", "üzücü", "vakit"," vale", "vasıta", "verimli", "verimsiz", "viral", "virüs", "web site", "webchat", "yağmur", "yakışıklı", "yalın", "yanlış", "yanıt", "yaptırım", "yarar", "yasal", "yatkın", "yavşak", "yazılım", "yerel", "yetersiz", "yetkili", "yoğunluk", "yollam", " yol", " yorum", "yoz", "yönetici", "yöntem", "yürütm", "zaman", " zam ", "zararlı", "zihniyet"]

Kredi Tahsisinde Kredi Temerrüdü Tahmini Modeli

Araştırma Makalesi/Research Article

Çiğdem TİMAS[1], Nuri BİNGÖL[2]

[1] Fen Bilimleri Enstitüsü,Yapay Zeka Mühendisliği,Üsküdar Üniversitesi,İstanbul,Türkiye
[2] Fen Bilimleri Enstitüsü, Bilgisayar Mühendisliği, Üsküdar Üniversitesi, İstanbul, Türkiye
cigdem.timas1981@gmail.com,nuri.bingol@uskudar.edu.tr
(Geliş/Received:11.11.2024; Kabul/Accepted:07.03.2025)
DOI: 10.17671/gazibtd.1583269

***Özet*—** Finans sektörünün ana öncüleri olan bankalar ve diğer finans kurumları piyasayı kullandırmış oldukları krediler sayesinde fonlamaktadırlar. Böylelikle fonlamanın sağlıklı yürütülebilmesi, fonlamanın en büyük kalemi olan kredi kullandırma işlevinin doğru yapılması ile sağlanır. Bu da ancak ve ancak tahsis edilen kredilerin bir kısmının ya da tamamının zamanında geri dönmeme olasılığını belirten "kredi riski" durumunun iyi idare edilebilmesi yani geri dönmeyen kredilerin mümkün mertebe en düşük seviyeye indirilmesi ile sağlanabilmektedir. Veri kümesine **Karar Ağacı ,CatBoost ,Yapay Sinir Ağları (YSA)** ve **Random Forest** ve karşılaştırma imkânı elde etme açısından da geleneksel istatistiki yöntemlerden **Lojistik Regresyon (LR)** yöntemleri uygulanarak kredi temerrüdü tahminleri yapılmıştır. (Özet)

***Anahtar Kelimeler*—** temerrüd, tahmin modeli ,kredi tahsisi

Credit Default Prediction Model in Credit Allocation

***Abstract*—** The main pioneers of the finance sector, banks and other financial institutions, fund the market through the loans they provide. Thus, the healthy execution of funding is ensured by correctly performing the loan provision function, which is the largest item of funding. This can only be achieved by managing the "credit risk" situation, which indicates the possibility of not returning some or all of the allocated loans on time, in other words, by reducing the non-performing loans to the lowest possible level. Credit default estimates were made by applying **Decision Tree**, **CatBoost**, **Artificial Neural Networks (ANN)** and **Random Forest** to the data set and **Logistic Regression (LR)** methods, which are among the traditional statistical methods, for comparison purposes.(Abstract)

***Keywords*—** default, prediction model, credit allocation

1. GİRİŞ (INTRODUCTION)

Bankalar tarafından tahsis edilen kredilerin, müşteri tarafından geri ödenememesi veya zamanında geri ödenmemesi hem kredi veren işletmenin sermaye kaybını hem de genel ekonomide oluşabilecek çeşitli riskli durumları beraberinde getirmektedir. En başta Bankaların yer aldığı finans dünyası, ekonomilerin devamlılığı açısından stratejik ve çok önemli bir işlev olan, "fonlama" işlevini yerine getiren, kritik derece de öneme sahip bir sektör niteliğindedir. Finans sektörünün ana öncüleri olan bankalar, diğer finans kurumları piyasayı kullandırmış oldukları krediler sayesinde fonlamaktadırlar. Böylelikle fonlamanın sağlıklı yürütülebilmesi, fonlamanın en büyük kalemi olan kredi kullandırma işlevinin doğru yapılması ile sağlanır. Bu da ancak ve ancak tahsis edilen kredilerin bir kısmının ya da tamamının zamanında geri dönmeme olasılığını belirten "kredi riski" durumunun iyi idare edilebilmesi yani geri dönmeyen kredilerin mümkün mertebe en düşük seviyeyeye indirilmesi sağlanabilmektedir. Kredi başvurusunda bulunan bu müşterilerden hangilerinin riskli kategorisine girmeyip kredi başvurusunun olumlu yanıt dönülmesi ve kredinin tahsis edilmesi, hangilerinin ise riskli bulunup taleplerinin reddedilmesi gerektiği anlaşılmaya çalışılmıştır. Belirtildiği gibi bu sorun, müşterileri "riskli" ve "risksiz" şeklinde iki sınıfa ayırma, yani bir çeşit sınıflandırma problemidir.Çalışmada veri kümesine Karar Ağacı ,CatBoost ,Yapay Sinir Ağları ve Random Forest ve karşılaştırma imkânı elde etme açısından da geleneksel istatistiki yöntemlerden Lojistik Regresyon yöntemleri uygulanmıştır. Birden fazla makine öğrenmesi modeli test edilmiş ve performansları karşılaştırılmıştır.Kredi temerrüdü tahmin modelinin performansını değerlendirmek için Confusion Matrix, ROC-AUC skoru ve Classification Report metriklerini kullandık. Confusion Matrix, modelin doğru ve yanlış tahminlerini analiz ederek False Negative (FN) ve False Positive (FP) değerlerini inceledik. ROC-AUC skoru, modelin sınıflandırma başarısını ölçerek temerrüde düşenleri ne kadar iyi ayırt ettiğini değerlendirdik. Ayrıca, Precision (Kesinlik), Recall (Duyarlılık) ve F1-Score gibi sınıflandırma metrikleriyle modelin yanlış pozitif ve yanlış negatif tahminleri arasındaki dengesini analiz ettik.Accuracy tek başına yeterli olmadığından, özellikle Recall ve AUC-ROC değerlerini karşılaştırarak en iyi modelin bulunması hedeflenmiştir.Çalışma makine öğrenmesiyle kredi risk tahmininde geleneksel lojistik regresyon yerine CatBoost, Random Forest gibi güçlü modeller de kullanılarak, literatürdeki klasik yöntemlere kıyasla bu modellerin de başarılı performanslar sunabileceğini göstermesi yönünden önemlidir.

2. LİTERATÜR (LITERATURE REVIEW)

Literatürde kredi riski tahminine yönelik çeşitli çalışmalar yapılmıştır. Bunlardan biri olan Yeh [1], Tayvan'daki kredi kartı kullanıcılarının temerrüt etme olasılığını, altı farklı veri madenciliği yöntemi ile analiz ederek bu yöntemlerin doğruluk oranlarını karşılaştırmıştır. Bu çalışmada değerlendirilen yöntemler; diskriminant analizi, lojistik regresyon, Bayes sınıflandırıcısı, en yakın komşu algoritması, yapay sinir ağları ve sınıflandırma ağaçlarıdır. Elde edilen sonuçlara göre, bu teknikler arasında en iyi performansı yapay sinir ağları göstermiş ve kredi puanlama uygulamalarında etkili bir yöntem olarak öne çıkmıştır. [1]

Budak ve Erpolat çalışmalarında [2] ,bankaların kredi risklerini en aza indirmek amacıyla, kredi başvurusunda bulunan müşterilerin ödeme potansiyellerini analiz etmek için yapay sinir ağları ve lojistik regresyon analizi karşılaştırılmıştır. Çalışmanın sonuçlarına göre, yapay sinir ağları yöntemi, müşterilerin ödeme alışkanlıklarının düzenli olup olmadığını tahmin etmede lojistik regresyon modeline kıyasla daha yüksek bir başarı sergilemiştir. [2]

Kavcıoğlu [3]'nun çalışmasında, kredi riskinin ölçümünde kullanılan modellerin genel olarak birbirine benzediği, ancak farklı değerlendirme kriterleri ve ölçüm yöntemleri içerdiği vurgulanmıştır. Bu modellerin bazı yönlerden birbirlerine avantaj sağladığı, ancak bazı eksikliklerinin de bulunduğu ifade edilmiştir. Bu sebeplerle, özellikle bankalar gibi kredi veren kurumların risklerini en aza indirebilmeleri için kendi yapılarına en uygun modeli ve yöntemi seçmeleri gerektiği belirtilmiştir. [3]

Hamori ve çalışma arkadaşları [4], Tayvan'daki temerrüt ödeme verilerini analiz ederek üç toplu öğrenme yöntemi (torbalama, rastgele orman ve artırma) ile çeşitli sinir ağı yöntemlerinin tahmin doğruluğu ve sınıflandırma yeteneğini karşılaştırmıştır. Çalışmanın sonuçlarına göre, güçlendirme yönteminin sınıflandırma başarısının, sinir ağları da dahil olmak üzere diğer makine öğrenimi yöntemlerine göre daha yüksek olduğu ortaya çıkmıştır. Ayrıca sinir ağı modellerinin performansının, seçilen aktivasyon fonksiyonuna ve gizli katman sayısına göre değişiklik gösterdiği bulunmuştur. [4]

İldaş [5]'ın çalışmasında, kredi riskinin ölçümüne yönelik birçok farklı yöntemin mevcut olduğu, klasik ve modern kredi risk ölçüm metotlarının yanı sıra, portföy kredi riskinin değerlendirilmesi amacıyla uluslararası finansal kuruluşlar tarafından geliştirilen modellerin de kullanılmasının önemli olduğu vurgulanmıştır. [5]

Mahbobi, Kimiagari ve Vasudevan [6] çalışmalarında, entegre bir tahmin doğruluğu algoritması geliştirirken, çalışmada makine öğrenimi sınıflandırıcıları olarak DNN (Derin Sinir Ağı), SVM (Destek Vektör Makineleri), KNN (En Yakın Komşu) ve ANN (Yapay Sinir Ağı) kullanılmıştır. 30.000 dengesiz veri kümesinden yararlanılarak, temerrüt ödemelerinin tahmin doğruluğunu artırmak amacıyla Sentetik Azınlık Aşırı Örnekleme Tekniği (SMOTE), SVM SMOTE, rastgele düşük örnekleme ve ALL-KNN gibi çeşitli aşırı ve yetersiz örnekleme stratejileri uygulanmıştır. Çalışmanın sonuçlarına göre, ALL-KNN örnekleme tekniği kullanılarak yapılan analizde, SVM modelinin %98,6 doğruluk ve en düşük çapraz entropi kaybı değeri olan 0,028 ile en iyi performansı sergilediği gözlemlenmiştir. [6]

Tütüncü ve Gürsakal [7], çalışmalarında kredi skorlama sistemlerinde en başarılı tahmini yapan algoritmayı belirlemek amacıyla makine öğrenmesi yöntemlerini kullanmışlardır. Gradyan Artırma, Yapay Sinir Ağları, Lojistik Regresyon, Rastgele Orman, Karar Ağacı, Destek Vektör Makineleri, K-En Yakın Komşu ve WOE dönüşümleriyle Lojistik Regresyon algoritmaları için modeller oluşturmuşlardır. Çalışmanın sonuçlarına göre, temerrüde düşen ve düşmeyen müşterilerin en iyi sınıflandırmasını sağlayan algoritmanın Gradyan Artırma olduğu tespit edilmiştir. [7]

Milli [8], tez çalışmasında riskli müşterilerin belirlenmesi için K-En Yakın Komşu, Naive Bayes, Lojistik Regresyon, Destek Vektör Makineleri, Çok Katmanlı Algılayıcı, Karar Ağaçları, Rastgele Ormanlar, Gradyan Artırma Karar Ağaçları, Extra Ağaçlar, Sert ve Yumuşak Oylama gibi çeşitli makine öğrenmesi sınıflandırma tekniklerini kullanmıştır. Ayrıca sınıf dengesizliği sorununu çözmek amacıyla, RUS (Random UnderSampling), ROS (Random OverSampling), SMOTE-ENN, SMOTE-Tomek Links ve Balanced Bagging Classifier gibi yeniden örnekleme ve hibrit yöntemlerle veri dengesini sağlamıştır. Üç farklı kredi veri kümesi üzerinde dört farklı senaryo ile yapılan çalışma sonucunda, sınıf dengesizliği için alt ve üst örnekleme yöntemlerinin bir arada kullanıldığı hibrit yöntemin, makine öğrenmesi tekniklerinde en iyi sınıflandırma performansına ulaşmada etkili olabileceği sonucuna varılmıştır. [8]

Zhu ve çalışma arkadaşlarının [9] yaptığı araştırmada, kredi temerrüdünü tahmin etmek amacıyla lojistik regresyon, karar ağacı, XGBoost ve LightGBM modelleri kullanılmıştır. Elde edilen sonuçlar, LightGBM ve XGBoost'un tahmin performansının lojistik regresyon ve karar ağacı modellerine göre daha yüksek olduğunu göstermiştir. LightGBM modeli için eğri altındaki alan 0,7213 olarak bulunmuş ve her iki modelin doğruluğu 0,8'in üzerinde çıkmıştır. Ayrıca, LightGBM ve XGBoost'un hassasiyetinin 0,55'in üzerinde olduğu belirlenmiştir. Sonuçlara göre, kredi vadesi, kredi notu, kredi ratingi ve kredi tutarı gibi değişkenlerin tahmin sonuçları üzerinde etkili olduğu görülmüştür. [9]

Literatür incelendiğinde, farklı makine öğrenmesi algoritmalarının kredi riskinin ölçümünde en iyi sonuçları verdiğine dair genel bir fikir birliği olmadığı görülmektedir. Bu araştırmayla Karar Ağacı ,CatBoost ,Yapay Sinir Ağları,Random Forest ve Lojistik Regresyon yöntemleriyle sınıflandırma performansları karşılaştırılarak kredi riski ölçümünde en iyi sonucu veren yöntemin bulunması hedeflenmiştir.

3. YÖNTEM (METHOD)

Veri kümesi Lending Club'a ait açık erişimde olan 27 değişkenden 396030 gözlemden oluşan veri setidir. Veri kümesi geçmişte kredi başvurusunda bulunanlar ve bunların temerrüt edip etmediğine ilişkin bilgileri içermektedir. Bu tarz modellerin kullanılmasındaki amaç, tekrardan kredi kullandırıldığında bir kişinin temerrüde düşme ihtimalinin olup olmadığını tespit edip ya krediyi reddetmek, ya da talep edilen kredi miktarından daha az tutarda bir kredi onaylamak, ya da riskli müşterilere daha yüksek faiz oranıyla borç vermek ya da direkt başvuruyu reddetmek şeklinde olabilir.

Araştırma kapsamında, Python'un sunduğu scikit-learn, pandas, numpy, matplotlib ve seaborn gibi güçlü kütüphanelerden faydalanılarak, geniş çaplı veri analizi ve makine öğrenmesi algoritmaları uygulanmıştır.Ham verinin temizlenmesi ve eksik değerlerin uygun yöntemlerle doldurulması sağlanmış, ardından değişkenlerin istatistiksel dağılımları detaylı bir şekilde incelenmiştir. Korelasyon analizleri ve değişken seçimi teknikleri kullanılarak modele en uygun özellikler belirlenmiş, böylece tahmin gücü yüksek ve doğruluğu artırılmış bir model oluşturulmuştur.

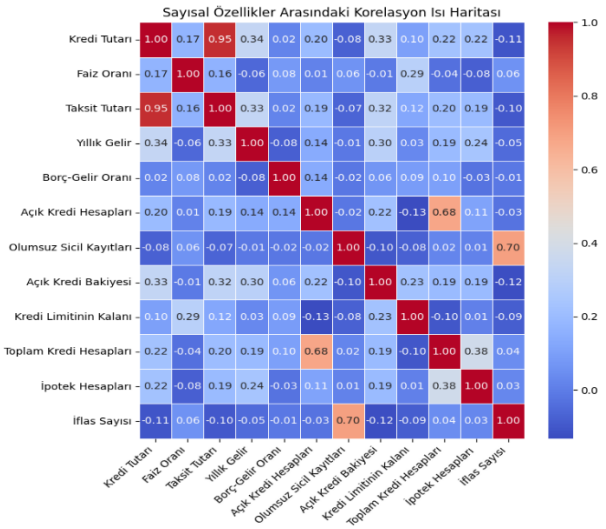
Araştırmada, lojistik regresyon, karar ağaçları, rastgele ormanlar ve yapay sinir ağları gibi farklı makine öğrenmesi algoritmaları karşılaştırılmış ve performans ölçütleri dikkate alınarak en iyi sonuç veren model seçilmiştir. Modelin başarı oranını değerlendirmek amacıyla, doğruluk (accuracy), kesinlik (precision), duyarlılık (recall) ve F1 skoru gibi çeşitli metrikler hesaplanmış, ayrıca ROC eğrisi ve AUC skoru ile modelin genel performansı görselleştirilmiştir.

Python'un esnek ve güçlü analiz araçları kullanılarak yapılan bu modelleme süreci, yalnızca veri setinden anlamlı içgörüler çıkarmakla kalmamış, aynı zamanda geleceğe yönelik doğru tahminlerin yapılmasını mümkün kılmıştır. Araştırmada geliştirilen bu metodoloji, veri bilimi alanında çalışan araştırmacılar ve sektör profesyonelleri için yol gösterici bir rehber niteliği taşımaktadır.

4. BULGULAR (FNDINGS)

4.1 Veri Analizi (Data Analysis)

Veri kümesinde kredi modelinin belirlenebilmesi için birçok değişken bulunmaktadır. Kredi faiz oranı, kredi miktarı, kredi notu, borçlunun kamu kayıtlarındaki iflas geçmişi, meslek bilgisi ve ev sahipliği durumu gibi çeşitli değişkenler yer almaktadır. Sayısal değerler arasındaki korelasyonu analiz etmek amacıyla ısı haritası kullanılmış ve kredi tutarı ile taksit değerleri arasında 0.95 gibi yüksek bir pozitif korelasyon katsayısı olduğu tespit edilmiştir.



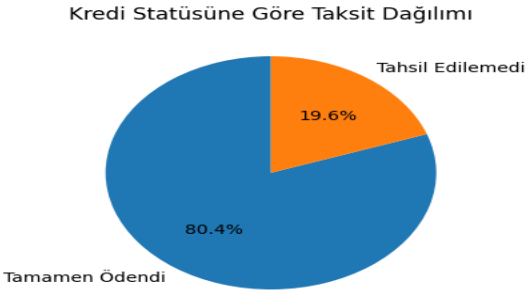
Şekil 1.Sayısal Özellikler Arasındaki Korelasyon Isı Haritası (*Correlation Heat Map Between Numerical Features)*

Veri setinde eksik değerlerin (missing values) dağılımına bakıldığında meslek bilgisi ,meslekte geçen süre,mortgage hesapları değişkenlerinde fazlaca sayıda eksik bilgi olduğu görülmüştür. Çok sayıda iş unvanının bulunması , iş unvanı değişkenindeki eksik verilerin sayıca fazla olması(22927) ve değişkenin modele katkısının çok olmayacağını değerlendirilerek veri setinden çıkarılmıştır. Aynı pozisyondaki geçen yıl sayısını gösteren bir fonksiyon döngüsü yazdırılarak yüzdelik oranları incelenmiş hemen hemen hepsinin yakın değerler olduğu analiz edilmiştir.Meslekte geçen çalışma sürelerinin de modele katkı sağlamayacağı değerlendirilerek veri setinden çıkarılmıştır.Kredi başvuru amacı değişkeni incelendiğinde kredi temerrüdü tahmini modelinin amacına uygun bir girdi sağlamadığı değerlendirilmiştir. Mortgage hesapları kredi için başvuran adayın ipotek hesaplarının sayısını gösteren bir değişken olup bu değişkendeki eksik değerler(37795) de sayıca fazladır.Mortgage hesapları değişkeni,kamu kayıtlarındaki iflas değişkeni ve kamu kayıtları değişkenindeki eksik değerler için fonksiyon oluşturulmuş ve fonksiyonla ilgili değişkenlerin kategorik(binary - ikili formata) hale getirilmesi sağlanmıştır. Kamu Kayıtlarındaki İflas değişkeni Sabit Değer ile Doldurma (Constant Imputation) yöntemi ile eksik değerlere sıfır atanarak doldurulmuştur. Ortalama ile Doldurma (Mean Imputation) yöntemiyle de Mortgage hesapları değişkenindeki eksik veriler doldurulmuştur.Kredi başvuru adayının adres değişkeni ,kredinin finanse edildiği ay bilgisi ve kredi kulübünün verdiği not değişkenleri veri setinden çıkarılmıştır.Kredi durumu kategorik değişkenleri 0 ve 1 sayısal değerlerine dönüştürülerek modelin işleyebileceği duruma dönüştürüldü.Bağımlı değişken; kredi statüsü (loan_status) değişkeni 1 kredi ödemelerinde gecikmesi(temerrüd) olan , 0 ise kredi ödemelerinde gecikmesi olmayan kredisini zamanında ödeyeni ifade etmektedir.

Şekil 1'den Şekil 9'a kadar olan görseller, yazarlar tarafından model oluşturulmadan önce veriyi tanımak ve modele hazırlık yapmak amacıyla gerçekleştirilen veri analizi çalışmalarına dayanmaktadır.

4.1.1 Kredi Statüsüne Göre Taksit Dağılımı (Distribution of Installment by Credit Status)

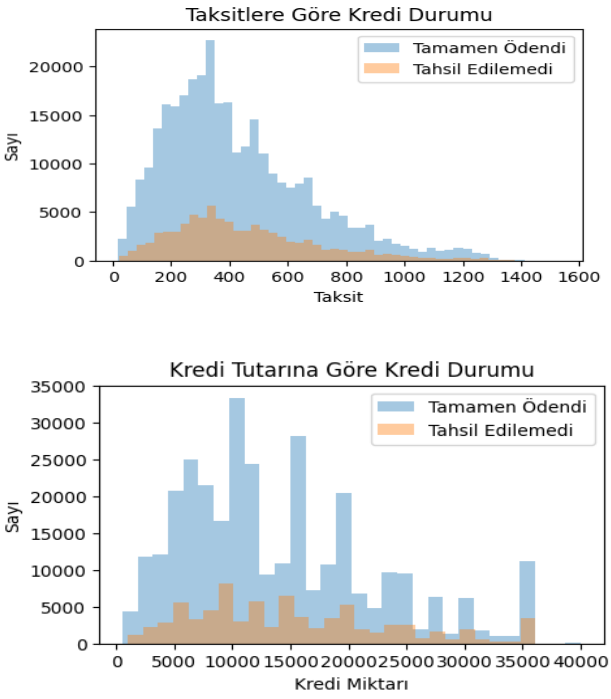
Veri kümesinde kredi statülerine göre, 77673 adedinde 19.6%'sında kredisi temerrüde düşmüş(Tahsil edilemedi),318357 adedinde 80.4 % 'sında kredisi tahsil edilmiş (Tamamı Ödendi) olarak veriler dağılım göstermektedir.



Şekil 2. Kredi Statüsüne Göre Taksit Dağılımı*(Installment Distribution According to Loan Status)*

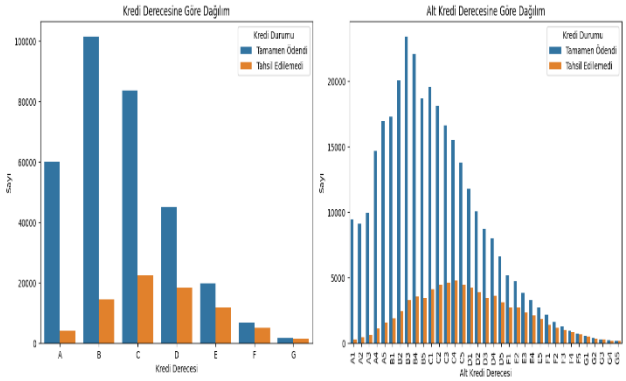
4.1.2 Taksit ve Kredi Tutarı &Kredi Durumu Grafiği (Installment and Loan Amount & Loan Status Chart)

Taksit tutarı ve kredi tutarı değişkenlerine göre kredi durumlarının görüntülenmesi sağlanmıştır. Histogramlarda taksit tutarı ve kredi tutarı arttıkça geri ödeme durumunun azaldığı görülmektedir. Kısa vadede ve daha düşük tutardaki kredilerin daha çok tahsil edilebildiği anlaşılmaktadır.



Şekil 3.Taksit ve Kredi Tutarı &Kredi Durumu Grafiği(*Installment and Loan Amount and Loan Status Graph)*

4.1.3 Derece(grade) ve Alt derece(sub_grade) Kredi Geri Ödeme Durumuna Göre Dağılımı (Distribution of Grade and Sub-grade by Loan Repayment Status)



Şekil 4.Derece ve Alt derecenin Kredi Geri Ödeme Durumuna Göre Dağılımı *(Distribution of Degree and Sub-degree According to Loan Repayment Status)*

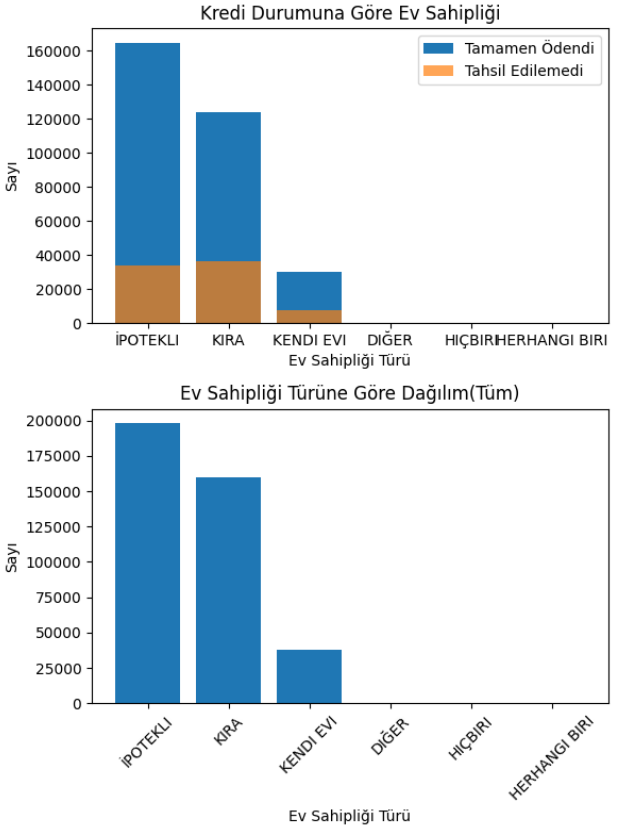
Derece: Borçlunun genel kredi güvenilirliğini veya risk seviyesini gösteren geniş bir kategoridir. Genellikle bir finans kurumu veya kredi derecelendirme kuruluşu tarafından borçlunun kredi skoru, finansal geçmişi ve diğer faktörlere dayanarak atanır. Harflerle ifade edilir (örneğin, A, B, C), burada "A", yüksek kredi skoruna sahip, düşük riskli bir borçluyu temsil ederken, "C" ve altı, daha yüksek riskli borçluları ifade eder. Yani, borçlunun krediyi geri ödeme olasılığını belirlemek için kullanılır.

Alt Derece: Derece içinde daha ayrıntılı bir sınıflandırmadır. Aynı derecedeki borçlular arasındaki risk farklarını daha net bir şekilde ayırt eder. Örneğin, "A" derecesi altında "A1", "A2", "A3" gibi alt dereceler olabilir. "A1", o kategorideki en güvenilir borçluyu temsil ederken, "A3" biraz daha fazla risk taşıyan borçluyu ifade eder. Bu alt dereceler, aynı kredi derecesine sahip borçlular arasında bile risk farklılıklarını ortaya koyar.

Derece ve altderece sütunlarındaki verilerin kredi ödeme durumu değişkenine göre nasıl dağıldığını gösterir. F ve G alt sınıflarının düzenli olarak geri ödenmediği görülmekle birlikte ,A,B ve C sınıflarının geri ödeme durumlarının başarılı olduğu görülmektedir.

4.1.4.Ev Sahipliği Durumu ile Kredi Durumu Arasındaki İlişki (The Relationship Between Home Ownership Status and Loan Status)

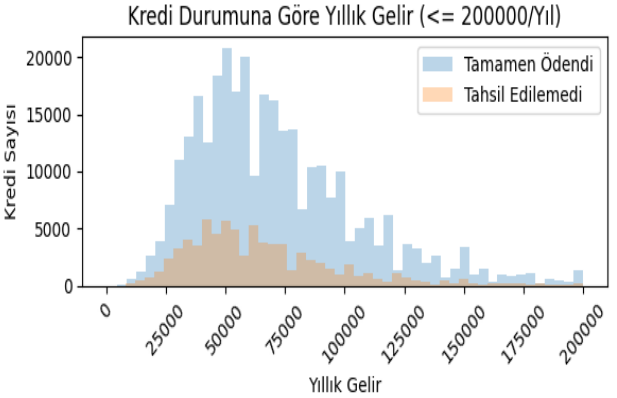
Kredi durumu ve ev sahipliği durumu değişkenleri arasındaki ilişkiyi görsel olarak incelemek ve ev sahipliği durumlarına göre kredi durumlarını karşılaştırmak amaçlanmıştır. Tamamı ödenmiş ve Ödenmemiş borç (Tahsil edilemedi) durumlarına sahip olanlar arasında ev sahipliği durumlarının dağılımına bakıldığında ,Tamamının ödeme durumunun ipotekli(Mortgage) ve kirada ikamet edenler arasında yoğunlaştığı görülmektedir.



Şekil 5. Ev Sahipliği Durumu ile Kredi Durumu Arasındaki İlişki*(The Relationship Between Homeownership Status and Loan Status)*

4.1.5.Yıllık Geliri <200.000 $ Olanlarda Kredi Ödeme Durumu(Loan Payment Status for Annual Income < $200,000)

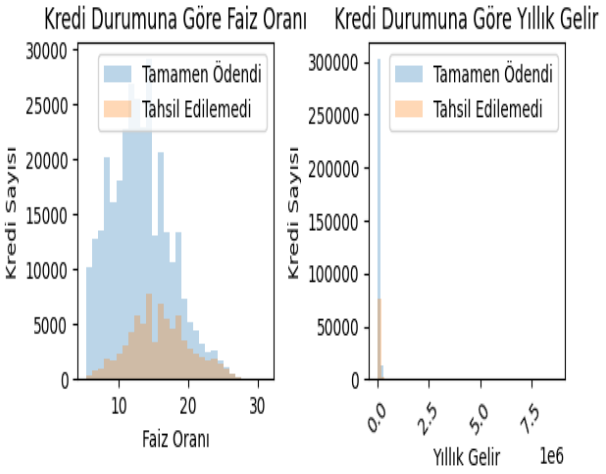
Yıllık geliri 200.000 $ veya daha az olan verilere göre kredi ödeme durumu kategorisine göre gruplanıp bir histogram grafiği oluşturuldu.Grafiğe bakıldığında büyük dağılımın yıllık geliri 100.000 $ ve altında olanlarda olduğu görülmektedir.



Şekil 6. Yıllık Geliri <200.000 $ Olanlarda Kredi Ödeme Durumu*(Loan Payment Status for Annual Income < $200,000*

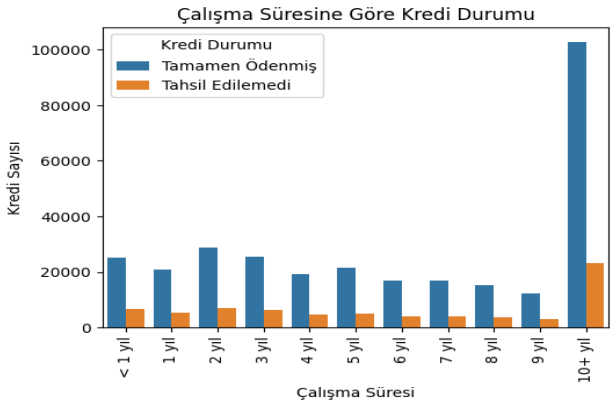
4.1.6 Faiz Oranı ile Yıllık Gelir ve Kredi Durumu Arasındaki İlişki (The Relationship Between Interest Rate, Annual Income, and Loan Status)

Faiz oranı ve yıllık gelir nitelikleri kredi durum değişkenine göre gruplanarak bir histogram çizdirilmesi sağlandı.Histograma göre faiz oranı arttıkça, kredi kullandırımının azaldığı ve temerrüd oranının arttığı tespit edilmiştir.



Şekil 7. Faiz Oranı ile Yıllık Gelir ve Kredi Durumu Arasındaki İlişki*(The Relationship Between Interest Rate, Annual Income, and Loan Status)*

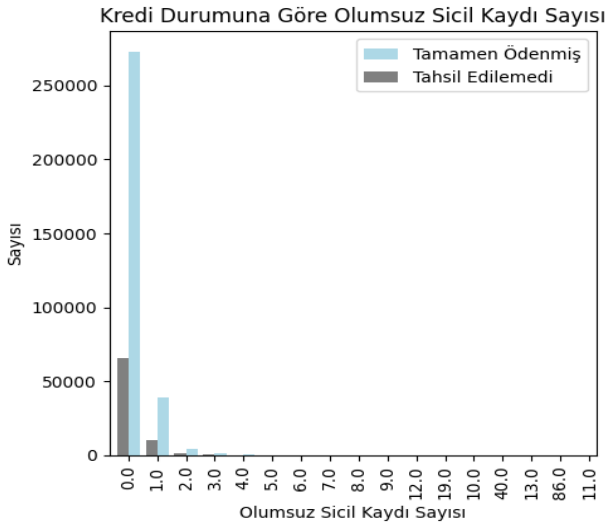
4.1.7 Meslekteki Tecrübe ve Kredi Durumu İlişkisi (Relationship Between Professional Experience and Loan Status)



Şekil 8.Meslekteki Tecrübe ve Kredi Durumu İlişkisi*(Relationship Between Professional Experience and Loan Status)*

Kredi başvurularında 10 yıl ve üzeri mesleki deneyime sahip bireylerin geri ödeme performansında belirgin bir pozitif farklılık gözlemlenmiştir.

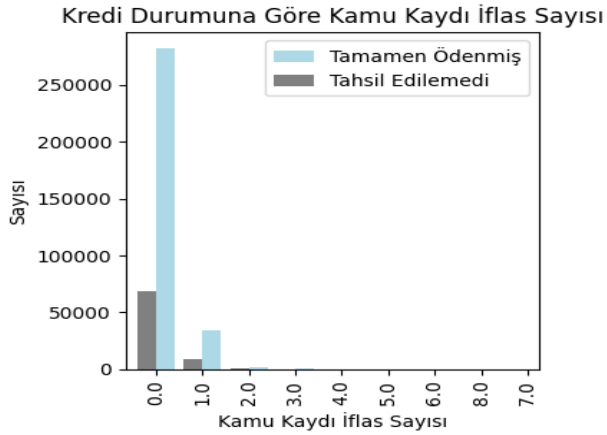
4.1.8 Olumsuz Sicil Kaydı ve Kredi Durumları (Negative Credit Record and Loan Status)



Şekil 9. Olumsuz Sicil Kaydı ve Kredi Durumları*(Negative Credit Record and Loan Status)*

Olumsuz sicil kaydı , bir borçlunun itibarını olumsuz etkileyen kamuya açık kayıtların sayısını ifade eder. Bu kayıtlar genellikle kredi geçmişi, borçların ödenmesi, icra takibi gibi finansal durumlarıyla ilgili olabilir.Bankaların kredi kullandırım kararı esnasında müşterinin Memzuç kayıtlarına ,KKB kayıtlarına bakarak karar vermesi örnek verilebilir. Bu olumsuz kayıtlar genellikle kredi notunun düşük olmasına neden olabilir ve finansal geçmişin incelenmesinde ve kredi tahsis kararının verilmesinde önemli bir etken olabilir. Geçmişte olumsuz sicil kaydı bulunmayanlarda gecikmeye düşme durumu gözle görülür derecede azdır. Fakat yine grafiğe bakıldığında geçmişte olumsuz sicil kaydı olmamasına rağmen ,kredisi temerrüde düşenlerin sayısı ,daha önce olumsuz kaydı olanlara göre fazladır.

4.1.9 Kamu Kayıtlarındaki İflas Kaydı & Kredi Durumları (Bankruptcy Records in Public Records & Loan Status)

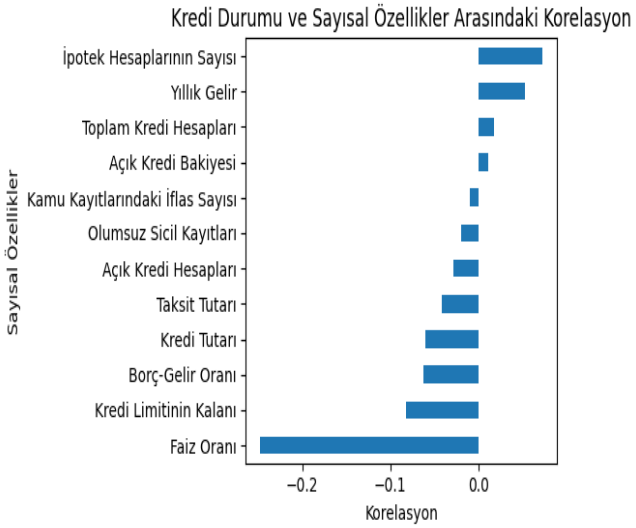


Şekil 10.Kamu Kayıtlarındaki İflas Kaydı & Kredi Durumları*(Bankruptcy Records in Public Records & Loan Status)*

Genel olarak başvuran kişinin kamu kayıtlarındaki iflas kaydı sayısının yüksek olması kredi geçmişinin olumsuz durumu olarak değerlendirilir ve kredi notunun düşmesine sebep olabilir. Çünkü bu tür iflaslar, kredi verenler için

riskli bir durum olarak kabul edilir ve gelecekteki kredi başvuruları için olumsuz bir etkiye sahip olur. Bu sebeple önemli bir değişkendir.Grafiğe bakıldığında geçmişte kamuda iflas kaydı olmayan müşterilerin kredi ödemelerinde büyük çoğunluğunda sorun olmadığı görülmüştür. Fakat yine grafiğe bakıldığında geçmişte kamuda iflas kaydı olmamasına rağmen ,kredisi temerrüde düşenlerin sayısı ,daha önce olumsuz kaydı olanlara göre fazladır.

4.1.10 Korelasyon - Kredi Durumu ve Sayısal Özellikler(Correlation Between Loan Status and Numerical Characteristics)



Şekil 11.Kredi Durumu ve Sayısal Özellikler Arasındaki Korelasyon*(Correlation Between Credit Status and Numerical Characteristics)*

Korelasyona bakıldığında faiz oranı ile kredinin ödenme durumu arasında negatif ve güçlü bir korelasyon olduğu görülmektedir. Faiz oranı arttıkça kredinin ödenme durumu da zorlaşmaktadır. İpotek hesaplarının sayısı ve yıllık gelir ile kredi durumları arasında da pozitif bir korelasyon olduğu görülmektedir. Bu değerler 1 e doğru yaklaştığı için aralarında güçlü bir ilişki olduğu görülmektedir.

4.2 Model Geliştirme(Model Development)

Veri analizi sonucunda belirlenen nihai öznitelik değişkenleri kullanılarak kredilerin temerrüt riski aşağıdaki algoritmalar ile belirlenmeye çalışılmış ve kullanılan algoritmaların sınıflandırma performansları karşılaştırılmıştır.

Makine öğrenmesi modeli oluşturmak için veri bölme, ölçeklendirme ve değişken ayrıştırma işlemleri gerçekleştirilmiştir. Veri kümesi %67 eğitim (train) ve %33 test (test) olmak üzere ikiye ayrıldı ve random_state=42 ile sonuçların her seferinde aynı olması için rastgelelik sabitlendi. Bağımlı değişken kredi durumu (loan_status) değişkeni, kredinin ödenip ödenmediğini gösteren hedef değişken olup kredi miktarı, faiz oranı,yıllık gelir,uygunsuz kamu sicil kayıtları ,açık kredi kart bakiyesi gibi değişkenler bağımsız değişkenler olarak ayrılmıştır. Min-Max Ölçeklendirme (Feature Scaling) gerçekleştirilmiş olup MinMaxScaler() ile tüm bağımsız değişkenleri 0 ile 1 arasına ölçeklendirildi.

Sınıflandırma modelinin tahmin performansını değerlendirmek için Confusion(karışıklık) matrisi kullanıldı.Böylelikle modelin doğru ve yanlış sınıflandırma sayıları gösterildi. Matris, tahminleri daha açık bir şekilde gösterir ve modelin performansının anlaşılmasına yardımcı olur.

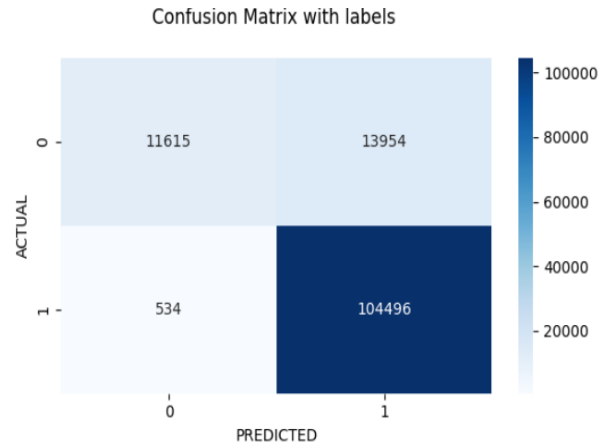
Karışıklık matrisi dört ana hücreye bölünür ve şu dört kategoriyi içerir:

1. **True Positive (TP)**: Modelin doğru bir şekilde pozitif olarak sınıflandırdığı örnekler.
2. **False Positive (FP)**: Modelin yanlış bir şekilde pozitif olarak sınıflandırdığı örnekler (pozitif olarak tahmin edilen ama gerçekte negatif olan örnekler)
3. **True Negative (TN)**: Modelin doğru bir şekilde negatif olarak sınıflandırdığı örnekler.
4. **False Negative (FN)**: Modelin yanlış bir şekilde negatif olarak sınıflandırdığı örnekler (negatif olarak tahmin edilen ama gerçekte pozitif olan örnekler)

Araştırmada ikili sınıflandırma problemlerinde, modelin pozitif sınıfı ne kadar iyi ayırt edebildiğini değerlendirebilmek için ROC (Receiver Operating Characteristic) eğrisi kullanılmıştır. ROC eğrisi, farklı eşiklerdeki performansı görmeye imkan tanımaktadır.Modellerin her birinin sınıflandırma raporu oluşturularak performans değerleri gösterilmiştir. Sınıflandırma raporu, genellikle her bir sınıf için hassasiyet, geri çağırma ve F1 skoru gibi değerleri gösterir. Bu değerler, modelin her bir sınıfı ne kadar doğru ve kapsamlı bir şekilde tahmin ettiğini gösterir. Bu rapor, modelin performansını daha ayrıntılı bir şekilde anlaşılmasına yardımcı olur ve sınıflandırma problemiyle ilgili değerli içgörüler sağlar.

4.2.1 Logistik Regresyon(Logistic Regression)

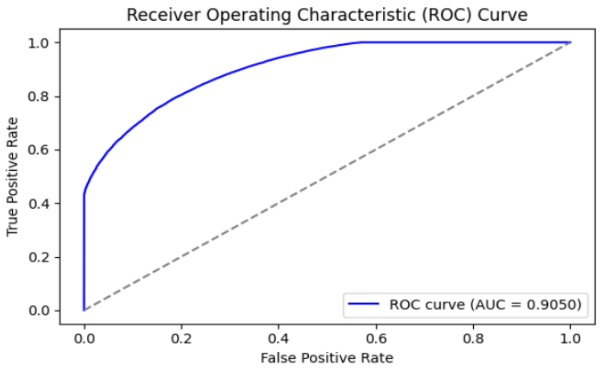
Scikit-learn kütüphanesinden LogisticRegression (lojistik regresyon modeli) ,confusion_matrix (karmaşıklık matrisi) ve classification_report(sınıflandırma raporu) temin edilerek kullanılmıştır.Seaborn ve matplotlib kütüphaneleri, veri görselleştirme amacıyla kullanılmıştır. LogisticRegression(solver='liblinear') fonksiyonu ile lojistik regresyon modeli oluşturulmuştur. Determinasyon katsayısı (R^2 skoru) hesaplanarak modelin eğitim seti üzerindeki performansı değerlendirilmiştir. Test verisi üzerinden tahminler yapılmış ve tahmin edilen değerlerin dağılımı analiz edilmiştir. Gerçek değerler ile tahmin edilen değerlerin karşılaştırılması için confusion matrix oluşturulmuştur.



Şekil 12.Logistik Regresyon Matrisi(*Logistic Regression Matrix)*

Lojistik Regresyon modelinin matrisdeki değerleri incelendiğinde 534 yanlış negatif (False Negative) tahmin yapıldığını.ve oldukça düşük bir değer olduğu için modelin genellikle kredi ödememe ihtimali olanları doğru tespit edebildiğini göstermektedir.

ROC eğrisi ve AUC (Area Under Curve) skoru hesaplanmıştır. Predict_proba() fonksiyonu ile olası tahminlerin pozitif sınıfa ait olasılıkları elde edilmiştir. AUC skoru, modelin tahmin doğruluğunu özetleyen önemli bir metriktir ve 1'e ne kadar yakınsa model o kadar başarılıdır.



Şekil 13.Logistik Regresyon ROC Eğrisi *(Logistic Regression ROC Curve)*

Modelin doğruluk(accuracy) skoru 0.89, Auc-Roc skoru 0.91 olup bu performans sonuçlarıyla güçlü bir tahmin yeteneğine sahip olduğunu göstermektedir.

Sınıflandırma raporu ile modelin hassasiyet (precision), duyarlılık (recall), F1 skoru gibi temel metrikleri hesaplanmıştır. Sınıf 0 için precision yüksek (0.96) ancak recall düşük (0.45) olarak gözlenmiştir. Bu, modelin gecikmesi olmayanları doğru tahmin etme eğiliminde olduğunu ancak bunları tam olarak yakalayamadığını gösteriyor. Sınıf 1 için oldukça iyi performans göstermiştir (f1-score = 0.94). Bu model, gecikmesi olanları iyi tahmin ederken gecikmesi olmayanları kaçırıyor olarak değerlendirilebilmektedir.

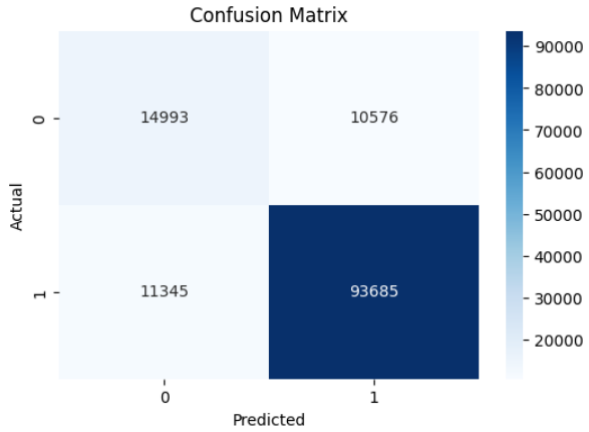
```
Classification Report:
              precision    recall  f1-score   support

           0       0.96      0.45      0.62     25569
           1       0.88      0.99      0.94    105030

    accuracy                           0.89    130599
   macro avg       0.92      0.72      0.78    130599
weighted avg       0.90      0.89      0.87    130599
```

Şekil 14.Logistik Regresyon Modeli Sınıflandırma Raporu*(Logistic Regression Model Classification Report)*

4.2.2 Karar Ağacı (Decision Tree)

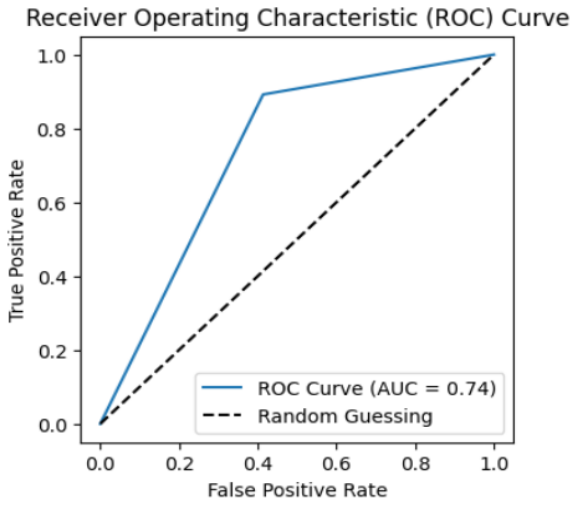
Karar Ağacı (Decision Tree) algoritması kullanılarak bir sınıflandırma modeli oluşturulmuş ve performans değerlendirmesi yapılmıştır. Model, ID3 (Entropy) kriteri ile eğitilmiş ve test verileri üzerinde tahminleme gerçekleştirilmiştir. Scikit-learn (sklearn) karar ağacı modeli (DecisionTreeClassifier) ve model başarım metrikleri kullanılmıştır.

Karar ağacı modelinin matrisdeki değerleri incelendiğinde pozitif sınıfları tespit etmede yani düşük False Negatif ve yüksek False Pozitif tahminlerinde Logistik regresyon modeline kıyasla daha başarılı olduğu gözlenmiştir.Ancak negatif sınıfları yanlış pozitif olarak tahmin etme eğilimi daha yüksektir.



Şekil 15.Karar Ağacı Matrisi(*Decision Tree Matrix)*

Karar ağacı modelinin performans sonucuna göre accuracy skoru 0.83 ,ROC Auc Score değeri 0.74 ‘dür.Logistik regresyon modeline kıyasla daha düşük performans sonuçları elde edilmiştir.



Şekil 16.Karar Ağacı ROC Eğrisi *(Decision Tree ROC Curve)*

Sınıflandırma raporuna göre sınıf 0 için precision ve recall dengeli ama düşük. Sınıf 1 için f1-score biraz düşük olsa da (0.90), recall biraz daha dengeli görülmektedir. Gecikmesi olmayanları Logistic Regression'a kıyasla biraz daha iyi yakalıyor, ancak genel doğruluk açısından Logistic Regression kadar güçlü görülmemiştir.

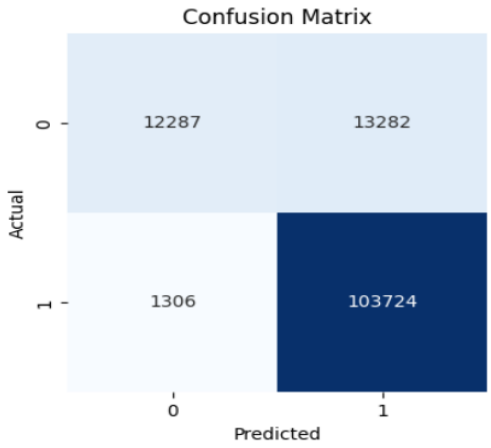
```
Classification Report:
              precision    recall  f1-score   support

           0       0.57      0.59      0.58     25569
           1       0.90      0.89      0.90    105030

    accuracy                           0.83    130599
   macro avg       0.73      0.74      0.74    130599
weighted avg       0.83      0.83      0.83    130599
```

Şekil 17. Karar Ağacı Modeli Sınıflandırma Raporu(*Decision Tree Model Classification Report)*

4.2.3 Catboost Algoritması(CatBoost Algorithm)

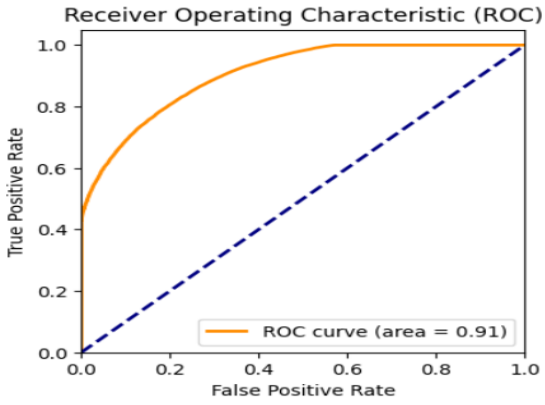
CatBoost (Categorical Boosting) algoritması kullanılarak bir sınıflandırma modeli oluşturulmuş ve performans değerlendirmesi gerçekleştirilmiştir. CatBoostClassifier, kategorik verilerle etkili çalışabilen bir gradient boosting algoritması olup, model eğitimi için varsayılan hiperparametreler kullanılmıştır.

Catboost modelinin matrisdeki değerleri incelendiğinde model, temerrüde düşen müşterileri büyük oranda doğru tahmin etmiş. Gerçekten temerrüde düşen müşterilerden sadece 1306'sını yanlış tahmin etmiş olarak görülmektedir.



Şekil 18.Catboost Matrisi*(CatBoost Matrix)*

Catboost modelinin performans sonucuna göre accuracy skoru 0.89 ROC Auc performans değeri 0.91 olup modelin güçlü bir tahmin yeteneğine sahip olduğunu gösterir.



Şekil 19. Catboost ROC Eğrisi *(CatBoost ROC Curve)*

Model genel olarak iyi bir eğitim ve test doğruluğuna sahip.Eğitim Doğruluğu 0.90 ,Test doğruluğu ise 0.89 'dür. Ancak, test ROC Auc değeri ,eğitim ROC Auc değerinden biraz daha düşüktür. Çıkan sonuçlar CatBoost'un test verisi üzerinde başarılı tahminler yaptığını ve geleneksel karar ağaçlarına kıyasla daha güçlü bir performans sunduğunu ortaya koymaktadır.

```
Test Classification Report:
              precision    recall  f1-score   support

           0       0.90      0.48      0.63     25569
           1       0.89      0.99      0.93    105030

    accuracy                           0.89    130599
   macro avg       0.90      0.73      0.78    130599
weighted avg       0.89      0.89      0.87    130599
```

Şekil 20.Catboost Modeli Sınıflandırma Raporu*(Catboost Model Classification Report)*

Sınıflandırma raporuna göre Sınıf 1'i Logistic Regression kadar iyi tahmin etmektedir. (F1-score = 0.93). Sınıf 0 için recall daha iyi (0.48), ancak precision biraz daha düşük olduğu görülmüştür.

4.2.4 Yapay Sinir Ağı (Artificial Neural Network)

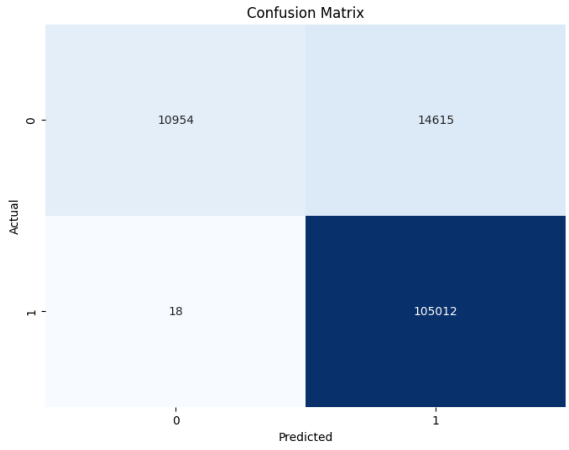
Çalışmada, Yapay Sinir Ağı (YSA) kullanılarak çok sınıflı bir sınıflandırma modeli oluşturulmuş ve değerlendirilmiştir. Model, Keras kütüphanesi ile Sequential API kullanılarak tasarlanmış ve Tam Bağlantılı Katmanlar (Fully Connected Layers) ile yapılandırılmıştır. Keras Yapay sinir ağı modeli (Sequential), tam bağlantılı katmanlar (Dense), aşırı öğrenmeyi önlemek için bırakma katmanı (Dropout) kullanılmıştır.

Modelin yapısı şu şekilde oluşturulmuştur:

- Giriş katmanı: 64 nöron, ReLU aktivasyon fonksiyonu.
- Gizli katmanlar: 64 nöronlu iki katman, ReLU aktivasyonu ve Dropout (0.5) kullanılarak aşırı öğrenme önlenmiştir.
- Çıkış katmanı: 3 sınıf için softmax aktivasyon fonksiyonu.

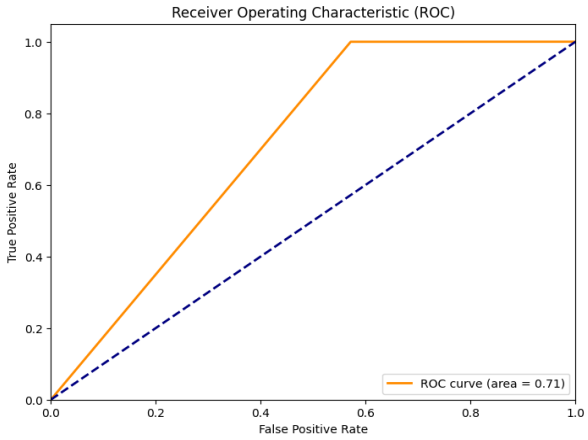
Model, Adam optimizasyon algoritması ve sparse_categorical_crossentropy kayıp fonksiyonu ile eğitilmiştir. 50 epoch ve 8 batch size kullanılarak eğitim gerçekleştirilmiş, doğrulama seti ile performans izlenmiştir.

Matrix 'e göre 10954 doğru negatif (True Negative), 14615 yanlış pozitif (False Positive), 18 yanlış negatif (False Negative) ve 105012 doğru pozitif (True Positive) tahmin yapıldığı görülmektedir. Model, borcunu ödeyemeyecek bireyleri (pozitif sınıf) çok iyi tespit ediyor (FN çok düşük, TP çok yüksek). Yanlış pozitiflerin (FP) fazla olması, bazı iyi müşterilere kredi vermeme riskini yaratabilir.



Şekil 21.Yapay Sinir Ağı Matrisi *(Artificial Neural Network Matrix)*

Yapay sinir ağı modelinin performans sonucuna göre Accuracy 0.89 ,ROC Auc Score değeri 0.71 'dir.



Şekil 22.Yapay Sinir Ağı ROC Eğrisi *(Artificial Neural Network ROC Curve)*

0 sınıfı için duyarlılık 0.43 ve 1 sınıfı için duyarlılık 1.00 olarak hesaplanmıştır. Bu, gerçek 0 sınıfına ait olan örneklerin yarısından azının doğru bir şekilde tahmin edildiğini ve gerçek 1 sınıfına ait olan örneklerin büyük bir kısmının doğru bir şekilde tahmin edildiğini gösterir. 1 sınıfı için yüksek bir F1-skoru elde edilmiştir.

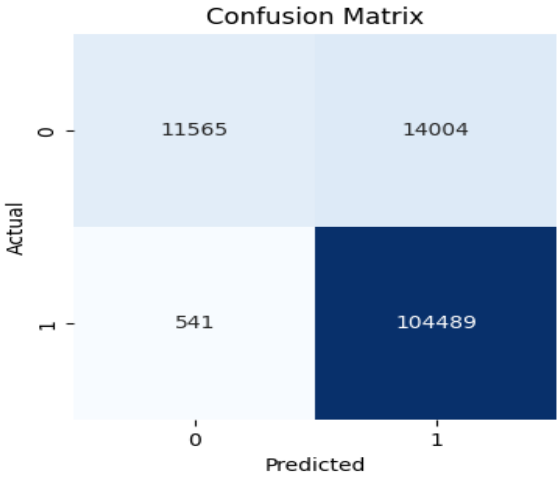
CLASSIFICATION REPORT:

	0	1	accuracy	macro avg	weighted avg
precision	1.00	0.88	0.89	0.94	0.90
recall	0.43	1.00	0.89	0.71	0.89
f1-score	0.60	0.93	0.89	0.77	0.87
support	25569.00	105030.00	0.89	130599.00	130599.00

Şekil 23.Yapay Sinir Ağları Modeli Sınıflandırma Raporu(*Artificial Neural Networks Model Classification Report)*

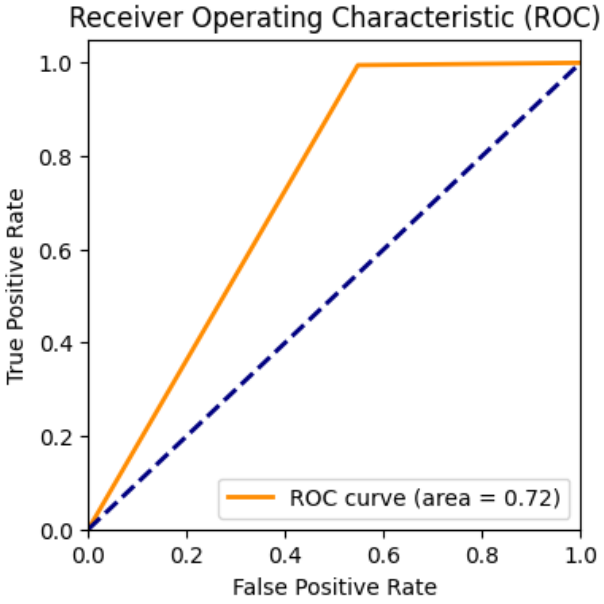
4.2.5 Random Forest Algoritması(Random Algoritması)

Rastgele Orman (Random Forest) algoritması kullanılarak bir sınıflandırma modeli oluşturulmuş ve performansı değerlendirilmiştir. RandomForestClassifier yöntemi, 100 karar ağacından (n_estimators=100) oluşan bir ansamble modeli olarak yapılandırılmıştır.Scikit-learn (sklearn): Random Forest modeli (RandomForestClassifier), modelin performans metrikleri (confusion_matrix, roc_curve, roc_auc_score, ConfusionMatrixDisplay) kullanılmıştır.



Şekil 24.Random Forest Matrisi*(Random Forest Matrix)*

Model temerrüde düşecek müşterileri (TP) büyük oranda doğru tahmin etmektedir. FN düşük olduğu için, riskli müşterileri kaçırma olasılığı düşük. Yanlış pozitif (FP) yüksek, yani bazı iyi müşterilere kredi verilmeme riski bulunmaktadır.



Şekil 25.Random Forest ROC Eğrisi *(Random Forest ROC Curve)*

Random Forest modelinin performans sonucuna göre Accuracy 0.89 ,ROC Auc Score değeri 0.72 'dir.

CLASSIFICATION REPORT:

	0	1	accuracy	macro avg	weighted avg
precision	0.96	0.88	0.89	0.92	0.90
recall	0.45	0.99	0.89	0.72	0.89
f1-score	0.61	0.93	0.89	0.77	0.87
support	25569.00	105030.00	0.89	130599.00	130599.00

Şekil 26.Random Forest Modeli Sınıflandırma Raporu(*Random Forest Model Classification Report)*

Sınıflandırma raporu incelendiğinde model eğitim verileri üzerinde çok yüksek bir doğruluk gösteriyor ancak test verileri üzerindeki doğruluk biraz daha düşük değerlerde görünüyor. Bu durum, modelin eğitim verilerine aşırı uyum sağlamış olabileceğini veya genelleme yeteneğinin biraz zayıf olabileceğini gösterebilir. Test sonuçlarına göre 0 sınıfı için kesinlik 0.95, 1 sınıfı için kesinlik 0.88 olarak hesaplanmıştır. Modelin 1 olarak tahmin ettiği örneklerin çoğunluğunun gerçekten 1 olduğunu ve 0 olarak tahmin ettiği örneklerin büyük bir kısmının da gerçekten 0 olduğunu gösterir. 0 sınıfı için duyarlılık 0.45 ve 1 sınıfı için duyarlılık 0.99 olarak hesaplanmıştır

5. SONUÇ VE ÖNERİLER (CONCLUSION AND RECOMMENDATIONS)

Veri analizi sonucunda seçilen öznitelik değişkenleri kullanılarak, kredilerin gecikme riski bir nevi kredi riski (0-1) bahsi geçen sınıflandırma algoritmaları ile tespit edilmeye çalışılmıştır. Kullanılan algoritmaların sınıflandırma performansları aşağıdaki tabloda belirtilerek karşılaştırılmıştır. Elde edilen nihai değişkenler ile eğitim, doğrulama ve test veri setleri üzerinden algoritmalar kredinin tahsis edilebilmesi için ,kredinin ödenebilme performansı yönünden müşterileri iyi ve kötü müşteri olarak sınıflandırarak, karmaşıklık matrisleri elde edilmiştir. Test veri seti üzerinden her bir algoritmaya ait kredi temerrüt (1) sınıfına ait karmaşıklık matrisinden elde edilen FN ,FP ve sınıflandırma raporundan elde edilen sınıflandırma ölçülerinin sonuçlarına aşağıdaki tabloda yer verilmiştir.

Model	Accuracy (%)	AUC-ROC	Precision	Recall	F1-Score	FN	FP
Lojistik Regresyon	89	0.91	0.88	0.99	0.94	534	13,954
Decision Tree	83	0.74	0.90	0.89	0.90	11,345	10,576
CatBoost	89	0.91	0.89	0.99	0.93	18	14,615
Yapay Sinir Ağları	89	0.71	0.88	1.00	0.93	18	14,615
Random Forest	89	0.72	0.88	0.99	0.93	541	14,004

Tablo 1 : Sınıflandırma Ölçütleri(*Classification Metrics)*

AUC-ROC (Receiver Operating Characteristic - Area Under Curve), modelin sınıfları ne kadar iyi ayırdığını gösterir. Lojistik Regresyon ve CatBoost en yüksek AUC-ROC değerine sahip modellerdir. Recall (1), gerçekten temerrüde düşen müşterilerin ne kadarını doğru tahmin ettiğini gösterir. Yapay Sinir Ağları (1.00) recall açısından en iyi modeldir. Ancak, diğer metrikleri de değerlendirmek gerekir. Precision (1), modelin "temerrüde düşecek" dediği kişilerin gerçekten temerrüde düşüp düşmediğini gösterir. Decision Tree bu metrikte en yüksek değere sahip olsa da genel başarısı düşüktür. FN, modelin temerrüde düşeceğini öngöremediği müşteri sayısını gösterir. Düşük olması istenir. En düşük FN değerine sahip modeller CatBoost ve Yapay Sinir Ağlarıdır.Genel olarak çalışma sonucu değerlendirildiğinde CatBoost modeli, kredi temerrüdü tahmini için en iyi seçenek olduğu söylenebilir.

☑ **Yüksek AUC-ROC (0.91):** Modelin temerrüde düşenleri ve düşmeyenleri ayırma gücü yüksektir.

☑ **Recall (1) = 0.99:** Temerrüde düşecek kişileri kaçırma oranı çok düşük.

☑ **FN = 18 (En düşük değerlerden biri):** Yanlış negatif

sayısı minimum seviyede.
☑ **Dengeli Precision ve F1-Score değerleri:** Overfitting yapmadan doğru tahmin yaptığı olarak değerlendirilmektedir.

Yapılan çalışma neticesinde hem yüksek AUC-ROC skoruna sahip hem de en düşük FN değerine sahip CatBoost modelinin kredi temerrüdünü tahmin etmek için en uygun model olduğu değerlendirilmiştir. Çalışma geleneksel lojistik regresyon yerine CatBoost gibi güçlü modeller de kullanılarak literatürdeki klasik yöntemlere kıyasla başarılı performanslar elde edilebileceğini göstermesi açısından önemlidir.

KAYNAKLAR (REFERENCES)

[1] Internet: C.Yeh & C.H Lien,The comparisons of data mining techniques for the predictive accuracy of probability of default of credit card clients , https://www.sciencedirect.com/science/article/abs/pii/S0957417407006719,11.11.2024.

[2] Internet: H. Budak & S. Erpolat, Kredi Riski Tahmininde Yapay Sinir Ağları ve Lojistik Regresyon Analizi Karşılaştırılması., https://dergipark.org.tr/tr/pub/ajit-e/issue/54448/741024 ,01.10.2024.

[3] Internet: Ş. Kavcıoğlu, Ticari Bankacılıkta Kredi Riskinin ve Kredi Riski Ölçüm Modellerinin Değerlendirilmesi, https://dergipark.org.tr/tr/download/article-file/3978, 02.10.2024.

[4] Internet: S. Hamori & M. Kawai & K. Takahiro & M.Yuji & C. Watanabe,Ensemble Learning or Deep Learning Application to Default Risk Analysis,https://epe.lacbac.gc.ca/100/201/300/jrn_risk_financial_mngmt/2018/v11no1/mdpi.com/1911-8074/11/1/12/htm.html ,11.10.2024.

[5] Internet: T. İldaş, Kredi Riski Ölçüm Modellerinin Değerlendirilmesi,https://dergipark.org.tr/en/download/article-file/1901150, 15.10.2024.

[6] M. Mahbobi & S. Kimiagari & M. Vasudevan, "Credit risk classification: an integrated predictive accuracy algorithm using artificial and deep neural networks", Annals of Operations Research, 330(1), 609–637, 2023.

[7] Internet: T.E Tütüncü & S. Gürsakal,Kredi Temerrüt Riskini Tahmin Etmede Makine Öğrenme Algoritmalarının Karşılaştırılması, https://dergipark.org.tr/tr/download/article-file/2635086 , 01.11.2024.

[8] Internet: M.E.F Milli , Sınıf Dengeleme Yöntemlerinin Makine Öğrenmesi Teknikleri Üzerine Etkisi: Kredi Risk Örneği, https://avesis.deu.edu.tr/yonetilen-tez/85127e15-e6b5-4804-a3b5-d6eae4aa02f9/sinif-dengeleme-yontemlerinin-makine-ogrenmesi-teknikleri-uzerine-etkisi-kredi-risk-ornegi, , 05.10.2024.

[9] Internet: X.Zhu & O. Chu & X.Song & P.Hu & L.Peng ,Explainable prediction of loan default based on machine learning models,https://www.sciencedirect.com/science/article/pii/S2666764923000218 , 11.10.2024.

X Platformunda İstenmeyen Gönderilerin Makine Öğrenmesi, Geniş Dil Modelleri ve Bilgisayarlı Görü ile Tespiti

Araştırma Makalesi/Research Article

Ebutalha CAMADAN[1], Mehmet ŞİMŞEK[2]

[1]Millî Savunma Üniversitesi, İstanbul, Türkiye
[2]Millî Savunma Üniversitesi, Kara Harp Okulu, Bilgisayar Mühendisliği Bölümü, Ankara, Türkiye

ebutalha.camadan@msu.edu.tr, mesimsek@kho.msu.edu.tr
(Geliş/Received:05.12.2024; Kabul/Accepted:11.03.2025)
DOI: 10.17671/gazibtd.1596990

Özet— Sosyal medya platformları, günümüzde bilgi paylaşımı ve iletişimde önemli araçlar haline gelirken, aynı zamanda istenmeyen gönderilerin (spam) yayılması da büyük bir sorun teşkil etmektedir. Bu çalışma, X sosyal medya platformundaki (eski adıyla Twitter) istenmeyen gönderilerin tespitine yönelik, makine öğrenmesi, geniş dil modelleri ve bilgisayarlı görü tekniklerini birleştiren yeni bir yaklaşım önermektedir. Türkiye'de popüler olan konulara dair görsel içeren gönderilerden bir veri kümesi oluşturularak, spam tespitinde en etkili makine öğrenmesi algoritmaları belirlenmeye çalışılmıştır. Gönderi içeriğinin etiketlerle ilişkisi ve birden fazla etiketin birbiriyle ilgisi gibi sosyal medya etkileşimini belirleyen öznitelikler geliştirilmiştir. Ayrıca, görsel içeriğin analizi için, görselin X platformunda ilk paylaşıldığı tarih ile internet üzerindeki diğer sayfalarda geçtiği metinle benzerliği gibi görsel odaklı öznitelikler de dahil edilmiştir. Bu öznitelikler, Google Gemini ve Cloud Vision AI araçları kullanılarak geliştirilmiştir. Beş farklı makine öğrenmesi algoritması (Karar Ağaçları, Rastgele Orman, SVM, Lojistik Regresyon, Çok Katmanlı Algılayıcı) ile yapılan deneylerde, Rastgele Orman algoritması en yüksek doğruluk ve F1 skoru değerlerine ulaşmıştır. Bu çalışma, X platformunda istenmeyen gönderi tespiti için makine öğrenmesi yöntemlerinin etkinliğini göstermiş ve Google Gemini ile Cloud Vision AI araçlarının etkin kullanımına dair yeni yöntemler sunmuştur. Ayrıca geliştirilen öznitelikler, spam içeriklerin doğru bir şekilde sınıflandırılmasında güçlü bir temel oluşturmaktadır.

Anahtar Kelimeler— X sosyal medya platformu, istenmeyen gönderi, makine öğrenmesi, geniş dil modelleri, bilgisayarlı görme

Detecting Spams on the X Platform with Machine Learning, Large Language Models, and Computer Vision

Abstract— While social media platforms have become crucial tools for information sharing and communication, the spread of unwanted content (spam) has also become a significant problem. This paper proposes a novel approach for spam detection on the social media platform X (formerly Twitter) by integrating machine learning, large language models, and computer vision techniques. A dataset containing posts with visual content on popular Turkish topics was created, aiming to identify the most effective machine learning algorithms for spam detection. Feature engineering was conducted to capture key aspects of social media interaction, including the relationship between post content and hashtags, as well as the relevance between multiple hashtags. Additionally, image-based features were introduced, such as the initial posting date of an image on X and its textual similarity to other web pages, to enhance visual content analysis. These features were developed using Google Gemini and Cloud Vision AI. Experimental evaluations with five machine learning algorithms (Decision Trees, Random Forest, SVM, Logistic Regression, and Multilayer Perceptron) demonstrated that the Random Forest algorithm achieved the highest accuracy and F1 score. This paper highlights the effectiveness of machine learning methods in spam detection on X and introduces new methodologies for leveraging Google Gemini and Cloud Vision AI. Furthermore, the engineered features provide a strong foundation for accurately classifying spam content.

Keywords— X social media platform, spam message, machine learning, large language models, computer vision

1. GİRİŞ (INTRODUCTION)

Bu bölümde, Online Sosyal Ağlar (OSA'lar) özellikle X platformu üzerinden iletişim ve etkileşimdeki rolünü inceleyerek, istenmeyen gönderilerin (spam) tespitine yönelik güncel yaklaşımları ve bu alandaki zorlukları ele alacağız. Ayrıca, bu çalışmanın temel hedefi, metin ve görsel içeriklerin birleşik analizine dayalı yeni bir spam tespit yaklaşımını tanıtmaktır.

Online Sosyal Ağlar (OSA'lar), özellikle geniş kullanıcı kitlesine sahip X platformu aracılığıyla, günümüzde iletişim ve etkileşimde vazgeçilmez bir araç haline gelmiştir [1-2]. Milyonlarca kullanıcı, X platformunu anılarını paylaşmak, gözlemlerini yayınlamak ve "Gönderi (Tweet)" adı verilen kısa mesajlarla etkileşimde bulunmak için kullanarak platformun toplumsal etkisini artırmaktadır [3].

Bu çalışmanın temel motivasyonu, sağlıklı bir online sosyal ağ ortamı oluşturmak için istenmeyen gönderilerin (spam) tespitinin kritik önem taşımasıdır. İstenmeyen gönderiler, X gibi platformlarda kullanıcı deneyimini olumsuz etkileyebilir. Ayrıca, dolandırıcılık ve kimlik avı gibi güvenlik risklerine yol açarak platformun güvenilirliğini zedeler [4-5]. İstenmeyen gönderiler, genellikle ticari veya kötü niyetli amaçlarla toplu halde paylaşılan, alıcı tarafından istenmeyen ve zararlı içerikler içerebilen mesajlardır. X platformunda bu içerikler, kullanıcıları tehlikeli web sitelerine yönlendiren URL'ler içerebilen kısa mesajlar ("Gönderiler/Tweetler") şeklinde yayılmaktadır [6]. E-postayla kıyaslandığında, X üzerindeki istenmeyen gönderilerin daha yüksek tıklanma oranına sahip olması [7], bu sorunun platform için ciddi bir tehdit olduğunu göstermektedir.

X platformunda istenmeyen gönderi tespitine yönelik güncel yaklaşımlar, Web of Science ve ScienceDirect veritabanları üzerinden 2017 sonrası yayınlar taranarak incelenmiştir. Arama terimi olarak "X/Twitter Spam Detection (X İstenmeyen Gönderi Tespiti)" kullanılmış ve derin öğrenme (DL), makine öğrenmesi, bulanık mantık gibi çağdaş yöntemlere odaklanan çalışmalar önceliklendirilmiştir. Literatürdeki çalışmalar, bu algoritmaların genellikle sözdizimi analizi, özellik çıkarımı ve kara listeleme yöntemlerine dayandığını göstermektedir.

X platformunda istenmeyen gönderi tespiti, mevcut makine öğrenmesi ve kara listeleme yöntemlerinin yetersiz kalması nedeniyle zorlu bir problem teşkil etmektedir. Bu sorunu çözmek için bazı araştırmacılar derin öğrenme tekniklerinden yararlanarak yeni yaklaşımlar önermiştir. Derin öğrenme tabanlı yaklaşımlar (YSA, CNN), geleneksel yöntemlere göre daha başarılı sonuçlar göstermiştir. Ancak, bu modellerin karmaşıklığı ve yüksek veri ihtiyacı, pratik uygulamalarda zorluklar yaratmaktadır [8]. Diğer araştırmacılar [9], kullanıcıların istenmeyen gönderi tespit sistemlerinden kaçınma stratejilerine yönelik hibrit bir teknik geliştirmiştir. Başka bir grup ise gönderi tabanlı analizlerin yetersiz olduğunu ve kullanıcı hesapları ile sosyal grafiğin önemini vurgulamıştır [10].

Bir grup araştırmacı [11], kullanıcı hesaplarına dayalı bir yaklaşım önererek gönderi sıklığının bir gösterge olabileceğini belirtmiştir. Ancak, bu yöntem LSA gibi tekniklere dayansa da, içeriğin anlamını tam olarak yakalayamamaktadır. Madisetty ve Desarkar'ın ensemble yöntemi ve kelime gömme tekniklerini kullandıkları çalışma umut verici olsa da [12], diğer araştırmacıların "Honey Profiles" metodundaki gibi veri toplama zorlukları devam etmektedir [13]. URL ve kara listeleme gibi basit yöntemlerin sınırlılıkları [14] ve sınıflandırma ile kümeleme yöntemlerinin değişken başarı oranları [15-16], tek bir çözümün yetersiz olduğunu göstermektedir.

Araştırmacıların doğal dil işleme ve makine öğrenmesi sınıflandırmaları üzerine çalışmaları [17] ve başka bir grup araştırmacının Naive Bayes önerisi [18], metin analizi ve olasılık hesaplamalarının önemini vurgulamaktadır. Ancak, Naive Bayes'in doğruluk oranı veri kalitesine bağlıdır. Öte yandan, dengesiz veri setlerinin yarattığı zorluklara ve SMOTE gibi tekniklerin önemine dikkat çeken araştırmalar da bulunmaktadır [19-20]. Sumathi ve Raja'nın [21] ETC ve VC kullandıkları yöntemi ile Thomas ve Meshram'ın [22] DNFNet yaklaşımı, karmaşık modellerin potansiyelini göstermektedir. Bununla birlikte, diğer bir grup araştırmacı [23], ensemble tekniklerinin, özellikle RF'nin, bireysel modellere göre daha iyi sonuçlar verdiğini belirtmiştir. Son yıllarda ise derin öğrenme tabanlı yaklaşımlar uygulanmış ve daha önceki yıllarda belirtilen veri ihtiyacı ve pratik uygulama zorluklarına değinilmiştir [24-25]. Diğer bir çalışmada ise araştırmacılar kuantum hesaplama prensiplerini kullanarak ikili kütleçekim arama algoritmasını rastgele orman modeliyle birleştirerek daha yüksek doğruluk oranı yakalamıştır. [26].

Sonuç olarak, mevcut çalışmalar, X platformunda istenmeyen gönderi tespitinin dinamik yapısı nedeniyle tek bir ideal çözümün bulunmadığını ve sürekli gelişen yeni yaklaşımlara ihtiyaç duyulduğunu ortaya koymaktadır. Mevcut çalışmaların çoğunluğu metin tabanlı gönderilere odaklansa da görsel içeriklerin, özellikle manipüle edilmiş, yanıltıcı veya uyumsuz görsellerin yaygınlaşmasıyla birlikte, istenmeyen gönderi tespitinde giderek daha önemli bir rol oynadığı unutulmamalıdır. Görsel içerikler, metin tabanlı gönderilere kıyasla kullanıcıların dikkatini daha kolay çekmektedir [27]. Bu nedenle, görsel içeren gönderilerin tespiti, istenmeyen içeriklerin engellenmesi açısından kritik bir öneme sahiptir.

Bu çalışmada, X sosyal medya platformundaki istenmeyen gönderilerin tespitine yönelik makine öğrenmesi, geniş dil modelleri ve bilgisayarlı görü tekniklerini birleştirerek yeni bir yaklaşım geliştirmeyi amaçladık. Bu doğrultuda, X platformundan veri topladık ve spam tespiti için zenginleştirilmiş bir veri seti oluşturduk. Literatürdeki mevcut çalışmalardan farklı olarak, bu veri seti hem metin hem de görsel içeriklerin analizini bir arada sunan özgün bir yapıya sahiptir. Mevcut çalışmalar genellikle yalnızca

metin tabanlı veya sınırlı görsel analiz tekniklerine odaklanırken, bizim yaklaşımımız, geniş dil modelleri ve bilgisayarlı görü tekniklerini entegre ederek istenmeyen içerik tespitine yeni bir bakış açısı getirmektedir. Ayrıca, sadece seçtiğimiz makine öğrenmesi algoritmalarının performansını kıyaslayarak, literatürdeki diğer çalışmaların sonuçlarıyla doğrudan bir karşılaştırma yapmamaktayız.

Çalışmamızın literatüre katkıları şunlardır:

- Metin ve görsel içeriklerin birleşik bir şekilde analiz edildiği özgün bir veri seti oluşturulması.
- Google Gemini ve Cloud Vision AI kullanarak geniş dil modelleri ile görsel analiz tekniklerinin entegre edilmesi.
- Türkiye'de trend olan konulara dayalı spam içeriklerinin tespiti için özgün bir veri seti ve analiz yaklaşımı geliştirilmesi.

Veri setimizi zenginleştirmek için, sosyal medya gönderilerinin içerik analizini gerçekleştirmek amacıyla Google'ın üretken dil modeli Gemini ve görsel analiz hizmeti Cloud Vision AI kullanılmıştır. Bu teknolojiler, geleneksel yöntemlerden farklı olarak metin ve görsel içerikleri bütünsel bir şekilde analiz etmeyi mümkün kılmaktadır. Bu sayede, gönderi içeriklerinin etiketlerle olan ilişkisini daha iyi anlayıp, görsel öğeleri de analiz edebildik. Türkiye'de trend olan konulara ait görsel içeren gönderileri inceleyerek, istenmeyen içeriklerin tespitine yönelik sağlam bir temel oluşturduk.

Veri setimizin sınıflandırılmasında beş farklı makine öğrenmesi algoritması kullanarak deneyler gerçekleştirdik. Denediğimiz yöntemler arasında Karar Ağaçları, Rastgele Orman, SVM, Lojistik Regresyon ve Çok Katmanlı Algılayıcı yer almaktadır. Deneyler sonucunda, Gemini'nin etiket analizi konusundaki güçlü performansı ve Cloud Vision AI'nın görsel analiz yetenekleri, Rastgele Orman algoritmasının veri seti üzerinde en yüksek doğruluğa ulaşmasını sağlamıştır.

Çalışmamızın kısıtları şunlardır:

1. Kullanılan veri seti yalnızca belirli bir sosyal medya platformu ile sınırlıdır ve dolayısıyla genelleştirilebilirliği başka platformlar için test edilmemiştir.
2. Çalışma sadece belirli makine öğrenmesi algoritmalarında test edilmiştir; derin öğrenme modelleri kapsam dışında bırakılmıştır.
3. Görsel analiz araçlarının sınırlamaları nedeniyle bazı görsellerin yanlış sınıflandırılması mümkündür.

Bu çalışma, X platformunda istenmeyen gönderi tespiti için makine öğrenmesi yöntemlerinin uygulanabilirliğini kanıtlamanın yanı sıra, bu alanda yeni ve etkili yöntemlerin geliştirilmesine de katkı sağlamaktadır.

Bu bölümde, X platformunda istenmeyen gönderi tespiti sorununu ve bu alandaki mevcut çözümleri inceledik. Ayrıca, bu çalışmanın metin ve görsel içerikleri birleştiren özgün yaklaşımını tanıttık. Bir sonraki bölümde, kullanılan veri seti ve yöntemlerin detaylarına odaklanacağız.

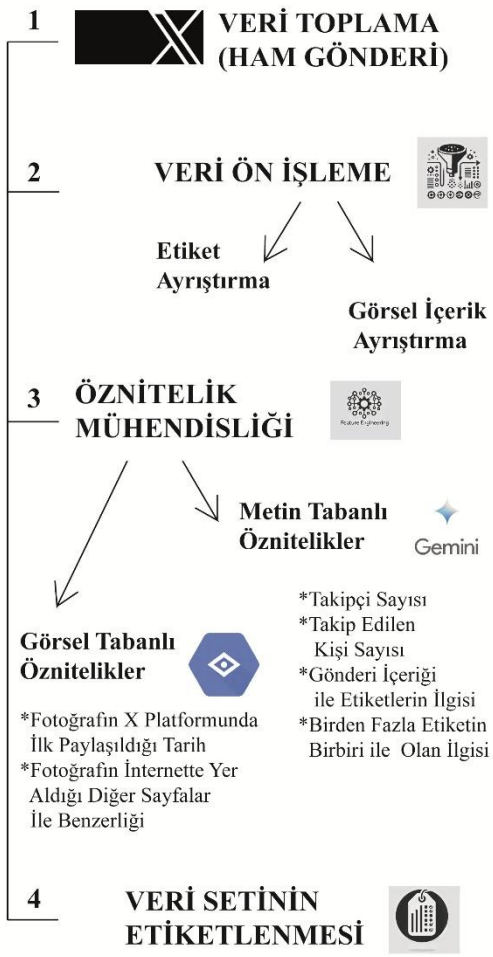
2. MATERYAL VE YÖNTEM (MATERIALS AND METHODS)

Bu bölümde, çalışmada kullanılan veri setinin oluşturulma süreci, özniteliklerin belirlenmesi ve bu özniteliklerin spam tespiti üzerindeki etkisi ele alınmaktadır. Fotoğraf içeren gönderilere odaklanılmasının nedenleri açıklanarak, bu tercihlerin analiz sürecine katkısı vurgulanmaktadır. Ayrıca, spam tespiti için kullanılan Karar Ağacı, Rastgele Orman, Destek Vektör Makineleri (SVM), Lojistik Regresyon ve Çok Katmanlı Algılayıcı (MLP) gibi makine öğrenmesi algoritmaları tanıtılmakta ve bu yöntemlerin veri setine nasıl uygulandığı detaylandırılmaktadır.

2.1. Üretilen Veri Seti

Twitter veri setlerine dayanan önceki çalışmalar, sosyal medya analizlerinde temel eksiklikler barındırmakta ve kapsamlı analizler için yetersiz kalmaktadır. Bu çalışmada kullanılan veri seti, 6 Şubat 2023 ile 17 Temmuz 2023 tarihleri arasında Türkiye'de trend olan 12 konuya ait 2481 gönderiden oluşmakta olup, tüm gönderiler fotoğraf içermektedir. Fotoğraf içeren gönderilere odaklanılmasının nedeni, istenmeyen içeriklerin çoğunlukla görsel medya üzerinden yayıldığının önceki çalışmalarda gösterilmesidir [28]. Ancak, önceki çalışmalar genellikle metin tabanlı analizlerle sınırlı kalmış ve görsel içeriklerin etkisi göz ardı edilmiştir. Bununla birlikte, görsel içerikler sosyal medya kullanıcı davranışları ve trend analizleri açısından önemli bilgiler sunmaktadır. Bu bağlamda, çalışmamızda fotoğraf içeren gönderilere yoğunlaşarak, spam içerikleri tespit etmek için yeni öznitelikler geliştirdik.

Öncelikle, takipçi sayısı, takip edilen kişi sayısı, gönderi içeriğinin etiketlerle ilişkisi, birden fazla etiketin birbirleriyle olan ilgisi gibi sosyal medya etkileşimini belirleyen önemli öznitelikler belirledik. Ayrıca, görsel içeriği analiz etmek için görselin X platformunda ilk paylaşıldığı tarih ve görselin internet üzerindeki diğer sayfalarda geçtiği metinle benzerliği gibi görsel odaklı öznitelikler de dahil ettik. Bu süreçte, X platformundan elde edilen ham verilerin (gönderilerin) özniteliklere dönüştürülmesi adımları, Şekil 1'de özetlenmektedir.



Şekil 1. Ham Veriden Öznitelik Çıkarımı Süreci
(Process of Feature Extraction from Raw Data)

Türkiye'ye özgü veri setleri üzerine yapılan çalışmalar ise sınırlıdır; Türkiye'deki yerel trendlerin, özellikle spam içerik gibi olgular üzerindeki etkisini analiz etmek için bu çalışmadaki gibi geniş kapsamlı bir veri setine ihtiyaç duyulmaktadır. Ayrıca, Google Gemini ve Cloud Vision AI gibi yeni teknolojiler, sosyal medya analizlerini daha derinlemesine inceleme fırsatı sunmasına rağmen, mevcut çalışmalarda bu araçlardan yeterince yararlanılmamıştır. Bu tür araçların entegrasyonu, daha zengin analizler yapılmasına olanak sağlayabilir. Son olarak, önceki çalışmalarda veri çeşitliliği sınırlı kalmış ve bu durum analizlerin kapsamını daraltmıştır. Çeşitli veri kaynaklarından elde edilen zenginleştirilmiş veri setleri, sosyal medya analizlerinde daha güvenilir ve ayrıntılı sonuçlar sunacaktır. Bu eksiklikler, mevcut X platformunda veri setlerinin neden yeterli olmadığını açıkça ortaya koymaktadır.

2.1.1.Veri Seti Üretiminde Kullanılan Yapay Zeka Araçları ve Alternatif Yöntemler

Bu çalışmada, metin ve görsel analizler için Google Gemini ve Google Cloud Vision AI araçları tercih edilmiştir. Bu araçlar, geniş dil ve görüntü işleme yetenekleri, API erişimi ve düşük maliyetli entegrasyon avantajları nedeniyle seçilmiştir. Ancak, Google Gemini bazı dillerde ve yerel konularda düşük doğruluk oranları sergileyebilmekte, ayrıca politikaları gereği bazı içerikleri filtreleyerek analiz dışı bırakabilmektedir. Google Cloud Vision AI ise görsel tanımada güçlü olmakla birlikte, sosyal medya spam tespiti için özel olarak tasarlanmadığından sınırlı bir performans göstermektedir.

Bu sınırlamaları aşmak için alternatif araçlar değerlendirilmiştir. OpenAI GPT gibi dil modelleri ve Microsoft Azure Computer Vision ile Amazon Rekognition gibi görsel analiz araçları potansiyel alternatifler olarak düşünülmüş, ancak çalışma kapsamına dahil edilmemiştir. Nihai tercih, Google araçlarının pratik avantajları ve yüksek doğruluk oranları nedeniyle bu yönde kullanılmıştır.

Sonuç olarak, çalışmada Google Gemini ve Cloud Vision AI'nın avantajlarından yararlanılmış, diğer araçlar literatür ışığında değerlendirilmiştir. Bu yaklaşım, spam içerik tespiti için kapsamlı bir analiz çerçevesi sunmuştur.

2.1.2.Takipçi Sayısı Özniteliği

Takipçi sayısı, X sosyal medya platformundaki hesapların etkileşim düzeyini ve popülerliğini anlamak için kullanılan önemli bir göstergedir. İstenmeyen gönderi paylaşan hesaplar, genellikle sahte veya bot hesaplar olduğundan, bu hesapların takipçi sayıları gerçek hesaplardan farklılık gösterebilir. Bu nedenle, takipçi sayısı, sahte hesapları tespit etmek için önemli bir öznitelik olarak kullanılmıştır. Bu veri, X API'si kullanılarak hesapların takipçi sayısına göre toplanmıştır.

2.1.3.Takip Edilen Kişi Sayısı Özniteliği

Takip edilen kişi sayısı, bir X sosyal medya hesabının etkileşim yapısını anlamak için önemli bir göstergedir. İstenmeyen gönderi paylaşan hesaplar genellikle rastgele veya çok sayıda kişiyi takip etme eğilimindedir. Bu durum, hesapları daha yaygın kullanıcı hesaplarından ayırabilir. Bu nedenle, takip edilen kişi sayısı, istenmeyen içerik paylaşan hesapları tespit etmede önemli bir gösterge sunar. Bu öznitelik, X API'si aracılığıyla elde edilmiştir.

2.1.4.Gönderi İçeriğinin Etiketler ile İlgisi Özniteliği

Gönderi içeriği ile etiketlerin ilişkisi, her bir gönderinin içeriği ile kullanılan etiketlerin uyumunu analiz ederek, hesapların sahte ya da bot olup olmadığına dair ipuçları sunar. İstenmeyen içerik yayan hesaplar, genellikle gönderilerini daha geniş bir kitleye ulaştırmak amacıyla popüler veya trend etiketleri kullanma eğilimindedir. Ancak, bu etiketlerin gönderi içeriğiyle uyumsuz veya alakasız olması, hesapları tespit etmek için belirleyici bir faktör olabilir.

Gönderi içeriği ile etiketlerin ilişkisinin analiz edilebilmesi için, gönderi metinleri ve etiketler ayrıştırılmış ve veriler

analiz edilebilir bir formata getirilmiştir. Bu adımda, her bir gönderi metni ve içerikteki etiketler ayrı sütunlara yerleştirilmiş, etiketlerin yoğunluğu analiz edilerek veri setinin yapısı hakkında bilgi edinilmiştir.

Gemini API, gönderinin içeriği ile ona atanan etiketlerin uyum derecesini belirlemek için kullanılmıştır. Bu API, gönderi içeriği ile ona atanan etiketlerin uyumunu 'Çok İlgisiz', 'İlgisiz', 'Nötr', 'İlgili' ve 'Çok İlgili' gibi derecelendirilmiş bir ölçekle sunar. Bu analiz, etiket ve içerik uyumunu belirleyerek, istenmeyen içerik paylaşan hesapları tespit etmeye yardımcı olur. Örneğin, tamamen alakasız etiketlere sahip bir gönderi "Çok İlgisiz" olarak sınıflandırılırken, konuya uygun etiketler taşıyan bir gönderi "Çok İlgili" olarak değerlendirilir.

Gemini API ayrıca, politikalarına aykırı içerikleri otomatik olarak filtreler ve şiddet, nefret söylemi veya cinsel içerik gibi unsurları engeller. Bu tür içeriklerin filtrelenmesi, veri setinin güvenilirliğini artırırken, engellenen içeriklerin yerine alternatif kategori etiketleri atanarak veri seti zenginleştirilmiştir.

2.1.5.Birden Fazla Etiketin Birbirleri ile Olan İlgisi Özniteliği

Etiketlerin birbirleriyle olan ilgisi özniteliği, X sosyal medya platformundaki gönderilerde yer alan etiketlerin içerik uyumunu analiz etmek amacıyla oluşturulmuştur. İstenmeyen gönderi paylaşan hesaplar genellikle çok sayıda alakasız veya popüler etiketi birlikte kullanarak dikkat çekmeye çalıştığından, etiketlerin birbiriyle olan ilgisi, bu tür hesapları tespit etmekte önemli bir göstergedir.

Bu öznitelik, Gemini API kullanılarak oluşturulmuştur. İlk olarak, her gönderideki etiket sayısı belirlenmiş, ardından etiketler ayrıştırılarak uyum analizine hazırlanmıştır. Gemini API, her gönderinin etiket uyumunu değerlendirirken 'Çok İlgisiz', 'İlgisiz', 'Nötr', 'İlgili' ve 'Çok İlgili' olmak üzere beş dereceden oluşan bir ölçek kullanmıştır. Bu ölçek, etiketlerin içerik ile ne kadar uyumlu olduğunu belirleyerek, gönderinin istenmeyen içerik olma olasılığını anlamaya yardımcı olmuştur.

2.1.6.Gönderi İçeriğinde Yer Alan Fotoğrafın X Platformu Üzerinde Paylaşılma Durumu Özniteliği

Gönderi içeriğinde yer alan fotoğrafın X üzerindeki ilk paylaşıldığı tarih, istenmeyen gönderi paylaşan hesapları tespit etmek için kullanılan önemli bir özniteliktir. Bu öznitelik, bir fotoğrafın başka bir kullanıcı tarafından daha önce paylaşılmış olup olmadığını ve yeniden paylaşılma oranını belirler. İstenmeyen hesaplar genellikle özgün içerik üretmek yerine mevcut içerikleri yeniden paylaşır, bu da fotoğrafın eski tarihlerde paylaşıldığını gösterir.

Veri setindeki fotoğraflar, Google Cloud Vision API aracılığıyla taranarak, her bir fotoğrafın X üzerindeki ilk paylaşıldığı tarih belirlenmiştir. İlk olarak, fotoğraf içeren gönderilerden URL'ler toplanmış ve ardından fotoğrafların önceki paylaşımlarına dair bilgiler elde edilmiştir. Eğer bir fotoğraf daha önce başka bir hesap tarafından paylaşılmışsa, bu durum fotoğrafın istenmeyen içerik olma olasılığını artırır. Son olarak, fotoğrafın paylaşım tarihi eskiyse 'Kaynak Kendisi-1', yeni ise 'Kaynak Başkası-2' etiketi verilmiştir.

Çalışmada kullanılan veri seti, 'Gönderi İçeriğinde Yer Alan Fotoğrafın X Platformu Üzerinde Paylaşılma Durumu' özniteliğinin Google Cloud Vision API ile oluşturulma durumuna göre ikiye ayrılmıştır: 293 gönderiden oluşan Veri Seti-I ve 2188 gönderiden oluşan Veri Seti-II.

Bir gönderideki görsel, X platformunda daha önce başka bir hesap tarafından paylaşılmamışsa, bu görselin kaynağı doğrudan "kendisi" olarak kabul edilir. Bu durum, Veri Seti-II'deki tüm görseller için geçerli olduğundan, 'Kaynak Kendisi-1' özniteliğinin eklenmesine gerek kalmamış ve bu öznitelik Veri Seti-II'den çıkarılmıştır.

2.1.7.Gönderi İçeriğinde Yer Alan Fotoğrafın İnternette Yer Alan Diğer Sayfalarda Birlikte Geçtiği Metin Arasındaki Benzerlik Özniteliği

Gönderi içeriğinde yer alan fotoğrafın internet üzerindeki diğer sayfalarda geçtiği metinle benzerliği özniteliği, X sosyal medya platformundaki içeriklerin anlamını daha iyi anlamak amacıyla kullanılmıştır. Bu öznitelik, bir fotoğrafın internet üzerindeki başka sayfalarda hangi metinlerle birlikte yer aldığını inceleyerek, gönderinin içeriği hakkında önemli bilgiler sağlar. Örneğin, bir fotoğrafın haber sitesinde aynı başlıkla yer alması, gönderinin haberle ilgili olduğunu düşündürebilir.

Bu öznitelik, Google Cloud Vision API'nin "Pages with matching images" özelliği ile oluşturulmuştur. API, görsellerin internet üzerindeki diğer sayfalarda hangi başlıklarla eşleştiğini belirler ve bu başlıkların metin içerikleriyle gönderi metnini Gemini API aracılığıyla karşılaştırır. Bu süreç, gönderi içeriği ve görsellerinin anlamını daha derinlemesine incelemeye olanak tanır. İlk aşamada, Vision API ile görsellerin eşleştiği sayfalardan "pageTitle" bilgisi toplanmış ve bu başlıkların metin içerikleri, Gemini API kullanılarak gönderi metniyle karşılaştırılmıştır.

Ardından, gönderi içeriği ile etiketlerin uyumunu belirlemek için kullanılan ölçek, 'Hiç Benzer Olmayan', 'Benzer Olmayan', 'Nötr', 'Benzer' ve 'Çok Benzer' olmak üzere beş dereceden oluşmaktadır. Bu dereceler, içerik ile etiketlerin uyumunu sırasıyla tanımlar: 'Hiç Benzer Olmayan' en düşük uyum seviyesini, 'Çok Benzer' ise en yüksek uyum seviyesini ifade eder.

2.1.8. Öznitelik Önem Analizi

Bu çalışmada kullanılan özniteliklerin spam tespitindeki önemini değerlendirmek amacıyla, Mutual Information

(MI) yöntemi kullanılarak özniteliklerin hedef değişken (spam veya spam değil) ile olan ilişkisi analiz edilmiştir. Mutual Information skoru, bir özniteliğin bağımsız değişken olarak hedef değişkeni (spam olup olmama durumu) ne kadar açıkladığını göstermektedir. Veri Seti-I için elde edilen Mutual Information skorları Tablo 1'de sunulmaktadır.

Tablo 1. Veri Seti-I için Mutual Information Skorları
(Mutual Information Scores for Dataset-II)

Öznitelik Adı	MI Skoru
Takipçi Sayısı	0.0234
Takip Edilen Kişi Sayısı	0.0146
Gönderi İçeriğinin Etiketler ile İlgisi	0.0591
Birden Fazla Etiketin Birbirleri ile Olan İlgisi	0.0569
Gönderi İçeriğinde Yer Alan Fotoğrafın X Platformu Üzerinde Paylaşılma Durumu	0.0634
Gönderi İçeriğinde Yer Alan Fotoğrafın İnternette Yer Alan Diğer Sayfalarda Birlikte Geçtiği Metin Arasındaki Benzerlik	0.0000

Veri Seti-I için yapılan MI analizi, modelin hangi değişkenlerden daha fazla bilgi edindiğini belirlemek açısından önemli bir katkı sunmuştur. Sonuçlara göre, "Gönderi İçeriğinde Yer Alan Fotoğrafın X Platformu Üzerinde Paylaşılma Durumu" (0.0634) en yüksek MI skoruna sahip olup, modelin tahmin gücüne en fazla katkıyı sağlamaktadır. Bunu "Gönderi İçeriğinin Etiketler ile İlgisi" (0.0591) ve "Birden Fazla Etiketin Birbirleri ile Olan İlgisi" (0.0569) değişkenleri takip etmektedir. Bu bulgular, içerik ve etiketler arasındaki ilişkinin modelin karar alma sürecinde kritik bir rol oynadığını göstermektedir. "Takipçi Sayısı" (0.0234) ve "Takip Edilen Kişi Sayısı" (0.0146) ise daha düşük bilgi katkısına sahip öznitelikler olarak belirlenmiştir.

Öte yandan, "Gönderi İçeriğinde Yer Alan Fotoğrafın İnternette Yer Alan Diğer Sayfalarda Birlikte Geçtiği Metin Arasındaki Benzerlik" (0.0000) değişkeninin model açısından anlamlı bir bilgi içermediği tespit edilmiştir. Bu nedenle, Veri Seti-I'den çıkarılarak modelin daha verimli hale getirilmesi sağlanmıştır. Bu tür düşük katkılı değişkenlerin elimine edilmesi, modelin sadece anlamlı değişkenlere odaklanmasını ve gereksiz hesaplama yükünü azaltmasını sağlamaktadır.

Veri Seti-II'ye yönelik yapılmış olan Mutual Information skorları ise Tablo 2'de gösterilmektedir.

Tablo 2. Veri Seti-II için Mutual Information Skorları
(Mutual Information Scores for Dataset-II)

Öznitelik Adı	MI Skoru
Takipçi Sayısı	0.0087
Takip Edilen Kişi Sayısı	0.0000
Gönderi İçeriğinin Etiketler ile İlgisi	0.0838
Birden Fazla Etiketin Birbirleri ile Olan İlgisi	0.0907
Gönderi İçeriğinde Yer Alan Fotoğrafın İnternette Yer Alan Diğer Sayfalarda Birlikte Geçtiği Metin Arasındaki Benzerlik	0.0179

Veri Seti-II için yapılan analizlerde, en yüksek MI skoruna sahip olan "Birden Fazla Etiketin Birbirleri ile Olan İlgisi" (0.0907) ve "Gönderi İçeriğinin Etiketler ile İlgisi" (0.0838) öznitelikleri, modelin öğrenme sürecinde en belirleyici faktörler olarak öne çıkmaktadır. Bu durum, gönderi içeriğinin etiketlerle ve etiketlerin birbirleriyle olan ilişkilerinin, modelin tahmin gücünü artırdığını göstermektedir. Buna karşılık, "Takipçi Sayısı" (0.0087) ve "Gönderi İçeriğinde Yer Alan Fotoğrafın İnternette Yer Alan Diğer Sayfalarda Birlikte Geçtiği Metin Arasındaki Benzerlik" (0.0179) gibi değişkenlerin etkisi daha sınırlı kalmıştır.

Özellikle "Takip Edilen Kişi Sayısı" (MI = 0.0000) özniteliğinin model için hiçbir bilgi sağlamadığı görülmüş ve Veri Seti-II'den çıkarılmıştır. Bu durum, ilgili değişkenin hedef değişken ile anlamlı bir ilişkiye sahip olmadığını ve modelin tahmin gücüne herhangi bir katkı sunmadığını ortaya koymaktadır.

2.1.9.Veri setinin Etiketlenmesi

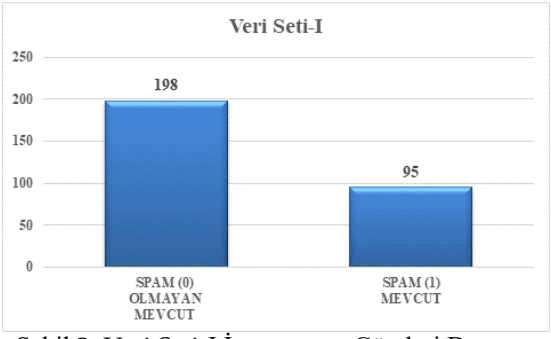
Veri seti, 'Gönderi İçeriğinde Yer Alan Fotoğrafın X Platformu Üzerinde Paylaşılma Durumu' özniteliğinin Google Cloud Vision API ile oluşturulma durumuna göre ikiye ayrılmıştır: 293 gönderiden oluşan Veri Seti-I ve 2188 gönderiden oluşan Veri Seti-II. Bu işlem sonucunda, Veri Seti-I'de 5 öznitelik, Veri Seti-II'de ise 4 öznitelik yer almaktadır.

Veri setinin etiketlenmesi sürecinde, gönderilerin istenmeyen gönderi olup olmadığını belirlemek için detaylı bir inceleme yapılmıştır. Bu inceleme, Cumhurbaşkanlığı Dijital Dönüşüm Ofisi'nin Sosyal Medya Kullanım Kılavuzu'ndaki 'spamcılar' tanımına dayandırılmıştır. Kılavuzda, spamcılar; *'Her türlü içeriğe ilgili ilgisiz yorum yapan, kendi sayfasını veya profilini ziyarete davet eden, takipçi kazanmak için takip eden, genellikle sadece beğeni ve repost yapıp içerik üretmeyen, paylaşımlarının altına konuyla tamamen ilgisiz iletiler yazan ve gündem hashtag'lerini kullanarak, bu hashtag'lerle alakasız paylaşımlar yapan'* kullanıcılar olarak tanımlanmaktadır [29]. Bu tanıma göre, hashtag içerikleriyle alakasız paylaşımlar istenmeyen gönderi olarak değerlendirilmiş ve bu tür paylaşımlar istenmeyen gönderi olarak sınıflandırılmıştır.

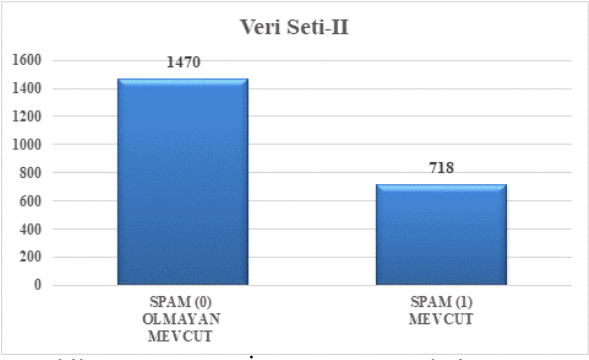
İkinci aşamada, veri setindeki tüm gönderiler detaylı bir şekilde incelenmiş ve her biri kılavuz kriterlerine göre etiketlenmiştir. Bu süreçte, gönderilerin içerdiği metinler ve fotoğraflar analiz edilmiş, özellikle hashtaglerin gönderi içeriğiyle uyumuna dikkat edilmiştir. Gönderinin hashtag ile doğrudan alakalı olmaması durumunda, gönderi istenmeyen olarak etiketlenmiştir.

Son olarak, etiketlenen veriler, makine öğrenmesi modellerinin eğitim ve test aşamalarında kullanılmak üzere düzenlenmiştir. İstenmeyen gönderiler (1) ve istenmeyen olmayan gönderiler (0) şeklinde iki ana kategoriye ayrılmıştır. Bu düzenleme, makine öğrenmesi modellerinin

istenmeyen gönderi tespitinde daha doğru sonuçlar elde etmesine yardımcı olmuştur. Veri Seti-I ve Veri Seti-II'ye yönelik spam ve spam olmayan gönderilerin dağılımı, Şekil 2 ve Şekil 3'te gösterilmektedir. Ayrıca, her iki veri setine ilişkin özet bilgiler Tablo 3'te sunulmaktadır.



Şekil 2. Veri Seti-I İstenmeyen Gönderi Durumu
(Data Set-I Spam Status)



Şekil 3. Veri Seti-II İstenmeyen Gönderi Durumu
(Data Set-II Spam Status)

Tablo 3. Veri Setleri Özet Durum
(Summary Status of Datasets)

Veri Seti Adı	Öznitelik Sayısı	İstenmeyen Gönderi (1)	İstenmeyen Olmayan Gönderi (0)	Genel Toplam
Veri Seti-I	5	95	198	293
Veri Seti-II	4	718	1470	2188

2.2. Makine Öğrenmesi Yöntemleri

Bu çalışmada Karar Ağacı, Rastgele Orman, Destek Vektör Makineleri (SVM - Support Vector Machine), Lojistik Regresyon ve Çok Katmanlı Algılayıcı (MLP) algoritmaları kullanılmıştır.

2.2.1.Karar Ağacı Algoritması

Karar Ağacı algoritması, makine öğrenmesinde sıkça kullanılan ve açıklanabilirliği yüksek olan bir yöntemdir. Verileri belirli özelliklere göre dallara ayırarak karar verme sürecini basitleştiren bu algoritma, sınıflandırma ve regresyon gibi çeşitli problemlerde yaygın olarak kullanılmaktadır. Karar ağaçları, veri setlerindeki en üst (kök) düğümden başlayarak her düğümde belirli bir özelliğe göre karar alır ve yaprak düğümlere ulaşana kadar dallanır. Yaprak düğümler, nihai bir sınıf veya değerin tahminini temsil eder. Açıklanabilirliği yüksek olan bu algoritma, özellikle sınıflandırma problemlerinde basit ve etkili bir çözüm sunar. Bu çalışmada, özellikle metin ve görsel özniteliklerin etkileşimini anlamak için tercih edilmiştir.

2.2.2.Rastgele Orman Algoritması

Rastgele Orman algoritması, makine öğrenmesinde yaygın olarak kullanılan bir ensemble (topluluk) sınıflandırıcıdır [30]. Her bir ağaç, veri kümesinin rastgele alt örnekleri ve özellikleri üzerine odaklanır. Tahmin sonuçları bir araya getirilerek, ensemble ortalaması alınır ve bu yöntemle daha doğru ve tutarlı bir sınıflandırma sağlanır [31]. Rastgele Orman algoritması, aşırı uyuma karşı dayanıklıdır ve diğer modellerden daha az hiperparametre ayarı gerektirir. Büyük ve karmaşık veri setlerinde etkilidir, bu nedenle tıbbi teşhislerden finansal tahminlere kadar çeşitli alanlarda başarıyla kullanılmaktadır. Gürültüyü azaltarak model performansını artırdığı için yüksek doğruluk gerektiren veri setlerinde sıklıkla tercih edilir. Bu nedenle, dengesiz veri setlerinde daha kararlı sonuçlar elde etmek için çalışmamızda tercih edilmiştir.

2.2.3.SVM Algoritması

SVM algoritması, makine öğrenmesinde özellikle sınıflandırma problemlerinde sıkça kullanılan önemli bir algoritmadır. SVM, veri noktalarını bir hiper düzlem ile ayırarak en iyi sınıflandırmayı sağlamak amacıyla destek vektörlerinden yararlanır. Sürekli özelliklerin olduğu veri setlerinde etkili olup, sınıflandırma sırasında veri noktalarını ayırmak için özel bir optimizasyon yöntemi kullanır. SVM, metin sınıflandırma, görüntü tanıma, biyomedikal veri analizi ve finansal analiz gibi alanlarda yaygın olarak kullanılmaktadır. Esnekliği ve yüksek performansı, bu algoritmayı geniş bir uygulama alanına sahip kılar. Özellikle yüksek boyutlu veri setlerinde etkili olan SVM, metin ve görsel özniteliklerin birleşik analizinde iyi bir performans sergilemesi beklenerek, bu çalışmada tercih edilmiştir.

2.2.4.Lojistik Regresyon Algoritması

Lojistik regresyon, ikili sınıflandırma problemlerinde yaygın olarak kullanılan güçlü ve açıklanabilir bir makine öğrenmesi algoritmasıdır. Bu yöntem, bağımsız değişkenlerle bağımlı değişken arasında doğrusal bir ilişki kurarak sonuçları 0 ile 1 arasında bir olasılık değeri olarak ifade eder. Model, bir veri noktasının özelliklerinin ağırlıklı toplamını hesaplayarak bu toplamı sigmoid fonksiyonuna uygular; böylece, pozitif sınıfa ait olma olasılığını tahmin eder. Lojistik regresyonun avantajlarından biri, katsayıların da yorumlanabilir olmasıdır; her katsayı, bağımsız değişkenin bağımlı değişken üzerindeki etkisi hakkında bilgi verir. Ancak, doğrusal olarak ayrılabilir olmayan veri setlerinde performansı sınırlıdır. Basit ve etkili bir algoritma olarak

lojistik regresyon, birçok alanda kullanılabilecek şekilde tasarlanmıştır. Bu çalışmada özellikle metin tabanlı özniteliklerin analizinde etkili olması beklendiği için tercih edilmiştir.

2.2.5.Çok Katmanlı Algılayıcı Algoritması

Çok Katmanlı Algılayıcı (MLP), makine öğrenmesinde kullanılan bir yapay sinir ağı türüdür ve karmaşık sınıflandırma problemlerini çözmekte oldukça etkilidir. MLP, girdi katmanı, bir veya daha fazla gizli katman ve çıktı katmanından oluşur. Girdi katmanındaki her nöron bağımsız değişkenleri temsil ederken, bu değerler gizli katmanlardaki nöronlara iletilir. Gizli katmanlar, girdileri işleyip aktivasyon fonksiyonlarını uygulayarak daha yüksek seviyede özellikler çıkarır. Son olarak, çıktı katmanı, sınıflandırma veya regresyon problemlerine yönelik nihai tahminleri üretir. Metin ve görsel özniteliklerin birlikte analiz edilmesi sürecinde daha derin öğrenme yetenekleri sunması nedeniyle bu çalışmada tercih edilmiştir.

2.3. Makine Öğrenmesi Yöntemlerinin Performans Değerlendirme Kriterleri

Bir algoritmanın başarısını değerlendirilebilmesi için eğitim aşamasından sonra, sınıflandırma modelinin tahmin doğruluğunun ölçülmesi gerekmektedir. Bu çalışmada, hata matrisi oluşturularak, doğruluk, duyarlılık, kesinlik ve F1 skoru gibi kriterler ile sınıflandırma modellerinin performansları karşılaştırılmaktadır.

2.3.1.Hata Matrisi

Hata matrisi, makine öğrenmesi ve özellikle sınıflandırma problemlerinde model performansını değerlendirmek için önemli bir araç olarak kullanılmaktadır. Bir hata matrisi, modelin doğru ve yanlış tahminlerini tablo şeklinde özetlemektedir ve bu sayede modelin başarısı hakkında bilgi sağlamaktadır. Hata matrisi, dört temel bileşenden oluşarak Tablo 4'te gösterilmektedir.

Tablo 4. Hata Matrisi
(Confusion Matrix)

Prensip	Tahmin Edilen (Pozitif)	Tahmin Edilen (Negatif)
Gerçek Pozitif	Doğru Pozitif (TP)	Yanlış Negatif (FN)
Gerçek Negatif	Yanlış Pozitif (FP)	Doğru Negatif (TN)

2.3.2.Doğruluk

Doğruluk, bir sınıflandırma modelinin tüm tahminlerinin ne kadar doğru olduğunu ölçen temel bir performans metriği olmaktadır. Doğruluk, doğru tahminlerin (hem TP hem de TN) toplam tahminlere oranı olarak hesaplanmaktadır. Yüksek doğruluk, modelin genel olarak iyi performans gösterdiğini belirtmektedir. Bu noktada doğruluk, diğer metriklerle birlikte desteklenerek değerlendirilmektedir. Doğruluk ölçütünün denklemi Eşitlik (1)'de ifade edilmiştir:

$$Doğruluk = \frac{TP+TN}{TP+TN+FP+FN} \quad (1)$$

2.3.3.Duyarlılık

Duyarlılık, aynı zamanda "hassasiyet" veya "true positive rate" olarak da isimlendirilir ve modelin, pozitif sınıfları ne kadar iyi tespit ettiğini ölçmektedir. Yüksek duyarlılık, modelin yanlış negatifleri (FN) minimize ettiğini göstermektedir. Duyarlılık ölçütünün denklemi Eşitlik (2)'de ifade edilmiştir:

$$Duyarlılık = \frac{\mathrm{TP}}{\mathrm{TP+FN}} \quad (2)$$

2.3.4.Kesinlik

Kesinlik, modelin pozitif tahminlerinin ne kadar doğru olduğunu göstermektedir. Yani, modelin pozitif olarak sınıflandırdığı örneklerin gerçekten pozitif olma oranını ifade etmektedir. Yüksek kesinlik, modelin yanlış pozitifleri (FP) minimize ettiğini ifade etmektedir ve özellikle istenmeyen gönderi filtreleme veya hastalık teşhisi gibi yanlış pozitiflerin maliyetli olduğu durumlarda kullanılmaktadır. Kesinlik ölçütünün denklemi Eşitlik (3)'te ifade edilmiştir:

$$Kesinlik = \frac{TP}{TP+FP} \quad (3)$$

2.3.5.F1 Skor

F1 Skor, kesinlik ve duyarlılığı tek bir metrikte birleştirerek modelin genel performansını özetlemeyi hedeflemektedir. F1 Skor, özellikle sınıflar arasında dengesizlik olduğunda faydalı olmakla birlikte kesinlik ve duyarlılık arasındaki dengeyi sağlamaktadır. F1 Skor ölçütünün denklemi Eşitlik (4)'te ifade edilmiştir:

$$F1\ Skor = 2 \times \frac{TP}{2\times TP+FP+FN} \quad (4)$$

2.3.6.ROC AUC (ROC Eğrisi Altındaki Alan)

ROC AUC (Receiver Operating Characteristic Area Under the Curve), bir modelin sınıflandırma performansını değerlendirmek için kullanılan önemli bir metrik olmaktadır. ROC Eğrisi, modelin farklı eşik değerlerinde duyarlılık (true positive rate) ve yanlış pozitif oranını (false positive rate) göstermektedir. ROC Eğrisi altındaki alan (AUC), modelin genel performansını özetlemekle birlikte 0.5 ile 1 arasında bir değer almaktadır. 0.5 değeri, modelin rastgele tahmin yaptığını gösterirken, 1 değeri mükemmel ayırma kapasitesini belirtmektedir.

Bu bölümde, veri setinin oluşturulma süreci, kullanılan makine öğrenmesi yöntemleri ve performans değerlendirme kriterleri açıklanmıştır. Elde edilen veri seti ve uygulanan algoritmalar, spam tespiti ve sosyal medya analizlerinde önemli bir katkı sağlamaktadır. Bir sonraki bölümde, bu yöntemlerle elde edilen sonuçlar ve modellerin karşılaştırmalı analizi sunulacaktır.

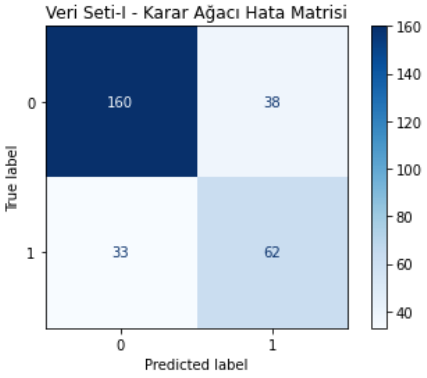
3. DENEYSEL SONUÇLAR (EXPERIMENTAL RESULTS)

Bu bölümde, makine öğrenmesi algoritmalarının performansını değerlendirmek amacıyla oluşturulan Veri Seti-I ve Veri Seti-II üzerinde gerçekleştirilen deneysel analizler sunulmaktadır.

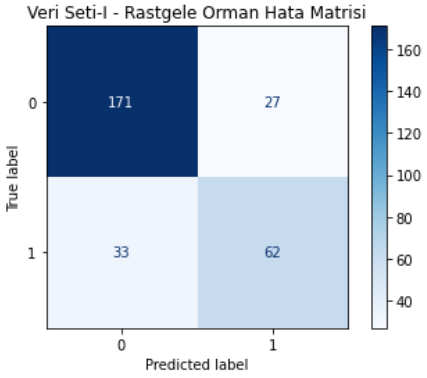
Veri seti, "Gönderi İçeriğinde Yer Alan Fotoğrafın X Platformu Üzerinde Paylaşılma Durumu" özniteliğinin Google Cloud Vision API ile oluşturulma durumuna göre ikiye ayrıldığından, bu bölümde Veri Seti-I ve Veri Seti-II olarak ayrı ayrı incelenmiştir.

3.1. Veri Seti-I

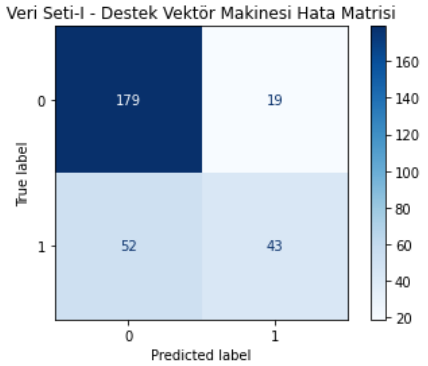
Makine öğrenmesi algoritmaları kullanılarak oluşturulan modelin test tahminleri, test verileri ile karşılaştırıldığında, Veri Seti-I için Şekil 4-8 aralığında gösterilmekte olan hata matrisi elde edilmektedir.



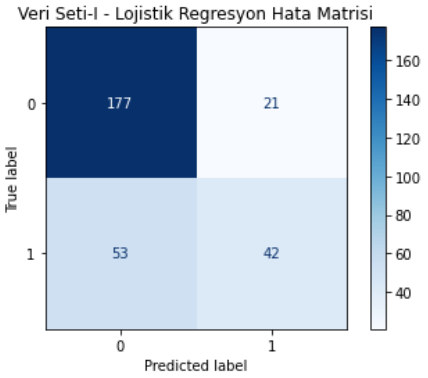
Şekil 4. Veri Seti-I Karar Ağacı Hata Matrisi
(Dataset-I Decision Tree Confusion Matrix)



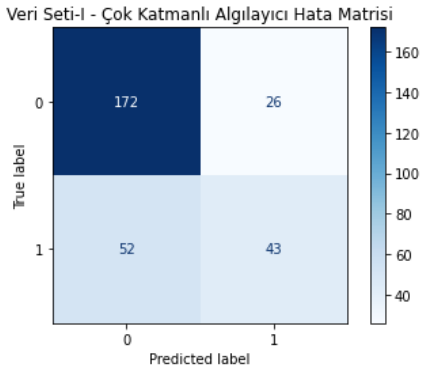
Şekil 5. Veri Seti-I Rastgele Orman Hata Matrisi
(Dataset-I Random Forest Confusion Matrix)



Şekil 6. Veri Seti-I SVM Hata Matrisi
(Dataset-I SVM Confusion Matrix)



Şekil 7. Veri Seti-I Lojistik Regresyon Hata Matrisi
(Dataset-I Logistic Regression Confusion Matrix)

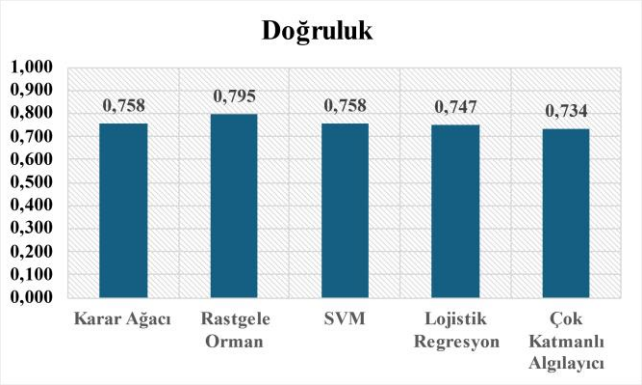


Şekil 8. Veri Seti-I Çok Katmanlı Algılayıcı Hata Matrisi
(Dataset-I MLP Confusion Matrix)

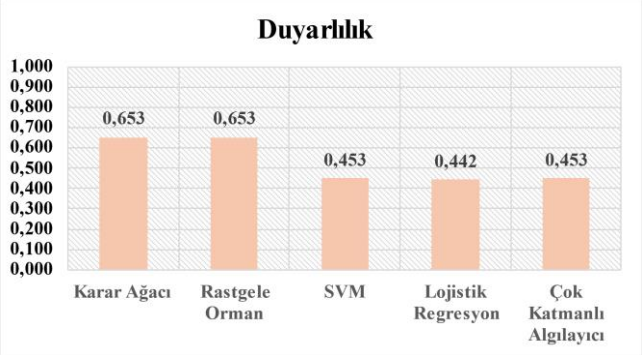
Veri Seti-I'e yönelik performans metrikleri Tablo 5'te ve Şekil 9-13 aralığında gösterilmektedir.

Tablo 5. Veri Seti-I Performans Metrikleri
(Dataset-I Performance Metrics)

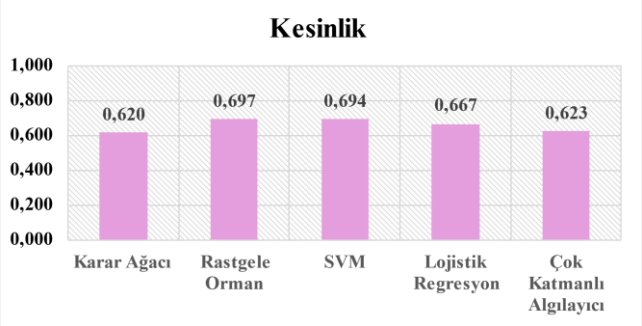
Algoritma	Doğruluk	Duyarlılık	Kesinlik	F1 Skor	ROC AUC
Karar Ağacı	0,758	0,653	0,620	0,632	0,727
Rastgele Orman	0,795	0,653	0,697	0,667	0,838
SVM	0,758	0,453	0,694	0,536	0,692
Lojistik Regresyon	0,747	0,442	0,667	0,526	0,772
MLP	0,734	0,453	0,623	0,519	0,754



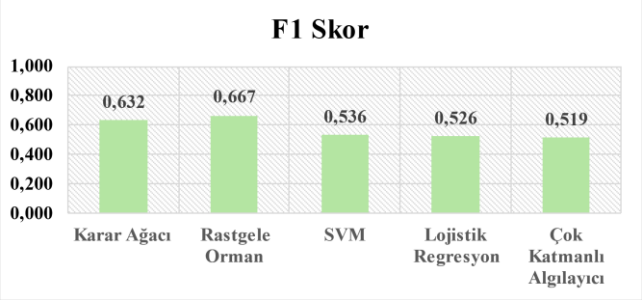
Şekil 9. Veri Seti-I Doğruluk Metrikleri
(Dataset-I Accuracy Metrics)



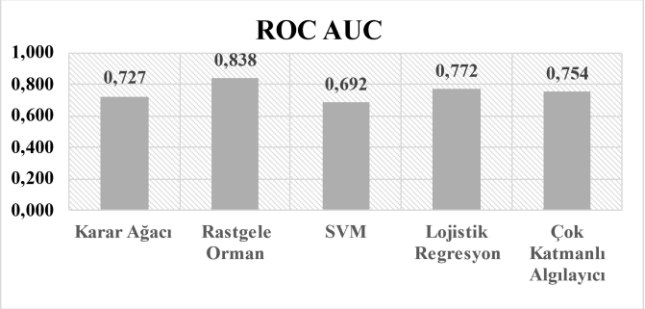
Şekil 10. Veri Seti-I Duyarlılık Metrikleri
(Dataset-I Recall Metrics)



Şekil 11. Veri Seti-I Kesinlik Metrikleri
(Dataset-I Precision Metrics)



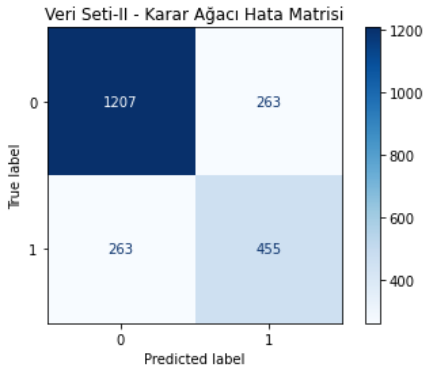
Şekil 12. Veri Seti-I F1 Skor Metrikleri
(Dataset-I F1 Score Metrics)



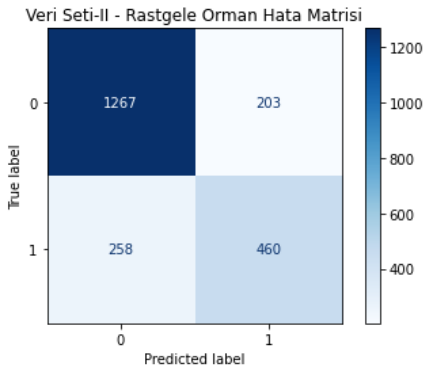
Şekil 13. Veri Seti-I ROC AUC Metrikleri
(Dataset-I ROC AUC Metrics)

3.2. Veri Seti-II

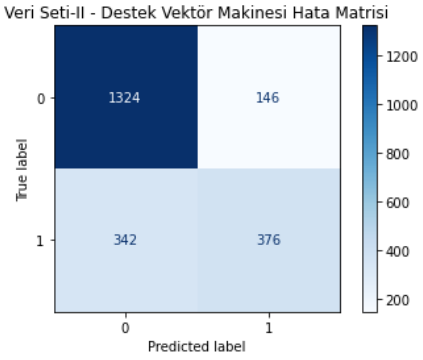
Makine öğrenmesi algoritmaları kullanılarak oluşturulan modelin test tahminleri, test verileri ile karşılaştırıldığında, Veri Seti-II için Şekil 14-18 aralığında gösterilmekte olan hata matrisi elde edilmektedir.



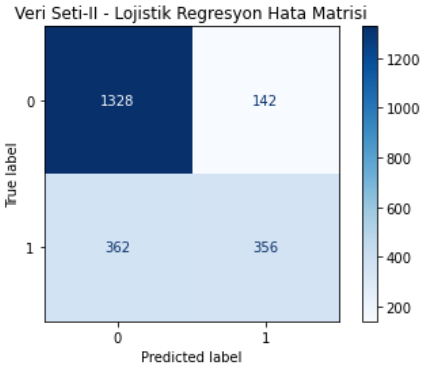
Şekil 14. Veri Seti-II Karar Ağacı Hata Matrisi
(Dataset-II Decision Tree Confusion Matrix)



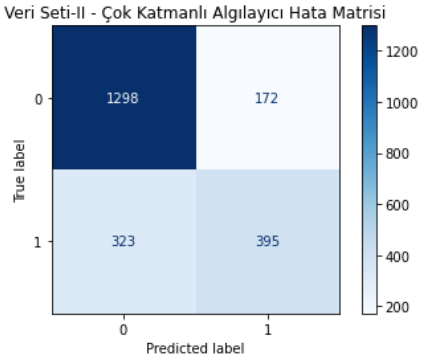
Şekil 15. Veri Seti-II Rastgele Orman Hata Matrisi
(Dataset-II Random Forest Confusion Matrix)



Şekil 16. Veri Seti-II SVM Hata Matrisi
(Dataset-II SVM Confusion Matrix)



Şekil 17. Veri Seti-II Lojistik Regresyon Hata Matrisi
(Dataset-II Logistic Regression Confusion Matrix)

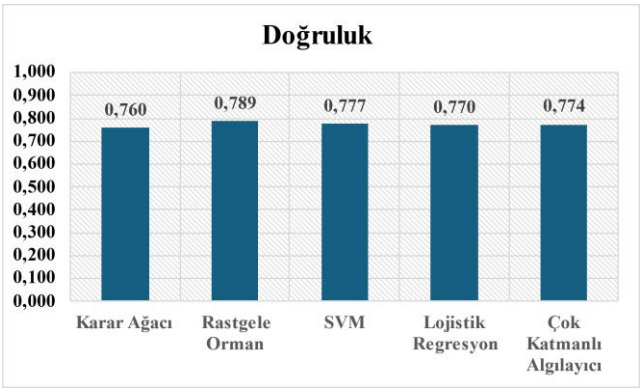


Şekil 18. Veri Seti-II Çok Katmanlı Algılayıcı Hata Matrisi
(Dataset-II MLP Confusion Matrix)

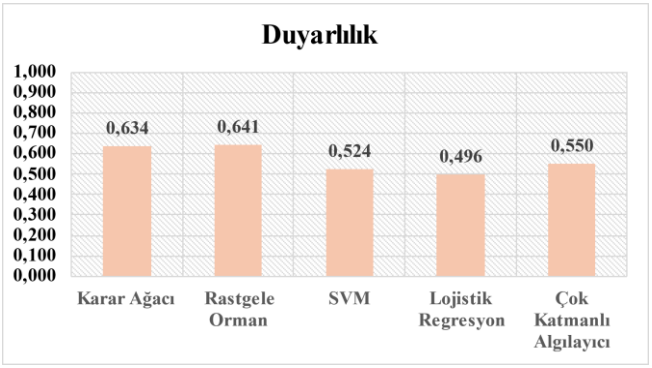
Veri Seti-II'e yönelik performans metrikleri Tablo 6'da ve Şekil 19-23 aralığında gösterilmektedir.

Tablo 6. Veri Seti-II Performans Metrikleri
(Dataset-II Performance Metrics)

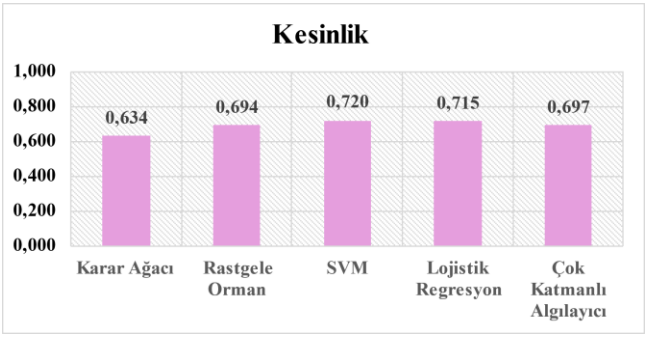
Algoritma	Doğruluk	Duyarlılık	Kesinlik	F1 Skor	ROC AUC
Karar Ağacı	0,760	0,634	0,634	0,635	0,728
Rastgele Orman	0,789	0,641	0,694	0,667	0,842
SVM	0,777	0,524	0,720	0,606	0,757
Lojistik Regresyon	0,770	0,496	0,715	0,585	0,763
MLP	0,774	0,550	0,697	0,615	0,802



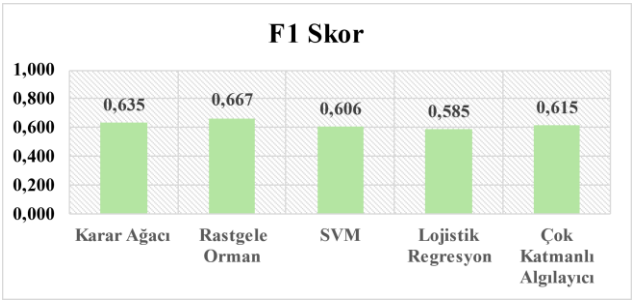
Şekil 19. Veri Seti-II Doğruluk Metrikleri
(Dataset-II Accuracy Metrics)



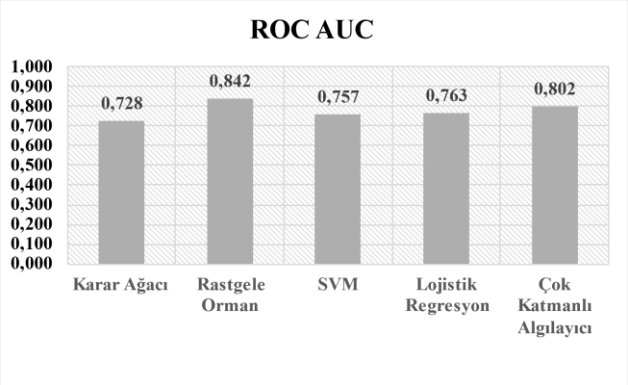
Şekil 20. Veri Seti-II Duyarlılık Metrikleri
(Dataset-II Recall Metrics)



Şekil 21. Veri Seti-II Kesinlik Metrikleri
(Dataset-II Precision Metrics)



Şekil 22. Veri Seti-II F1 Skor Metrikleri
(Dataset-II F1 Score Metrics)



Şekil 23. Veri Seti-II ROC AUC Metrikleri
(Dataset-II ROC AUC Metrics)

3.3. Algoritmaların Veri Setleri Üzerindeki Performans Farklılıkları ve Nedenleri

Bu çalışmada kullanılan iki farklı veri seti, içerik ve büyüklük açısından farklılık göstermektedir. Veri Seti-I (293 gönderi) daha küçük ve dengesiz bir yapıya sahipken, Veri Seti-II (2188 gönderi) daha büyük ve dengesiz bir dağılıma sahiptir. Yapılan analizler, makine öğrenmesi algoritmalarının bu farklı veri setleri üzerinde farklı performans gösterdiğini ortaya koymuştur.

Özellikle Rastgele Orman algoritması, her iki veri setinde de en yüksek genel performansa ulaşmış, doğruluk, duyarlılık ve F1 skoru açısından en iyi sonuçları elde etmiştir. Bunun temel nedenleri şunlar olabilir:

- Veri Seti-II'nin daha büyük olması, bazı algoritmaların (SVM, Lojistik Regresyon, MLP) daha iyi genelleme yapmasını sağlamış, ancak Rastgele Orman ve Karar Ağacı için belirgin bir fark yaratmamıştır.
- Veri dengesizliği, özellikle Veri Seti-I üzerinde SVM ve Lojistik Regresyon'un performansını olumsuz etkilemiştir. Veri Seti-II'de bu algoritmaların F1 skorlarında iyileşme görülmüştür.
- Karar Ağaçları ve Rastgele Orman gibi ağaç tabanlı yöntemler, veri dengesizliği ile daha iyi başa çıkabilmiştir. Ağaç tabanlı yöntemler, her sınıfa özel ayrım noktaları belirleyerek dengesiz veri setlerinde daha kararlı sonuçlar üretebilir.
- Veri Seti-I'de F1 skorlarının düşük olması, azınlık sınıfının (spam) yeterince iyi öğrenilememesinden kaynaklanmıştır. Bu durum, Veri Seti-I için SMOTE (Synthetic Minority Over-sampling Technique) gibi veri dengeleme yöntemlerinin kullanılmasını gerektirebilir.

Bu bölümde elde edilen sonuçlar doğrultusunda, farklı makine öğrenmesi algoritmalarının spam tespiti üzerindeki performansları karşılaştırılmıştır. Özellikle Destek Vektör Makineleri, Lojistik Regresyon ve Çok Katmanlı Algılayıcı (MLP), Veri Seti-II üzerinde daha yüksek F1 skoru ve duyarlılık göstermiştir. Rastgele Orman ve Karar Ağacı algoritmalarında ise veri setleri arasında belirgin bir üstünlük gözlenmemiştir. Bir sonraki bölümde, elde edilen bulguların genel bir değerlendirmesi yapılarak sonuçlar tartışılacaktır.

4. SONUÇ VE TARTIŞMA (RESULTS AND DISCUSSION)

Bu çalışmada, X sosyal medya platformundaki istenmeyen gönderilerin tespitine yönelik makine öğrenmesi, geniş dil modelleri ve bilgisayarlı görü tekniklerini birleştiren yeni bir yaklaşım geliştirilmiştir. Yapılan analizlerde, dengesiz yapıya sahip iki farklı veri seti (Veri Seti-I ve Veri Seti-II) üzerinde çeşitli makine öğrenmesi algoritmalarının performansları değerlendirilmiştir. Veri Seti-I 293 örnekten (95 istenmeyen gönderi, 198 istenmeyen olmayan gönderi) oluşurken, Veri Seti-II 2188 örnekten (718 istenmeyen gönderi, 1470 istenmeyen olmayan gönderi) oluşmaktadır. Performans değerlendirmelerinde, özellikle spam tespiti gibi dengesiz veri setlerinde, yalnızca doğruluğa (accuracy) odaklanmak yanıltıcı olabileceğinden, sınıflar arası dengesizliklere duyarlı olan F1 skoru, ROC AUC, duyarlılık (recall) ve kesinlik (precision) gibi kapsamlı ölçütler tercih edilmiştir. Bu yaklaşım, sınıflar arasındaki dengesizliklerin yanlış yorumlanmasını engellemeyi ve sonuçların daha tutarlı bir şekilde değerlendirilmesini sağlamayı hedeflemiştir.

Her iki veri seti üzerinde yapılan analizler sonucunda, Rastgele Orman algoritmasının en dengeli ve güçlü sonuçları sunduğu görülmüştür. Veri Seti-I'de ROC AUC değeri (0,838), F1 skoru (0,667) ve kesinlik değeri (0,697) ile en iyi sonuçları elde ederken, Veri Seti-II'de ROC AUC değeri (0,842), F1 skoru (0,667) ve kesinlik değeri (0,694) ile diğer modelleri geride bırakmıştır. Bu bulgular, Rastgele Orman algoritmasının hem spam tespiti (duyarlılık) hem de yanlış pozitifleri (FP) minimize etme konusunda dengeli bir performans sunduğunu göstermektedir. Veri Seti-I'de duyarlılık değeri (0,653), Veri Seti-II'de ise (0,641) olarak ölçülmüş olup, her iki veri setinde de FN oranlarının daha da azaltılması için iyileştirme potansiyeli bulunmaktadır. Ayrıca, düşük FP oranları sayesinde iletişim uzmanlarının yanlış pozitifleri inceleme iş yükü azalmakta ve zamanlarını daha verimli kullanmaları sağlanmaktadır.

Karar Ağacı algoritması, her iki veri setinde de ortalama bir performans sergilemiştir. Veri Seti-I'de 0,758 doğruluk ve 0,632 F1 skoru elde edilirken, Veri Seti-II'de 0,760 doğruluk ve 0,635 F1 skoru ile benzer bir başarı yakalanmıştır. Ancak, ROC AUC değerleri (Veri Seti-I: 0,727; Veri Seti-II: 0,728) ve duyarlılık değerleri (Veri Seti-I: 0,653; Veri Seti-II: 0,634), bu algoritmanın dengesiz veri setlerinde spam tespiti açısından yeterince güçlü olmadığını göstermektedir. Özellikle azınlık sınıfına (istenmeyen gönderi) yönelik öğrenme kapasitesinin sınırlı olması nedeniyle spam içerikleri belirleme konusunda zayıf kalmıştır. Bu durum, Karar Ağacı algoritmasının FN oranlarını düşürme konusunda yetersiz kaldığını göstermektedir.

Destek Vektör Makineleri (SVM), her iki veri setinde de düşük duyarlılık değerleri nedeniyle spam tespitinde yetersiz kalmıştır. Veri Seti-I'de duyarlılık değeri (0,453), Veri Seti-II'de ise (0,524) olarak ölçülmüş olup, diğer modellere kıyasla belirgin şekilde daha düşük performans göstermiştir. Bu durum, SVM'nin dengesiz veri setlerinde azınlık sınıfını ayırt etmede zorlandığını ve FN oranlarını artırdığını göstermektedir. Ayrıca, yüksek FP oranları iletişim uzmanlarının yanlış pozitifleri inceleme sürecini uzatarak iş yükünü artırabilir.

Lojistik Regresyon, her iki veri setinde de orta düzeyde bir performans sergilemiştir. Veri Seti-I'de 0,747 doğruluk ve 0,526 F1 skoru, Veri Seti-II'de ise 0,770 doğruluk ve 0,585 F1 skoru elde edilmiştir. Ancak, duyarlılık değerleri (Veri Seti-I: 0,442; Veri Seti-II: 0,496) algoritmanın spam tespitinde zayıf kaldığını göstermektedir. Bu durum, Lojistik Regresyon'un FN oranlarını azaltma konusunda etkisiz olduğunu ve iletişim uzmanlarının iş yükünü artırabilecek yanlış pozitifler (FP) ürettiğini göstermektedir.

Çok Katmanlı Algılayıcı (MLP), Veri Seti-I'de 0,734 doğruluk ve 0,519 F1 skoru ile en düşük performansı sergilerken, Veri Seti-II'de 0,774 doğruluk ve 0,615 F1 skoru ile nispeten daha iyi bir performans göstermiştir. ROC AUC değerleri (Veri Seti-I: 0,754; Veri Seti-II: 0,802) ve duyarlılık değerleri (Veri Seti-I: 0,453; Veri Seti-II: 0,550), bu algoritmanın dengesiz veri setlerinde spam tespiti açısından daha başarılı olduğunu göstermektedir. Bununla birlikte, FP oranlarının daha da optimize edilmesi gerekmektedir.

Elde edilen bulgular değerlendirildiğinde, Rastgele Orman algoritmasının dengesiz veri yapıları için en etkili çözümü sunduğu görülmektedir. Bu algoritma, duyarlılık ve kesinlik arasında optimal bir denge sağlayarak spam tespitinde yüksek performans göstermiştir. Özellikle spam içerikleri kaçırmama hedefi doğrultusunda, Rastgele Orman'ın yüksek duyarlılık (recall) değerleri FN oranlarını minimize etme konusunda etkili olmuştur. Ayrıca, düşük FP oranları sayesinde iletişim uzmanlarının yanlış pozitifleri inceleme yükü azalarak iş süreçlerinde verimlilik sağlanmaktadır.

Bu çalışma, makine öğrenmesi ve bilgisayarlı görü yöntemlerinin entegrasyonu sayesinde, spam tespiti alanında hem teorik hem de pratik önemli katkılar sunmaktadır. Teorik açıdan, metin ve görsel içeriklerin birlikte analiz edilmesi, geleneksel yaklaşımlara kıyasla daha kapsamlı bir çözüm önermektedir.

Pratikte ise, özellikle sosyal medya platformlarında otomatik spam tespit sistemlerinin geliştirilmesine yönelik önemli çıktılar sağlamaktadır. Ancak, modelin genelleştirilebilirliği konusunda sınırlamalar bulunmaktadır. Çalışmanın veri seti yalnızca Türkiye'de belirli bir zaman aralığında paylaşılan görselleri içermekte ve farklı platformlar, diller ve kültürel bağlamlarda test edilmemiştir. Bu durum, modelin yerel bağlamda etkili sonuçlar vermesini sağlamakla birlikte, uluslararası platformlarda uygulanabilirliği konusunda bazı sınırlamalar ortaya çıkarmaktadır. Özellikle, farklı kültürel bağlamlar, dil yapıları ve sosyal medya kullanım alışkanlıkları, modelin performansını etkileyebilir. Ancak, geliştirilen yöntemlerin (metin ve görsel analiz tekniklerinin entegrasyonu gibi) evrensel olarak uygulanabilir olduğu düşünülmektedir. Bu nedenle, geniş ölçekli uygulamalar için çapraz platform analizleri ve uzun vadeli verilerle doğrulama yapılması gerekmektedir.

Gelecekteki çalışmalarda, daha derin sinir ağı tabanlı yaklaşımlar (örneğin, Transformer modelleri veya CNN tabanlı görsel analiz teknikleri) ile mevcut yöntemin karşılaştırılması önemli bir adım olabilir. Özellikle BERT veya CLIP gibi modellerin metin ve görsel içerikleri birlikte analiz etme yetenekleri, spam tespit sistemlerinin doğruluğunu artırabilir. Ayrıca, gerçek zamanlı spam tespiti için işlem süresi ve hesaplama maliyetlerini optimize edecek stratejiler geliştirilmesi, bu tür sistemlerin pratik uygulanabilirliğini artıracaktır. Daha büyük ölçekli ve dinamik olarak güncellenen veri setleri üzerinde çapraz doğrulama yapılması, modelin genelleştirilebilirliğini ve güvenilirliğini artırabilir. Bunun yanı sıra, farklı platformlardan (Instagram, Facebook gibi) toplanacak çeşitli veri setleri üzerinde karşılaştırmalı analizler yapılarak, platformlar arası spam davranışlarının benzerlikleri ve farklılıkları incelenebilir. Bu tür çalışmalar, spam tespit sistemlerinin farklı sosyal medya ortamlarında daha etkili ve tutarlı bir şekilde çalışmasını sağlayabilir.

Ayrıca çalışmada kullanılan Google Gemini ve Cloud Vision AI gibi araçların alternatifleriyle (örneğin OpenAI GPT, Microsoft Azure Computer Vision, Amazon Rekognition) karşılaştırmalı incelemeler yapılması da önemli bir katkı sağlayabilir. Farklı modellerin performanslarının aynı veri setleri üzerinde test edilmesi, hangi araçların belirli senaryolarda daha etkili olduğunu ortaya koyabilir. Bu tür karşılaştırmalar, spam tespiti için en uygun araç seçimini kolaylaştırabilir ve daha geniş bir perspektif sunabilir.

Sonuç olarak, bu çalışma istenmeyen içerik tespiti alanında önemli bir temel oluşturarak, daha kapsamlı araştırmalar için bir yol haritası sunmuştur. Gelişmiş algoritmalar, daha zengin veri setleri, gerçek zamanlı analiz yöntemleri ve alternatif araçların karşılaştırmalı değerlendirilmesiyle, bu alanda daha etkili çözümler geliştirilmesi mümkün olacaktır.

KAYNAKLAR (REFERENCES)

[1] P. Sharma, T. Nagpal, G. Shrivastava, J. D. Kumar, "A Systematic Review on Social Bots Account Detection Using Machine Learning," *2023 5th International Conference on Advances in Computing, Communication Control and Networking* (ICAC3N), pp. 862–866, 2023.

[2] S. Ghosh, B. Viswanath, F. Kooti, N. K. Sharma, K. P. Gummadi, B. Krishnamurthy, "Understanding and combating link farming in

the twitter social network," *The Web Conference* (WWW), pp. 61–70, 2012.

[3] N. L. Jin, N. Y. Chen, N. T. Wang, N. P. Hui, A. V. Vasilakos, "Understanding user behavior in online social networks: a survey," *IEEE Communications Magazine*, vol. 51, no. 9, pp. 144–150, 2013.

[4] K. Thomas, C. Grier, D. Song, V. Paxson, "Suspended accounts in retrospect," 12th ACM Workshop on Hot Topics in Networks (HotNets), pp. 1–7, 2011.

[5] A. Aggarwal, A. Rajadesingan, P. Kumaraguru, "PhishAri: Automatic real-time phishing detection on twitter," eCrime Researchers Summit, pp. 1–12, 2012.

[6] D. Wang, S. Navathe, L. Liu, D. Irani, A. Tamersoy, C. Pu, "Click Traffic Analysis of Short URL Spam on Twitter," **IEEE International Conference on Big Data, pp**. 1–8, 2013.

[7] C. Grier, K. Thomas, V. Paxson, M. Zhang, "@spam," **Proceedings of the 2010 ACM Conference on Computer and Communications Security**, pp. 27–37, 2010.

[8] T. Wu, S. Liu, J. Zhang, Y. Xiang, "Twitter spam detection based on deep learning," Proceedings of the Australasian Computer Science Week Multiconference (ACSW), pp. 1–9, 2016.

[9] M. Mateen, M. A. Iqbal, M. Aleem, M. A. Islam, "A hybrid approach for spam detection for Twitter," **2022 19th International Bhurban Conference on Applied Sciences and Technology** (IBCAST), pp. 106–111, 2017.

[10] S. Sedhai, A. Sun, "Semi-Supervised Spam Detection in Twitter Stream," IEEE Transactions on Computational Social Systems, vol. 5, no. 1, pp. 169–175, 2017.

[11] I. Inuwa-Dutse, M. Liptrott, I. Korkontzelos, "Detection of spam-posting accounts on Twitter," Neurocomputing, vol. 315, pp. 496–511, 2018.

[12] S. Madisetty, M. S. Desarkar, "A Neural Network-Based Ensemble Approach for Spam Detection in Twitter," IEEE Transactions on Computational Social Systems, vol. 5, no. 4, pp. 973–984, 2018.

[13] R. Kaur, S. Singh, H. Kumar, "Rise of spam and compromised accounts in online social networks: A state-of-the-art review of different combating approaches," *Journal of Network and Computer Applications, vol.* 112, pp. 53–88, 2018.

[14] K. Binsaeed, G. Stringhini, A. Eleyan, "Detecting Spam in Twitter Microblogging Services: A Novel Machine Learning Approach based on Domain Popularity," *International Journal of Advanced Computer Science and Applications, vol*. 11, no. 11, pp. 1–7, 2020.

[15] K. S. Adewole, T. Han, W. Wu, H. Song, A. K. Sangaiah, "Twitter spam account detection based on clustering and classification methods," *The Journal of Supercomputing, vol.* 76, no. 7, pp. 4802–4837, 2018.

[16] N. El-Mawass, P. Honeine, L. Vercouter, "SimilCatch: Enhanced social spammers detection on Twitter using Markov Random Fields," Information Processing & Management, vol. 57, no. 6, p. 102317, 2020.

[17] Y. Kontsewaya, E. Antonov, A. Artamonov, "Evaluating the effectiveness of machine learning methods for spam detection," Procedia Computer Science, vol. 190, pp. 479–486, 2021.

[18] K. U. Santoshi, S. S. Bhavya, Y. B. Sri, B. Venkateswarlu, "Twitter Spam Detection Using Naïve Bayes Classifier," *International Conference on Inventive Computation Technologies, pp*. 773–777, 2021.

[19] S. B. Abkenar, E. Mahdipour, S. M. Jameii, M. H. Kashani, "A hybrid classification method for Twitter spam detection based on differential evolution and random forest," Concurrency and Computation: Practice and Experience, vol. 33, no. 21, 2021.

[20] C. Kumar, T. S. Bharti, S. Prakash, "A hybrid Data-Driven framework for Spam detection in Online Social Network," Procedia Computer Science, vol. 218, pp. 124–132, 2023.

[21] M. Sumathi, S. P. Raja, "Machine learning algorithm-based spam detection in social networks," Social Network Analysis and Mining, vol. 13, no. 1, 2023.

[22] M. Thomas, B. B. Meshram, "ChSO-DNFNet: Spam detection in Twitter using feature fusion and optimized Deep Neuro Fuzzy Network," Advances in Engineering Software, vol. 175, p. 103333, 2022.

[23] S. J. Alsunaidi, R. T. Alraddadi, H. Aljamaan, "Twitter spam accounts detection using machine learning models," **2022 14th International Conference on Computational Intelligence and Communication Networks (CICN)**, vol. 15, pp. 525–531, 2022.

[24] S. Kaddoura, S. Henno, "Dataset of Arabic spam and ham tweets," Data in Brief, vol. 52, p. 109904, 2023.

[25] D. SivaKrishna, G. Srinivas, "StopSpamX: A multi-modal fusion approach for spam detection in social networking," MethodsX, vol. 12, p. 103227, 2025.

[26] K. P. Sharma, M. Gupta, A. Mishra, "Quantum Behaved Binary Gravitational Search Algorithm with Random Forest for Twitter Spammer Detection," Results in Engineering, vol. 20, p. 103993, 2025.

[27] A. Nekrasov, S. H. Teoh, S. Wu, "Visuals and attention to earnings news on Twitter," *SSRN Electronic Journal*, 2019.

[28] C. Boididou, S. Papadopoulos, Y. Kompatsiaris, J. Schifferes, N. Newman, "Verifying information with multimedia content on Twitter," Multimedia Tools and Applications, vol. 77, no. 12, pp. 15545–15571, 2017.

[29] Cumhurbaşkanlığı İletişim Başkanlığı, Sosyal Medya Kullanım Kılavuzu, Türkiye, 2020.

[30] Q. Ren, H. Cheng, H. Han, "Research on machine learning framework based on random forest algorithm," **AIP Conference Proceedings, vol.** 1864, no. 1, p. 020164, 2017.

[31] A. Gupte, S. Joshi, P. Gadgul, A. Kadam, "Comparative Study of Classification Algorithms used in Sentiment Analysis," *International Journal of Computer Science and Information Technologies, vol.* 5, no. 5, pp. 6261–6264, 2014.

Bulanık Hedef Programlama Yöntemi İle Ele Alınan Lisansüstü Tez Çalışmalarının Bibliyometrik Analizi

Araştırma Makalesi/Research Article

Elif AKDAŞ[1], Tamer EREN[2]*

[1,2]Endüstri Mühendisliği Anabilim Dalı, Kırıkkale Üniversitesi, Kırıkkale, Türkiye
meakdas9@gmail.com, tamereren@gmail.com
(Geliş/Received:05.06.2024; Kabul/Accepted:01.04.2025)
DOI: 10.17671/gazibtd.1495999

Özet— Bu makalede, Bulanık Hedef Programlama (BHP) yöntemi ile yapılan lisansüstü tez çalışmalarının bibliyometrik analizi sunulmaktadır. BHP, belirsizlik içeren karar verme problemlerini ele almak için kullanılan bir matematiksel modelleme yaklaşımıdır. Belirsizliklerin ve kesin olmayan bilgilerin bulunduğu durumlarda, karar verme sürecini etkili bir şekilde yönetmeyi sağlar. Bu yaklaşımın lisansüstü tez çalışmalarında nasıl kullanıldığını ve bu alandaki araştırma trendlerini belirlemek için bibliyometrik analiz yönteminin nasıl uygulandığı incelenmektedir. Bibliyometrik analiz için, Yükseköğretim Kurulu (YÖK) Tez Merkezi'nde 1996-2024 yılları arasında BHP ile ele alınan lisansüstü tez çalışma dokümanları taranmıştır. Elde edilen 51 adet tez çalışmasının yazarı, yayın tarihi, danışman adı, ele alındığı konu ve anahtar kelimeleri gibi temel özellikleri toplanmıştır. Bu çalışmanın bulguları, BHP'nin özellikle endüstri mühendisliği, işletme ve istatistik gibi ana bilim dallarında yoğun olarak kullanıldığını göstermektedir. Ayrıca, BHP'nin özellikle karar destek sistemleri, çok kriterli karar verme ve optimizasyon teknikleri gibi çeşitli alanlarla entegrasyonu göze çarpmaktadır.

Anahtar Kelimeler— bulanık hedef programlama, bibliyometrik analiz, lisansüstü tez çalışmaları

Fuzzy Goal Programming Approach For Bibliometric Analysis Of Postgraduate Thesis Studies

Abstract— This article presents a bibliometric analysis of graduate thesis studies conducted using the Fuzzy Goal Programming (FGP) method. FGP is a mathematical modeling approach used to address decision-making problems involving uncertainty. It effectively manages the decision-making process in situations where uncertainties and imprecise information exist. The study examines how this approach is utilized in graduate thesis studies and applies bibliometric analysis to determine research trends in this field. For bibliometric analysis, graduate thesis documents addressed with FGP in the Higher Education Council (YÖK) Thesis Center between 1996 and 2024 were scanned. Basic characteristics such as the author, publication date, advisor name, topic addressed, and keywords of the 51 obtained thesis studies were collected. The findings of this study show that BHP is used extensively especially in industrial engineering, business administration and statistics. In addition, the integration of BHP with various fields such as decision support systems, multi-criteria decision making and optimization techniques is particularly noticeable.

Keywords— fuzzy goal programming, bibliometric analysis, postgraduate thesis studies

1. GİRİŞ (INTRODUCTION)

Çok amaçlı doğrusal programlama modellerinin çözümünde çeşitli yöntemler kullanılmaktadır. Bunlardan birisi de hedef programlama yöntemidir. Söz konusu yöntemde, matematiksel modelde yer alan kısıtlarla hedefler belirlenmektedir [1]. Hedef programlamanın amacı, sapma değişkenlerinden mümkün olduğu kadar az sapma olmasını sağlamaktır [2]. Ancak hedeflerin tam olarak belirlenemediği veya kesin olmadığı koşullarda, problemin karar vericisi hedeflerin belirsizliğini ifade etmek için bulanık kümeler ve bulanık kurallar kullanmaktadır. Bu sayede, karar verme sürecindeki belirsizlikler ve kesin olmayan bilgiler daha etkin bir şekilde ele alınabilir. BHP, geleneksel hedef programlama yönteminden farklı olarak gerçek hayat problemlerindeki belirsizliklerle başa çıkmak için kullanılmakta ve karar verme sürecine esneklik sağlayan bir matematiksel modelleme yaklaşımıdır [3]. Hedef programlama kesin hedeflere odaklanırken, BHP belirsizlik içeren koşullar için daha uygun bir yaklaşım sunmaktadır [2].

Literatürde, BHP konusu ile ele alınan lisansüstü tez çalışmalarının hangi yılda, hangi yazar ve danışman tarafından, hangi üniversite ve üniversite türünde, hangi ana bilim dalında ve hangi problem türlerinde incelendiğine dair yapılmış herhangi bir çalışmaya rastlanmamıştır. Bu nedenle bu çalışma, BHP yöntemi ile gerçekleştirilen lisansüstü tez çalışmalarının bibliyometrik analizini sunmaktadır. Araştırma verileri, YÖK Tez üzerinden elde edilmiştir. Tarama terimlerinin "bulanık hedef programlama" ve "fuzzy goal programming" olduğu ve aranacak alanın "özet" seçildiği lisansüstü tez çalışma dokümanlarının taraması yapılmıştır.

BHP, belirsizlik içeren karar verme problemlerine etkili bir matematiksel modelleme yaklaşımı sunmaktadır. Yapılan analiz, BHP'nin lisansüstü düzeydeki araştırmalarda kullanımını ve araştırma eğilimlerini göstermektedir. Çalışmanın bulguları, BHP yönteminin gelecekteki araştırmalar için önemli bir temel oluşturduğunu ve karar verme süreçlerini iyileştirmek isteyen akademisyenler ve profesyoneller için değerli bilgiler sunduğunu göstermektedir.

2. BİBLİYOMETRİK ANALİZ (BIBIOMETRIC ANALYSIS)

Bibliyometri terimi, kitap anlamında olan 'biblio' kelimesi ile ölçü birimi anlamında olan 'metrics' kelimesi bir araya getirilerek türetilmiştir. Bu terim, matematiksel ve istatistiksel yöntemlerin bilimsel kitaplar ve diğer iletişim araçlarına uygulanması olarak tanımlanmıştır. Bibliyometri, çeşitli veri tabanlarından elde edilen verilerin istatistiksel ve matematiksel yöntemlerle analiz edilmesine dayanmaktadır. Başka bir deyişle, bibliyometri; kitaplar, makaleler ve diğer yayınların istatistiksel bir analizidir (Oxford Dictionary, 2017).

Bibliyometrik analiz, belirli bir alanda belirli bir dönemde ve belirli bir coğrafi bölgede kişiler veya kurumlar tarafından üretilmiş yayınların ve bu yayınlar arasındaki ilişkilerin sayısal olarak incelenmesidir. Bu analiz, bilimsel yayınlar üzerinde istatistiksel ve matematiksel yöntemler kullanarak bir bilim dalının gelişimini, içeriğini, yayın eğilimlerini ve etkileşimlerini değerlendirerek mevcut durumunu belirlemeyi ve gelecekteki eğilimleri tahmin etmeyi amaçlamaktadır [4]. Güçlü bir yöntem ve veri seti sunan bibliyometrik analiz, belirli bir zaman dilimindeki makale sayısı, yazarlar ve ortak çalışmalar gibi faktörler üzerinden bir çalışma alanının genel yapısını ortaya çıkarabilir. Ayrıca, bir çalışmanın kendisinden sonra yapılan çalışmalar üzerindeki etkisini değerlendirmek için de kullanılabilir [5].

Bibliyometrik analizin temel kullanım alanlarından biri de, araştırma eğilimlerini belirlemektir. Belirli bir araştırma alanındaki eğilimleri ve konuların popülerliğini izlemek için kullanılmaktadır. Bu sayede, gelecekteki araştırma yönergelerini belirlemede yardımcı olabilir.

3. BULANIK HEDEF PROGRAMLAMA (FUZZY GOAL PROGRAMMING)

BHP, belirsizlik içeren karar verme problemlerini ele almak için kullanılan bir matematiksel modelleme yaklaşımıdır. Bu yöntem, karar verme sürecini etkili bir şekilde yönetmeyi sağlamanın yanı sıra belirsizliklerin ve kesin olmayan bilgilerin bulunduğu durumlarda kullanılır.

Karar verme sürecinde kullanılan hedefler ve kısıtların belirsizliği ifade etmek için bulanık kümeler ve bulanık mantık kullanılır. Bu sayede, gerçek hayattaki karmaşık problemleri modellemek ve çözmek mümkün olmaktadır [6].

BHP, [7] tarafından geliştirilen bulanık küme teorisine ve [8] tarafından sunulan üyelik fonksiyonları kavramına dayanmaktadır. Karar vericinin yalnızca belirsiz ve kesin olmayan hedef değerlerini kullanabildiği durumların çözümü için geliştirilmiştir [9].

Hedef programlama yönteminde bulanık küme teorisinin ilk kullanımı [10] numaralı çalışmada ele alınmıştır. Sonrasında, eşit öncelikli belirsiz hedefler veya kısıtlar içeren tek bir hedef programlama probleminin nasıl çözülebileceğini [11] göstermiştir. Genelleştirilmiş bir BHP modeli aşağıdaki gibi ifade edilmektedir. Bu hedeflerde, hedeflerin sol taraf parametreleri birer sabitken, sadece hedeflerin istenen seviyeleri bulanık sayı olarak ele alınmıştır [12].

Bulanık Hedefler:

$$(AX)_i \leq \sim b_i \; i = 1, 2, \dots, i_0$$

$$(AX)_i \geq \sim b_i \; i = i_0 + 1, \dots, j_0$$

$$(AX)_i = \sim b_i \; i = j_0 + 1, 2, \ldots, k$$

Bulanık Olmayan Kısıtlar:

$$(AX)_l (=, \leq, \geq) b_l \;\; l = 1, 2, \ldots, p$$

$$x_j \geq 0 \; i = 1, 2, \ldots, m$$

Burada '~' i-inci hedefin istenilen seviyeleri olan b_i'nin bulanık sayıdan oluştuğunu göstermektedir [13].

BHP ile yapılan çalışmalara örnek verecek olursak;

[14] hastanelerde yatan diyabet, mide gibi rahatsızlıkları olan hastalara bir aylık menü planlaması hazırlarken değer aralıkları kesin olmadığından dolayı BHP yönteminden yararlanmışlardır. [15] tedarikçi seçimi karar sürecindeki belirsizlikler ve çoklu hedefler dikkate alınarak, BHP ile çözüm önerisinde bulunmuşlardır. [16] BHP yönteminden yararlanarak, sigorta şirketlerinde firmaya ait değişkenlerin dönem kârının ya da zararının üzerindeki etkilerini incelemeyi amaçlamışlardır.

4. LİTERATÜR ANALİZİ (LITERATURE ANALYSIS)

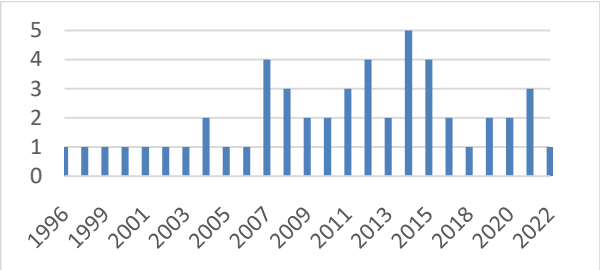
Bu çalışma, Yükseköğretim Kurulu (YÖK) Tez Merkezi'nde yayınlanan BHP ile ele alınan lisansüstü tez çalışmalarının doküman analizi yöntemiyle incelendiği nitel bir araştırmadır.

Veri toplama sürecinin ilk aşamasında; YÖK Tez Merkezi'nin internet sayfasından çevrimiçi tarama yapılmıştır. Tarama terimi alanına "fuzzy goal programming", "bulanık hedef programlama" ve "bulanık amaç programlama" kelimeleri yazılmıştır ve aranacak alan "özet" seçilerek arama yapılmıştır.

Arama işlemi sonucunda 1996-2024 yılları arasında taranan 51 adet lisansüstü çalışmalara ulaşılmıştır. Taranan çalışmalar; yazar, yıl, lisansüstü tez çalışmalarının türü, etkin danışmanlar, etkin üniversiteler ve türleri, etkin enstitüler ve ana bilim dalı, ele alınan konular, entegre edilen yöntemler ve anahtar kelimelerin analizi ile kategorize edilmiştir. 2023 ve 2024 yıllarında yapılmış herhangi biz çalışma bulunmamaktadır.

4.1. Yıllara Göre Toplam Çalışma Sayısının Dağılımı (Distribution of Total Number of Studies by Year)

Yapılan lisansüstü tez çalışmalarının yıllara toplam çalışma sayısı dağılımı Şekil 1'de verilmiştir.

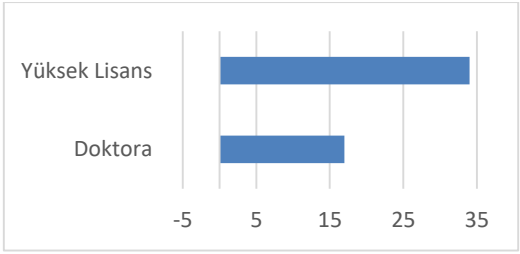


Şekil 1. Yıllara Göre Dağılım
(Distribution by Year)

Şekil 1'de görüldüğü üzere, BHP yöntemiyle yapılan lisansüstü çalışmaların sayısında yıllar içinde sürekli bir artış gözlemlenmektedir. BHP yöntemine olan ilgi zamanla artmış, özellikle 2014 yılından itibaren belirgin bir yükseliş kaydedilmiştir. Bu durum, söz konusu alanın daha fazla araştırmacı tarafından tercih edildiğini ve gelecekte de araştırma ilgisinin devam edeceğini göstermektedir. Ancak, 2023 ve 2024 yıllarında bir duraklama döneminin yaşandığı tespit edilmiştir. Bu duraklama, gelecekte yapılacak araştırmalarda alandaki yeniden canlanma eğilimlerini ortaya koyabilir. Başka bir deyişle, araştırmaların zaman içindeki evrimini ve mevcut duraklama döneminin gelecekteki araştırmalara yön verebilecek potansiyelini ortaya koymaktadır.

4.2. Lisansüstü Tez Çalışmalarının Türü (Type of Graduate Thesis Studies)

Yapılan çalışmaların tez türüne göre dağılımı Şekil 2'de verilmiştir.

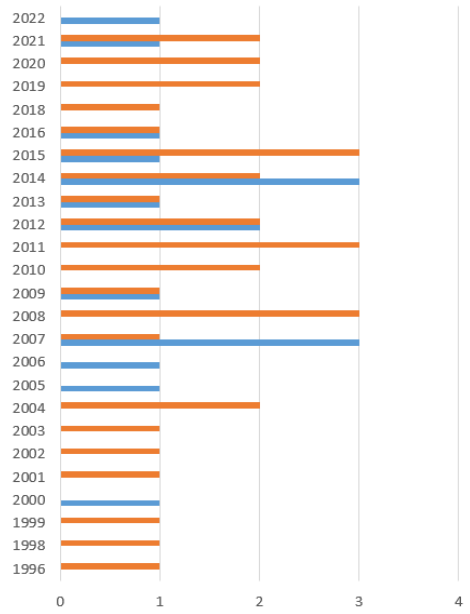


Şekil 2. Çalışma Türüne Göre Dağılım
(Distribution by Type of Work)

Şekil 2'ye yönelik analizde, incelenen 51 lisansüstü çalışmadan 17'sinin doktora tezi ve 34'ünün ise yüksek lisans tezi olduğu belirlenmiştir.

4.3. Yıllara Göre Çalışma Türü Sayılarının Eğilimi (Trends in the Number of Types of Work by Year)

Lisansüstü tez çalışmalarının türlerine ve gerçekleştirildikleri yıllara göre dağılımı, Şekil 3'te sunulmaktadır.



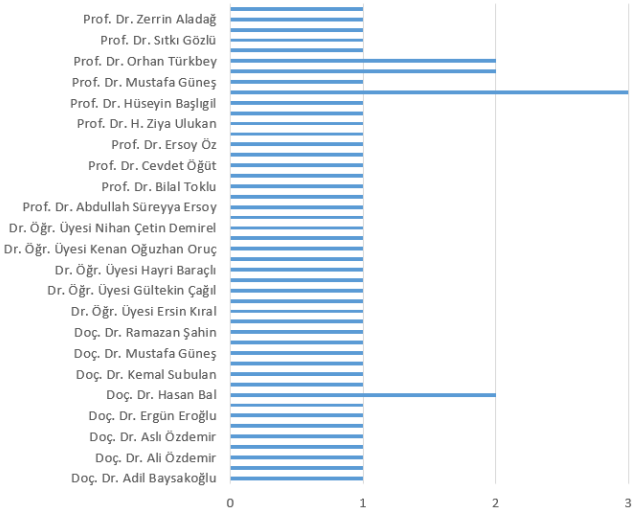
(Turuncu: Yüksek Lisans, Mavi: Doktora)

Şekil 3. Çalışma Türü Sayılarının Yıllara Göre Dağılımı
(Distribution of the Number of Study Types by Year)

Şekil 3 incelendiğinde, Yüksek Lisans tez çalışmalarında neredeyse her yıl Bibliyometrik Hızlandırma (BHP) yöntemine başvurulduğu gözlemlenmektedir. Bununla birlikte, Yüksek Lisans tezlerine kıyasla, Doktora tezlerinde BHP yönteminin kullanım sıklığı daha düşük seviyelerde kalmaktadır.

4.4. Alanındaki Etkin Danışmanlar (Effective Consultants in the Field)

Lisanüstü öğrencileri ile BHP yöntemi üzerine çalışan danışmanların toplam çalışma sayılarına göre dağılımı Şekil 4'te görüldüğü gibidir.

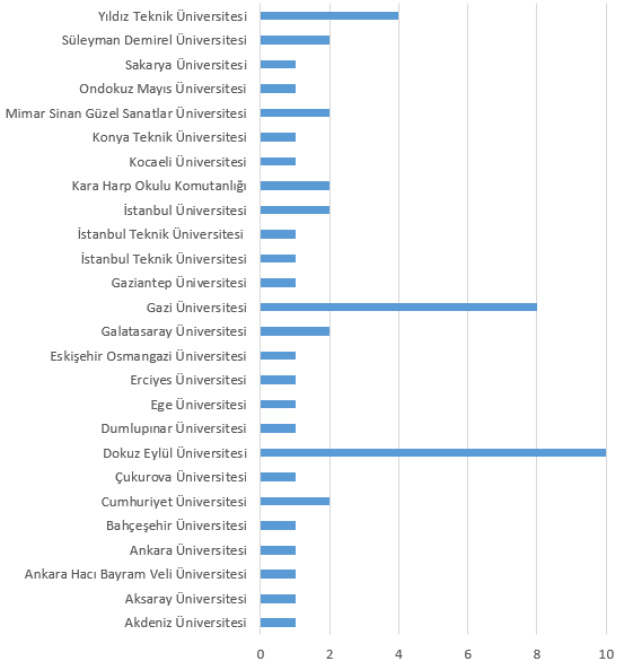


Şekil 4. Etkin Danışmana Göre Dağılım
(Distribution by Active Consultant)

Şekil 4'e bakıldığında, söz konusu yöntem ile en fazla çalışma yapan danışmanın 4 çalışma ile Prof. Dr. İrem Özkarahan olduğu görülmektedir. Diğer danışmanların ise BHP ile 1 veya 2 çalışma gerçekleştirdiği görülmektedir.

4.5. Alanındaki Etkin Üniversiteler ve Türleri (Active Universities and Types in the Field)

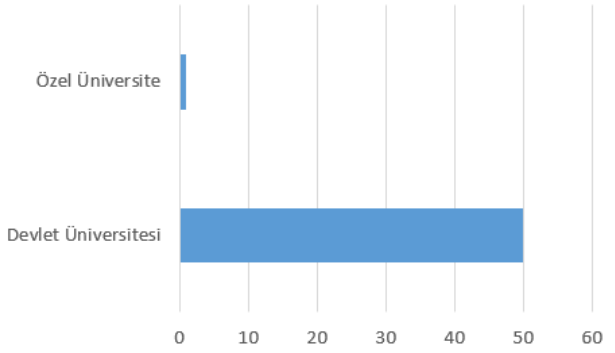
BHP yöntemini ele alarak yapılan tez çalışmalarının üniversitelere göre dağılımı Şekil 4'te görüldüğü gibidir. Bu dağılım, her bir üniversitenin BHP yöntemiyle yapılan tez çalışmalarına olan katkısını ve söz konusu yöntemin hangi üniversitelerde daha yoğun olarak kullanıldığını göstermektedir. Şekil 4, her üniversiteye ait tez çalışmalarının sayısını belirleyerek, BHP yönteminin akademik araştırmalar arasındaki yaygınlık derecesini ortaya koymaktadır. Ayrıca, bu dağılımda üniversitelerde görevli etkin danışmanların katkıları da dikkate alınarak, hangi danışmanların BHP yöntemiyle en fazla çalışmayı gerçekleş0tirdiği hakkında bilgiler sunulmaktadır.



Şekil 5. Üniversitelere Göre Dağılım
(Distribution by Universities)

Şekil 5'e göre, BHP yönteminin 10 adet çalışma ile en fazla Dokuz Eylül Üniversitesi'nde ve daha sonra 8 adet çalışma ile de Gazi Üniversitesi'nde ele alındığı görülmektedir.

BHP yöntemini ele alarak yapılan tez çalışmalarının üniversite türlerine göre dağılımı, Şekil 6'da açıkça sunulmaktadır.

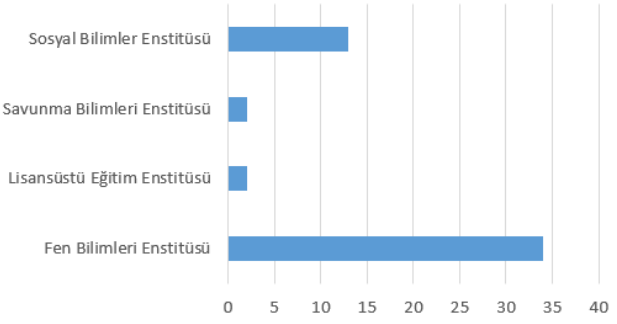


Şekil 6. Üniversite Türüne Göre Dağılım
(Distribution by Type of University)

Şekil 6'ya göre bu dağılımda, devlet üniversitelerinde toplamda 50 adet, özel üniversitelerde ise yalnızca 1 adet BHP ile ilgili çalışma gerçekleştirildiği görülmektedir. Şekil 5, BHP yönteminin daha çok devlet üniversitelerinde tercih edildiğini ortaya koymakta ve özel üniversitelerde bu yöntemin kullanımının sınırlı kaldığını göstermektedir. Bu durum, devlet üniversitelerinin araştırma kapasitesinin ve kaynaklarının bu tür yöntemlerin uygulanmasında daha belirleyici bir rol oynadığını, buna karşın özel üniversitelerde BHP yöntemine yönelik araştırmaların daha az olduğunu göstermektedir.

4.6. Etkin Enstitüler ve Ana Bilim Dalı (Active Institutes and Departments)

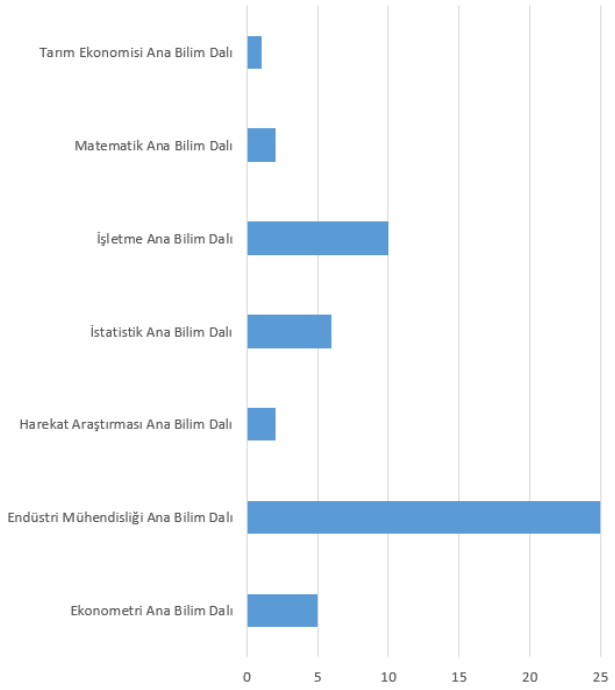
BHP yöntemi ile yapılan lisansüstü çalışmaların enstitülere göre dağılımı, Şekil 7'de ayrıntılı bir şekilde sunulmaktadır.



Şekil 7. Enstitülere Göre Dağılım
(Distribution by Institutes)

Şekil 7'ye göre, BHP yöntemi ile gerçekleştirilen lisansüstü çalışmaların enstitülere göre dağılımı incelendiğinde, en fazla çalışmanın Fen Bilimleri Enstitüsü'nde yapıldığı görülmektedir. Bu enstitüde toplamda 34 adet BHP yöntemiyle hazırlanan tez çalışması bulunmaktadır. Sosyal Bilimler Enstitüsü ise 13 adet çalışma ile ikinci sırada yer almaktadır. Bu dağılım, BHP yönteminin daha çok Fen Bilimleri alanında tercih edildiğini, ancak Sosyal Bilimler alanında da belirli bir düzeyde ilgi gösterildiğini ortaya koymaktadır. Bu durum, BHP yönteminin farklı akademik disiplinlerde nasıl kullanıldığına dair önemli bir gösterge sunmaktadır.

BHP yöntemi ile hangi ana bilim dalları tarafından ne kadar sayıda lisansüstü çalışma yapıldığının sayısal dağılımı, Şekil 8'de detaylı bir şekilde sunulmuştur. Bu dağılım, BHP yönteminin farklı ana bilim dallarında kullanım oranlarını ve bu ana bilim dallarında yapılan tez çalışmalarının sayısını açıkça göstermektedir.

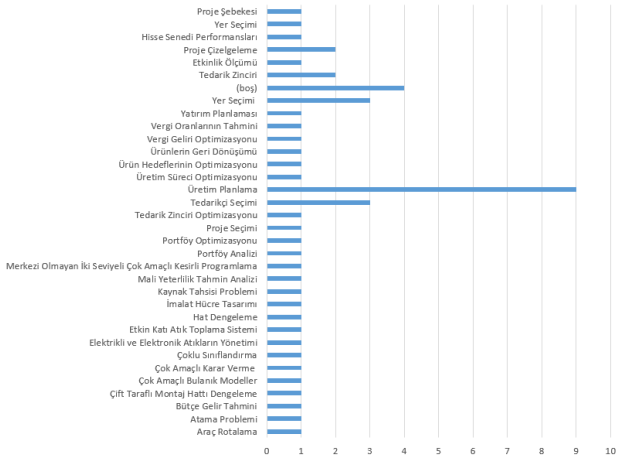


Şekil 8. Ana Bilim Dalına Göre Dağılım
(Distribution According to Main Department)

Şekil 8'deki verilere göre, lisansüstü tez çalışmalarında BHP yöntemini odaklayarak en fazla çalışma yapan ana bilim dalı, 25 çalışma ile Endüstri Mühendisliği Ana Bilim Dalı olarak belirlenmiştir. Daha sonra, İşletme Ana Bilim Dalı 10 çalışma ile ikinci sırada yer almaktadır. En düşük çalışma sayısı ise 1 adet ile Tarım Ekonomisi Ana Bilim Dalı'nda gerçekleştirilmiştir.

4.7. Ele Alınan Konular (Topics Addressed)

BHP yöntemi kullanılarak yapılan tez çalışmalarında ele alınan konular Şekil 9'da verilmiştir. Ancak bu çalışmalar arasından 4 tanesinin konusu belirlenemediği için bu 4 çalışma analiz dışında bırakılmıştır.



Şekil 9. Konulara Göre Dağılım
(Distribution by Subjects)

Şekil 9 incelendiğinde, BHP yöntemiyle ele alınan tez çalışmalarında "Üretim Planlama" konusunun 9 adet çalışma ile en fazla ele alındığı gözlemlenmiştir. Ardından, "Tedarik Seçimi" ve "Yer Seçimi" konularında 3 çalışma gerçekleştirilmiştir. Diğer konular ise genellikle 1 veya 2 çalışma ile temsil edilmiştir.

4.8. Anahtar Kelime Analizi (Keyword Analysis)

BHP yöntemi ile gerçekleştirilen lisansüstü tez çalışmalarında yer alan anahtar kelimelerin kullanım sıklıklarına dayalı olarak bir kelime bulutu oluşturulmuştur. Bu kelime bulutu, Şekil 10'da verilmiştir. Ancak bu çalışmalar arasından 7 tanesinin anahtar kelimelerine ulaşılamadığından dolayı bu 7 çalışma hariç bırakılmıştır.

Şekil 10. Anahtar Kelime Analizi
(Keyword Analysis)

Şekil 10'da görüldüğü gibi, en fazla kullanılan anahtar kelimeler arasında "Bulanık Hedef Programlama" ve "Hedef Programlama" öne çıkmaktadır.

Yapılan bibliyometrik analiz çalışmasının sonuçlarını değerlendirecek olursak; BHP yöntemine olan ilginin artışı ve bu alandaki önemli araştırma eğilimleri vurgulanmıştır. Bu tür bir çalışma, literatüre önemli katkılar sağlayabilir. Çalışmanın sağladığı katkılar:

- ✓ Zaman içindeki eğilimlerinin tespiti
- ✓ Alan içi farklılaşma ve konuların derinlemesine incelenmesi
- ✓ Araştırma alanlarındaki potansiyel boşlukların belirlenmesi
- ✓ Literatürdeki bilgi akışının sağlanması

5. SONUÇ VE ÖNERİLER (CONCLUSIONS AND RECOMMENDATIONS)

Bu araştırma, BHP yöntemi kullanılarak gerçekleştirilen 51 lisansüstü tez çalışmasının bibliyometrik analizini sunmaktadır. 1996 ve 2024 yılları arasında neredeyse her yıl birçok ana bilim dalının bu yöntem ile farklı konularda çalışmalar gerçekleştirdiği görülmektedir. Yapılan analiz, BHP'nin zaman içerisinde lisansüstü düzeydeki araştırmalarda nasıl yaygınlaştığını ve farklı disiplinlerde nasıl uygulandığını ortaya koymaktadır. Bu durum, BHP yönteminin karar verme süreçlerinde belirsizlikleri yönetme yeteneğinin, çok çeşitli alanlarda kabul gördüğünü ve gelecekte de genişlemeye devam edebileceğini işaret etmektedir.

İncelenen tez çalışmalarının analizi; BHP'nin endüstri mühendisliği, işletme, istatistik ve diğer ana bilim dallarında sıklıkla tercih edildiğini göstermektedir. Bu değerlendirme, BHP'nin lisansüstü düzeydeki araştırmaların geniş bir uygulama alanına sahip olduğunu ve bu yöntemin farklı disiplinlerde karşılaşılan belirsizlikleri etkili bir şekilde yönetme potansiyelini yansıtmaktadır. Ayrıca, BHP'nin özellikle karar verme süreçlerinde önemli bir araç olarak kabul edildiği ve gelecekte bu alandaki araştırmalara yön verme açısından belirleyici bir rol oynayabileceğini vurgulamaktadır. Endüstri mühendisliği ve işletme ana bilim dalında üretim planlama, yer seçimi, tedarikçi seçimi, çok amaçlı karar verme problemleri gibi birçok konuda kullanılan BHP'nin uygulama alanının geniş olmasına rağmen, son yıllarda bu yöntemin kullanımı sınırlı sayıda kalmıştır.

Yapılan bibliyometrik analiz, BHP'nin belirsizlik içeren karar verme problemlerine etkili bir matematiksel modelleme yaklaşımı sunduğunu ve karar verme süreçlerini yönetmede kullanılabileceğini desteklemektedir. Çalışma, BHP'nin lisansüstü tezlerde nasıl uygulandığını ve bu yöntemin akademik araştırmalarda daha fazla nasıl kullanılabileceği gösteren önemli bulgular sunmaktadır. Bu bulgular, gelecekteki araştırmalar için sağlam bir temel oluşturarak, BHP'nin hangi alanlarda daha fazla araştırma yapılması gerektiğine dair ipuçları sağlamaktadır.

Özetleyecek olursak lisansüstü tez çalışmalarında ele alınan BHP yöntemi, son yıllarda popülerliğini korumakta olan ve devlet üniversitelerinde en çok ele alınan bir yöntem olmuştur. Fen Bilimleri Enstitüsü'nde ve özellikle de Endüstri Mühendisliği Ana Bilim Dalı'nda 25 çalışmada kullanılan bir yöntem olmuştur. Literatürde yer alan 9 adet üretim planlama problemlerinin ele alındığı lisansüstü tez çalışmalarında yine bu yöntemden yararlanıldığı görülmektedir.

Bibliyometrik analiz araştırmacılar, akademisyenler, yayıncılar ve karar vericiler için bu yöntemin etkisini, görünürlüğünü ve önemini anlamayı amaçlamıştır. Yapılan bibliyometrik analiz, ilgili alandaki bilimsel araştırma stratejilerini belirleme konusunda bir kılavuz olarak kullanılabilir.

KAYNAKLAR (REFERENCES)

[1] B. Demirel, A. Yelek, H. M. Alağaş, T. Eren, "Ankaray Güvenlik Personelinin Vardiya Çizelgeleme Probleminin Hedef Programlama Yöntemi İle Çözümü", *Demiryolu Mühendisliği*, (8), 1-17, 2018.

[2] C. H. Kağnıcıoğlu, "Hedef Programlama ve Bulanık Hedef Programlama Arasındaki İlişki", *Gazi Üniversitesi İktisadi ve İdari Bilimler Fakültesi Dergisi*, 7(2), 17-38, 2005.

[3] A. Lotfi, A. Dorra, B. Kaddour, K. Abdessamad, "Fuzzy goal programming to optimization the multi-objective problem", *Science Journal of Applied Mathematics and Statistics*, 2(1), 14-19, 2014.

[4] G. D. Şakar, A. G. Cerit, "Uluslararası Alan İndekslerinde Türkiye Pazarlama Yazını: Bibliyometrik Analizler ve Nitel Bir Araştırma", *Atatürk Üniversitesi İktisadi ve İdari Bilimler Dergisi*, 27(4), 274201337-274201362, 2013.

[5] U. Al, U. Sezen, İ. Soydal, "Hacettepe Üniversitesi bilimsel yayınlarının sosyal ağ analizi yöntemiyle değerlendirilmesi", *Hacettepe Üniversitesi Edebiyat Fakültesi Dergisi*, 29(1), 2012.

[6] İ. Ertuğrul, "Bulanık Hedef Programlama ve Bir Tekstil Firmasında Uygulama Örneği", *Eskişehir Osmangazi Üniversitesi Sosyal Bilimler Dergisi*, 6(2), 45-79, 2005.

[7] L.A. Zadeh, "The concept of a linguistic variable and its application to approximate reasoning—I", *Information Sciences*, 8(3), 199-249, 1975.

[8] H.J. Zimmermann, "Fuzzy programming and linear programming with several objective functions", *Fuzzy sets and systems*, 1(1), 45-55, 1978.

[9] C. Colapinto, R. Jayaraman, D. La Torre, "Goal Programming Models For Managerial Strategic Decision Making", *In Applied mathematical analysis: theory, methods, and applications*, 487-507, 2020.

[10] R. Narasimhan, "Goal Programming in a Fuzzy Enviroment", *Decision Sciences*, 11, 325-336, 1980.

[11] E.L. Hannan, "Linear Programming with Multiple Fuzzy Goals", *Fuzzy Sets and Systems*, 6(3), 235-248, 1981.

[12] R. N. Tiwari, S. Dhamar, J. R. Rao, "Priority Structure in Fuzzy Goal Programming", *Fuzzy Sets and Systems*, 19, 251-259, 1986.

[13] M. M. Özkan, Bulanık Hedef Programlama, Ekin Kitabevi, Bursa, 2003.

[14] T. Eren, S. Ö. Kaçmaz, N. Şengül, "Hastanelerde Özel Hastalar İçin Bulanık Hedef Programlama İle Menü Planlaması", *Beykent Üniversitesi Fen ve Mühendislik Bilimleri Dergisi*, 11(1), 1-37, 2018.

[15] E. Aydın, Çağıl, G., "Bulanık Ahp ve Bulanık Hedef Yaklaşımı ile Hammadde Tedarikçisi Seçimi", *Itobiad: Journal of the Human & Social Science Researches/İnsan ve Toplum Bilimleri Araştırmaları Dergisi*, 9(5), 3568-3597, 2020.

[16] Y. Akgül, F. Çamlıbel, S. Çamlıbel, "Hayat dışı sigorta sektöründe kârı etkileyen firma içi faktörlerin incelenmesi: Bulanık hedef programlama örneği", *Ekonomi Politika ve Finans Araştırmaları Dergisi*, 6(2), 332-355, 2021.