



# JOURNAL OF INFORMATION TECHNOLOGIES

## BİLİŞİM TEKNOLOJİLERİ DERGİSİ

Volume / Cilt **17** Number / Sayı **4** Year / Yıl **2024** Month / Ay **October / Ekim**





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): 17

Sayı (Issue): 4

Ekim (October) 2024

Sahibi (Owner)  
**Dr. Uğur ÜNAL**  
Rektör (Rector)

**Genel Yayın Yönetmeni & Baş Editör**  
(General Publication Director & Editor in Chief)  
**Dr. Ashhan TÜFEKÇİ**  
Bilişim Enstitüsü Müdürü  
Director of Institute of Informatics

**Yardımcı Editör**  
(Associate Editor)  
**Dr. Öner BARUT**  
Bilişim Enstitüsü Müdür Yardımcısı  
Vice Director of Institute of Informatics

**Yardımcı Editör**  
(Associate Editor)  
**Dr. Nadircan DAĞLI**  
Bilişim Enstitüsü Müdür Yardımcısı  
Vice Director of Institute of Informatics

**Editörler (Editors)**

❖ <b>Dr. Serdar KULA</b>	Gazi Üniversitesi (Gazi University)
❖ <b>Dr. Hüseyin POLAT</b>	Gazi Üniversitesi (Gazi University)
❖ <b>Dr. Resul DAŞ</b>	Fırat Üniversitesi (Fırat University)
❖ <b>Dr. Muzaffer KAPANOĞLU</b>	Eskişehir Osmangazi Üniversitesi (Eskişehir Osmangazi University)
❖ <b>Dr. Mehmet ŞİMŞEK</b>	Milli Savunma Üniversitesi (National Defence University)
❖ <b>Dr. Mehmet Sıraç ÖZERDEM</b>	Dicle Üniversitesi (Dicle University)
❖ <b>Dr. Mustafa Mahir ÜLGÜ</b>	Sağlık Bakanlığı
❖ <b>Dr. Murat YILMAZ</b>	Gazi Üniversitesi (Gazi University)
❖ <b>Dr. Oktay YILDIZ</b>	Gazi Üniversitesi (Gazi University)
❖ <b>Dr. Olgun DEĞİRMENCİ</b>	TOBB ETÜ (TOBB Economics and Technology University)
❖ <b>Dr. Recep BENZER</b>	Gazi Üniversitesi (Gazi University)
❖ <b>Dr. Ö. Tolga PUSATLI</b>	Çankaya Üniversitesi (Çankaya University)
❖ <b>Dr. Cihangir TEZCAN</b>	Orta Doğu Teknik Üniversitesi (Middle East Technical University)
❖ <b>Dr. Mehmet SEVRİ</b>	Recep Tayyip Erdoğan Üniversitesi (Recep Tayyip Erdoğan University)
❖ <b>Dr. Muhammed Ali KOŞAN</b>	Kahramanmaraş İstiklal Üniversitesi (Kahramanmaraş Istiklal University)
❖ <b>Dr. Levent ÇETİNKAYA</b>	Çanakkale Onsekiz Mart Üniversitesi (Çanakkale Onsekiz Mart University)
❖ <b>Dr. Sinan TOKLU</b>	Gazi Üniversitesi (Gazi University)

**Yayın Danışma Kurulu (Editorial Advisory Board)**

<b>Dr. Ahmet COŞAR</b> <i>Turkish Aeronautical Association University, Turkey</i>	<b>Dr. Aslanbek NAZİEV</b> <i>Ryazan State University, Russia</i>	<b>Dr. Bogdan PATRUT</b> <i>Alexandru Ioan Cuza University of Iasi, Romania</i>
<b>Dr. Deepak GUPTA</b> <i>Maharaja Agrasen Institute of Technology, India</i>	<b>Dr. Jafar A. ALZUBİ</b> <i>Al-Balqa Applied University, Jordan</i>	<b>Dr. Jolanta SABAITYTĖ</b> <i>Vilnius Gediminas Technical University, Lithuania</i>
<b>Dr. Ilya LEVİN</b> <i>Tel Aviv University, Israel</i>	<b>Dr. Pınar KARAGÖZ</b> <i>Middle East Technical University, Turkey</i>	<b>Dr. Ufuk ÇAĞLAYAN</b> <i>Yaşar University, Turkey</i>
<b>Dr. Veysi İŞLER</b> <i>Hasan Kalyoncu University, Turkey</i>	<b>Dr. Victor Hugo Costa DE ALBUQUERQUE</b> <i>Universidade de Fortaleza, Brazil</i>	<b>Dr. Vijender Kumar SOLANKİ</b> <i>CMR Institute of Technology, India</i>
<b>Dr. Ebrahim KHOSRAVI</b> <i>Clayton State University, United States</i>		

**Dil Editörü**

*(Language Editor)*

**Dr. Çağla Gizem AKKAŞ**

*Bilişim Enstitüsü  
Institute of Informatics*

**Teknik Sorumlu**

*(Technical Assistant)*

**Dr. Candan TÜMER**

*Bilişim Enstitüsü  
Institute of Informatics*

**Teknik Sorumlu**

*(Technical Assistant)*

**Yasemin İÇTÜZER**

*Bilişim Enstitüsü  
Institute of Informatics*

**Sekreterlik**

*(Secretary)*

**Bilişim Teknolojileri Dergisi**

*Bilişim Enstitüsü  
Institute of Informatics*

**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).

# Industrial Metaverse as a New Component of Digital Transformation: A Bibliometric Analysis

*Araştırma Makalesi/Research Article*

 İsmail YOŞUMAZ<sup>1</sup>

<sup>1</sup>Kütahya Dumlupınar Üniversitesi Kütahya Sosyal Bilimler MYO İşletme Yönetimi Bölümü Merkez / Kütahya / Türkiye

[ismaily@dpu.edu.tr](mailto:ismaily@dpu.edu.tr)

(Geliş/Received:20.05.2024; Kabul/Accepted:18.07.2024)

DOI: 10.17671/gazibtd.1487061

**Abstract**— The digital transformation process in businesses operates within the data and knowledge cycle. Industrial metaverse technology has recently emerged as a new catalyst within this transformation. Industrial metaverse technology aims to develop a virtual space for living and working that closely resembles the physical environment, facilitated through collaboration with multiple technologies. In this context, this study aims to identify the technologies or processes that industrial metaverse technology interacts with and to provide a framework for researchers and businesses engaging in this field. A bibliometric analysis was conducted on studies using the keyword “industrial metaverse” in the Web of Science and Scopus databases to accomplish this aim. A total of 44 studies were identified from the Scopus database and 16 from the Web of Science database, amounting to 60 studies. A complete census was performed based on these 60 studies, which constituted the research sample. Data were manually collected, and a keyword analysis was conducted using the VOSviewer application. The study's findings suggest that industrial metaverse technology is connected to processes such as smart production, employee-technology collaboration, employee training, and environmental sustainability. Additionally, technologies such as digital twins, blockchain, non-fungible tokens (NFTs), virtual reality, augmented reality, the industrial internet of things, and artificial intelligence were identified as forming the infrastructure of industrial metaverse technology.

**Keywords**— digital transformation, industrial metaverse, knowledge management, technology and innovation management, Industry 5.0

## Dijital Dönüşümün Yeni Bileşeni Olarak Endüstriyel Metaverse: Bibliyometrik Analiz

**Özet**— İşletmelerde dijital dönüşüm süreci veri ve bilgi döngüsü içerisinde işleyen bir süreçtir. Son zamanlarda dijital dönüşüm sürecine Endüstriyel metaverse teknolojisi yeni bir katalizör olarak dahil olmuştur. Endüstriyel metaverse teknolojisi birçok teknoloji ile iş birliği içerisinde, sanal ortamda fiziksel ortama benzer yeni bir sanal yaşam ve çalışma alanı oluşturmayı hedefleyen teknolojidir. Bu kapsamda, bu çalışmanın amacı, Endüstriyel metaverse teknolojisinin hangi teknoloji veya süreçlerle iş birliği içerisinde olduğunu belirleyebilmek ve bu sayede bu alanda çalışma yapacak araştırmacılara ve işletmelere bir vizyon sunabilmektir. Bu amaca ulaşabilmek için Web of Science ve Scopus veritabanlarında endüstriyel metaverse anahtar kelimesi kapsamında yapılan çalışmaların bibliyometrik analizi yapılmıştır. Scopus veritabanından 44, WoS veritabanından 16 ve toplamda 60 çalışmaya ulaşılmıştır. Araştırmanın evrenini oluşturan bu 60 çalışma kapsamında tam sayım yapılmıştır. Veriler manuel derlenmiş ve çalışmaların anahtar kelime analizleri VOSviewer uygulaması ile yapılmıştır. Çalışma sonucunda Endüstriyel metaverse teknolojisinin akıllı üretim, çalışan ve teknoloji iş birliği, çalışan eğitimi, çevresel sürdürülebilirlik gibi süreçlerle ilişkili olduğu sonucuna varılmıştır. Ayrıca bu ilişkinin sağlanabilmesi için dijital ikiz, blok zinciri, Nun-fungible token (NFT), sanal gerçeklik, arttırılmış gerçeklik, endüstriyel nesnelerin internet, yapay zekâ gibi teknolojilerin endüstriyel metaverse teknolojisinin altyapısını oluşturduğu görülmüştür.

**Anahtar Kelimeler**— dijital dönüşüm, endüstriyel metaverse, bilgi yönetimi, teknoloji ve yenilik yönetimi, Endüstri 5.0

## 1. INTRODUCTION

With the advent of Industry 4.0, businesses have begun automating all their processes to optimise data and knowledge through technological contributions, thereby initiating efforts within the scope of digital transformation. Before the Industry 4.0 process, digitalisation efforts were generally carried out in the context of injecting technology and automation into businesses. However, with the Industry 4.0 process, the data and knowledge needs of technologies such as big data, internet of things (IoT), cloud computing, artificial intelligence (AI), metaverse, blockchain, and NFTs have brought digital transformation efforts within the purview of the data and knowledge cycle [1].

Despite these advancements, the role of employees in the technology-supported digital transformation within the data and knowledge cycle during Industry 4.0 was not fully elucidated. The emergence of Industry 5.0, building on the technological foundations of Industry 4.0 and incorporating advanced technologies like productive artificial intelligence and industrial metaverse, has clarified the status of employees in the digital transformation process. Employees are now recognised as valuable assets in digital transformation. Alongside employees, environmental sustainability and business resilience against cyber security breaches, ecological disasters, and supply chain disruptions have become key components of Industry 5.0. Thus, the essential elements of a digital transformation suitable for Industry 5.0 are technology, employees, environmental sustainability, and business resilience.

A literature review reveals that industrial metaverse technology is interconnected with all components of the digital transformation process appropriate for Industry 5.0. Specifically, Kumar et al. [5] explore the relationship between the industrial metaverse and organisational agility and business performance, while Kshetri [6] examines the economic returns of the industrial metaverse. Additionally, Ren et al. [7], Xiang et al. [8], and Yang, Wang, and Zhao [9] investigate the link between industrial metaverse and smart manufacturing. Kshetri and Dwivedi [10] study the relationship between the industrial metaverse and environmental sustainability, and Huang et al. [11] and Zheng et al. [12] address the challenges and opportunities the industrial metaverse presents to businesses. Gattullo et al. analyse the relationship between the industrial metaverse and knowledge sharing, while Opperman et al. [13] focus on the relationship between the industrial metaverse and remote support and training. Furthermore, Muller, Bohne, and Jamboula [14] examine the scope and components of industrial metaverse, Bharti and Sharma [15] discuss the interaction between avatars and humans in decision-making processes, and Aromaa et al. [16] explore the relationship between industrial metaverse and human factors.

This study aims to identify the technologies and processes with which industrial metaverse technology collaborates,

thereby providing insights for researchers and businesses interested in this field. A bibliometric analysis of studies indexed under the "industrial metaverse" keyword in the Web of Science and Scopus databases was conducted to achieve this goal. The analysis encompassed 44 studies from Scopus and 16 from WoS, totaling 60 studies. A comprehensive census of these 60 studies was performed, with data manually compiled and analysed using the VOSviewer application. To further support the study's objectives, figures (Figures 1 and 2) were included to illustrate the conceptual framework and infrastructure of the industrial metaverse.

In this context, the study's research question was determined as follows: Which technologies or processes are most closely aligned with the industrial metaverse?

The study concluded that industrial metaverse technology is associated with processes such as smart processes, employee-technology collaboration, maintaining the functionality of the data and knowledge cycle infrastructure, employee training, and environmental sustainability. Furthermore, technologies like digital twins, blockchain, NFT, virtual reality (VR), augmented reality (AR), the industrial internet of things (IIOT), and AI form the foundational infrastructure of industrial metaverse technology.

A limitation of this study is that only the studies indexed in the Web of Science (WoS) and Scopus databases were examined. Future research encompassing different databases could enhance the findings of this study.

## 2. CONCEPTUAL FRAMEWORK

This section aims to explain the digital transformation process of enterprises within the scope of the data and knowledge cycle. Following this, conceptual knowledge will be provided on metaverse technology and industrial metaverse technology, which are key technological infrastructures required to establish the data and knowledge cycle. The conceptual understanding is supplemented with figures drawn explicitly for this topic.

### 2.1. Digital Transformation

The digital transformation process in enterprises accelerated and took on a new dimension with the announcement of Industry 4.0 in 2011. This initiative introduced a new wave of digital transformation focused on the data and knowledge cycle in the industry [17], [18]. Digital transformation has evolved from merely focusing on technology and automation to emphasising the data and knowledge cycle, which these technological advancements support. By integrating and analysing data obtained from physical and virtual environments, enterprises generate new data and insights, aiming to make processes such as production, marketing, and finance more transparent and efficient. This approach seeks to improve decision-making, predictive maintenance, customer relationship

management, employee skill enhancement, production processes, and waste reduction [19], [20].

The data and knowledge cycle involves the acquisition, storage, analysis, sharing, and preservation of data within enterprises and is considered within the context of technological collaboration [21]. Establishing this cycle requires the development of technological infrastructure for each process within the cycle. Figure 1 illustrates the data and knowledge cycle underpinning businesses' digital transformation.

While the Industry 4.0 process explained digital transformation within a technology-supported data and knowledge cycle, it did not fully address the role of employees. The subsequent Industry 5.0 process expanded on this by incorporating employees, environmental sustainability, and business resilience against various threats into the data and knowledge-oriented digital transformation [22], [23]. This approach builds on the Industry 4.0 foundation by emphasising collaboration between technology and human factors, ensuring a more holistic and sustainable transformation.

Humans interact with their physical environment through the five senses. To fully leverage the metaverse infrastructure, human sensory organs need to be stimulated [27]. This stimulation allows people to experience the virtual environment in the metaverse as if it were real. The technology that facilitates this experience is known as VR technology [28]. Currently, metaverse infrastructures can stimulate the senses of sight and hearing, while research on olfactory stimulation is ongoing. Additionally, no special equipment is necessary to access the metaverse; interactions can occur via computer, tablet, or phone screens using keyboards, mouse, or voice commands. However, technologies such as VR and AR are essential to fully immerse in the metaverse and experience the virtual environment realistically [29].

The technologies underpinning the metaverse infrastructure have gained prominence in the digital transformation of enterprises within the Industry 4.0 framework. Digital twin technology generates analogs of

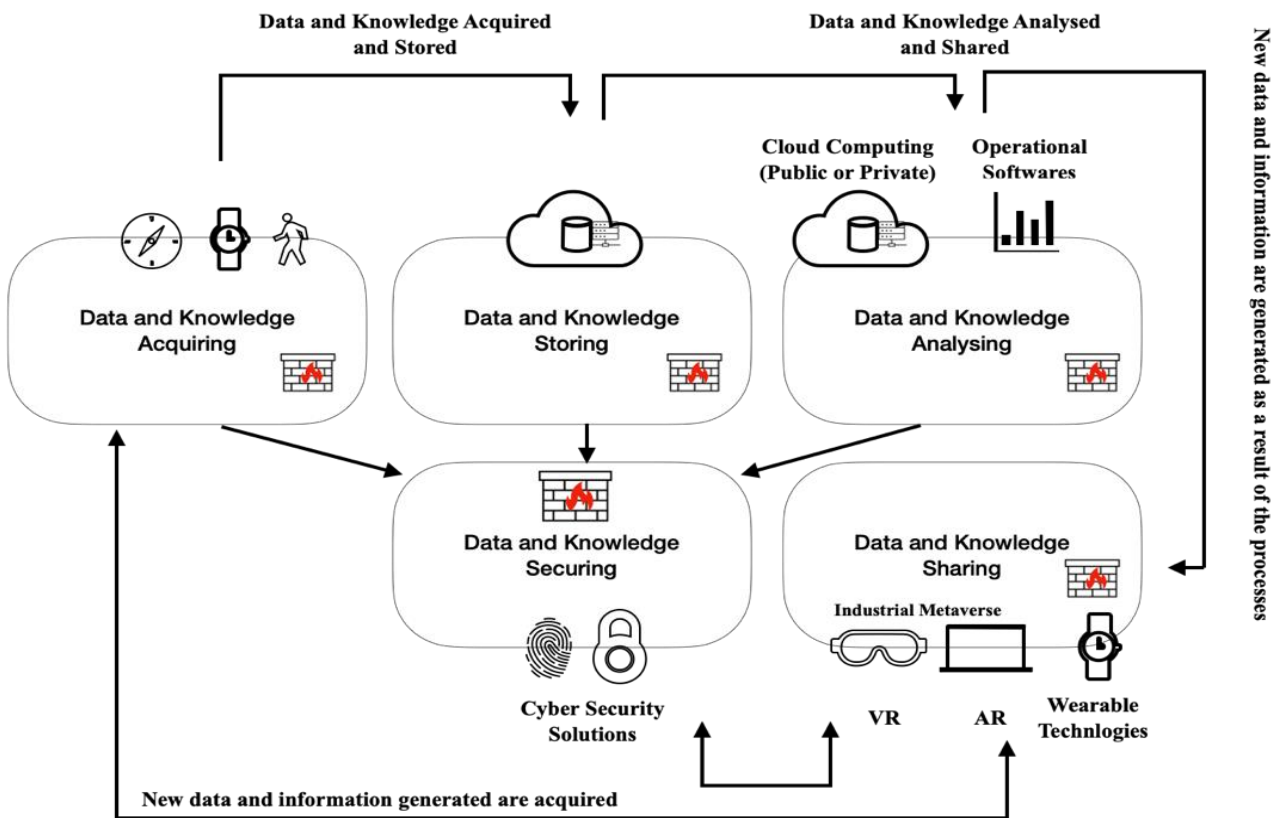


Figure 1. Data and Knowledge Cycle (Source: Author Elaboration)

### 2.2. Metaverse

The term "metaverse" first appeared in the novel "Snow Crash" in 1992, describing a virtual world where users from the physical environment interact with avatars through cyber components [24]. The metaverse can be defined as a virtual environment that emerges from the interaction between the physical and cyber environments through technological infrastructure [25], [26].

physical objects in the virtual environment. The interfaces connecting the metaverse to the physical world are VR and AR technologies. AI and three-dimensional (3-D) modelling solutions configure objects and facilitate user interaction in the metaverse. Blockchain-based digital currencies support the trade infrastructure within the metaverse, and blockchain technology also secures metaverse data. Ownership of virtual objects is managed

through NFT technology, closely related to the blockchain infrastructure [30].

The primary application areas for metaverse technology include sectors such as production, marketing, healthcare, and education. Adopting metaverse technology in

technologies are generally categorised into two groups: operational technologies and information technologies.

**i. Operational technologies:** This category encompasses all the hardware, software, and technologies required to structure, maintain, and ensure the efficiency of operational business processes such as production, procurement, logistics, and finance [31], [32]. Included

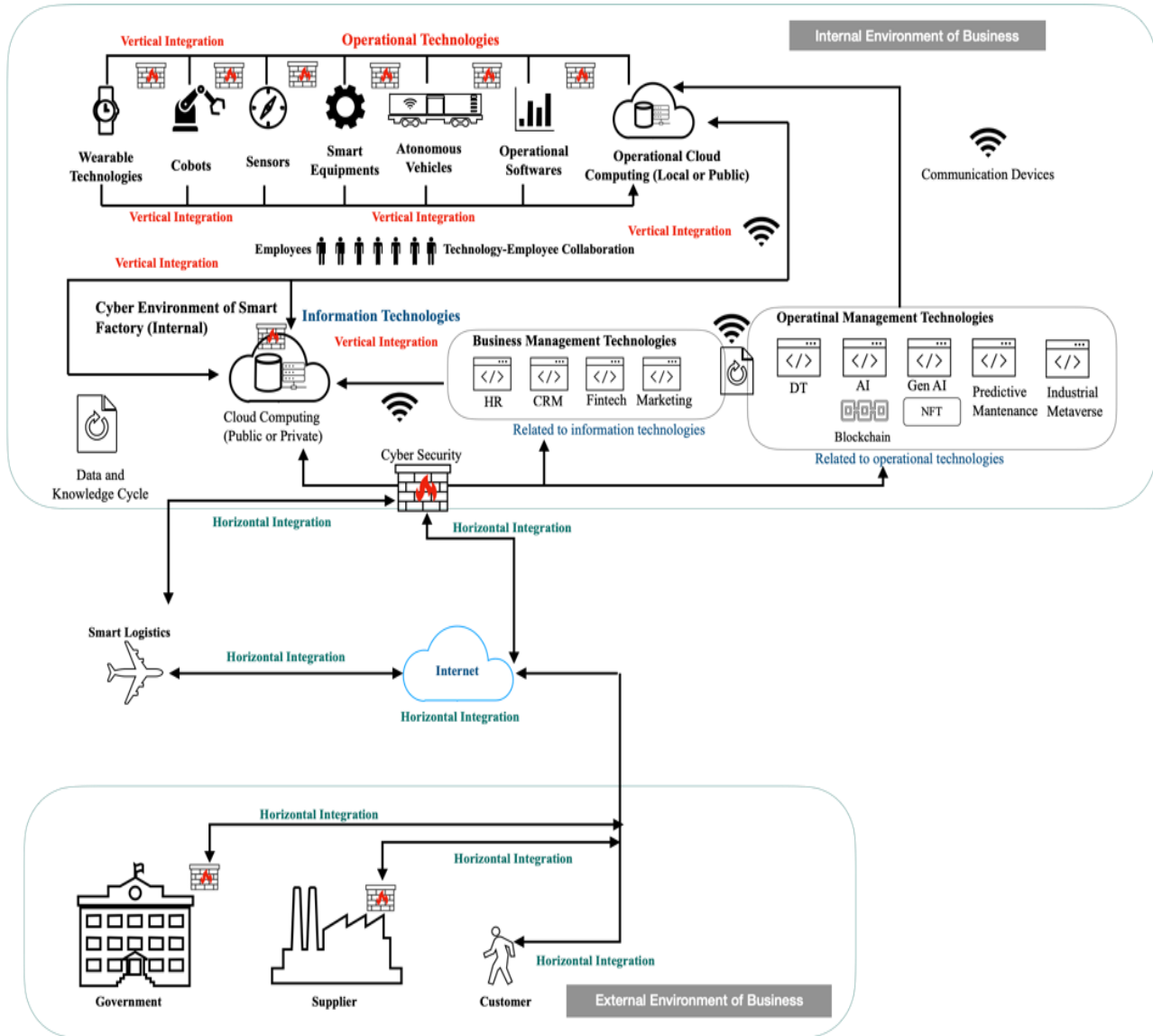


Figure 2. Operational and Information Technologies (Source: Author Elaboration)

manufacturing, engineering, and related processes has given rise to the concept of the industrial metaverse. As the industrial metaverse has emerged, the broader concept of the metaverse has expanded to address issues relevant to individuals and society. In contrast, the industrial metaverse has begun to focus on topics pertinent to various industries.

2.2. Industrial Metaverse

The industrial metaverse refers to the application of metaverse technologies for industrial purposes. To fully define the industrial metaverse, it is necessary to explain the technologies used in the industry conceptually. These

within operational technologies are wearable technologies, cobots, IIOT (sensors and related technologies), autonomous vehicles, machinery, equipment, and associated software. Technologies such as AI, generative AI, and digital twins are employed to maximise efficiency. The benefits from these technologies can be measured through improvements in predictive maintenance, decision-making processes, and the emergence of new business models [32].

**ii. Information technologies:** This category forms the communication and information system infrastructure necessary for establishing the data and knowledge cycle, enabling all business functions to collaborate. The data and knowledge cycle involves integrating data obtained from

operational technologies with other data sources within the business, transforming this data into knowledge, and sharing the transformed knowledge with the necessary systems within the business [19]. Technologies such as cloud computing, cybersecurity equipment, and communication equipment (such as Wi-Fi v6 and 5G) are all considered within the scope of information technologies. Additionally, technologies and software used in human resources, marketing, finance, and accounting processes fall under information technologies. These technologies also support the benefits derived from operational technologies and help eliminate silos between departments within the business [33].

Operational and information technologies play a crucial role in achieving vertical integration within the internal environment of a business and horizontal integration in the external environment. Vertical integration involves the collaboration among departments, machines, employees, and all other internal components within the business. Horizontal integration, on the other hand, refers to the cooperation between the business and its external environment, including suppliers, government units, and customers. This cooperation can lead to the creation of new value networks and innovative business models, thereby significantly contributing to the digital transformation efforts of enterprises [22], [34], [35]. The interplay between operational technologies, information technologies, and the concepts of horizontal and vertical integration is illustrated in Figure 2.

Upon explaining these concepts, industrial metaverse technology can be defined as follows: Industrial metaverse is a technology that integrates various technologies, such as digital twin, blockchain, VR, and AR, through operational and information technologies, involving the transfer of data and knowledge obtained from cobots, autonomous vehicles, employees, machinery, and equipment used in the industry to information technologies via an IIOT infrastructure, to make the operations of enterprises in the physical environment traceable, configurable, and sustainable in the virtual environment, thereby enhancing operational efficiency.

The industrial metaverse concept collaborates with multiple technologies, including IIOT [37], digital twin [38], blockchain, NFTs [39], simulation and modelling software, and AI [14]. VR and AR technologies facilitate interaction between the industrial metaverse environment and employees [13].

By implementing an industrial metaverse infrastructure, enterprises can achieve various benefits, such as improved decision-making processes [11], structured predictive maintenance activities, increased efficiency in production processes, enhanced collaboration between employees and technology [36], more effective marketing activities through a better understanding of customer demand [6] and strengthened customer relations.

Industrial metaverse technology has been studied in the literature since 2021. No bibliometric analysis study has

been identified in this field. Existing studies generally focus on the application areas of industrial metaverse technology [4-5], [9], [11]. However, this study aims to fill the gap in the literature by providing an overview of the components of industrial metaverse technology and its related technologies. This will be achieved by analysing all relevant studies published in the Web of Science (WoS) and Scopus databases. Given its novelty, the following research question (RQ) has been posed to determine which technologies or processes it collaborates with:

RQ1. Which technologies or processes are most closely aligned with the industrial metaverse?

### 3. METHODOLOGY

This study aims to conduct a bibliometric analysis of the studies undertaken in the scope of the keyword “industrial metaverse” in the Web of Science and Scopus databases to address the research question and identify the technologies or processes most frequently collaborated with Industrial metaverse technology. Keywords include those chosen by the authors. For this purpose, studies in the WoS and Scopus databases were analysed from January 1, 2022, to May 15, 2024. The concept of the industrial metaverse began gaining prominence in 2021. For this reason, the first studies on this subject have been included in WoS and Scopus academic databases since 2022. There were no relevant studies in the WoS or Scopus databases before this date. The 60 documents examined are provided in the appendix section of the study.

The study population consists of 44 studies from the Scopus database and 16 studies from the WoS database, totaling 60 studies under the keyword “industrial metaverse”. A complete census was conducted, and no sample was selected. The data were compiled manually and VOSviewer application.

The bibliometric analysis, which reveals the technologies and processes related to industrial metaverse technology, was conducted using the VOSviewer application. In this context, the frequency of word repetition in the article titles and abstracts of the examined studies and the frequency of keyword co-occurrence were analysed. A total of 250 keywords from the related studies (authors’ keywords) and the title and abstract of each study were analysed separately. According to the frequency of co-occurrence of keywords, 22 keywords were selected. In the title and abstract analysis, the words used at least seven times in total in all titles and abstracts were selected to identify the areas related to industrial metaverse technology. The Linear-Logarithmic (Lin/Log Modularity) normalisation method was used to analyse keywords, article titles, and abstracts. Lin/Log modularity can help to visualise more clearly the relationships between research areas, keywords or clusters of topics [40]. The findings of the analysis and the discussion of these findings are presented in the findings and discussion section.



Additionally, four publications were published in both WoS and Scopus databases. These joint publications are shown in dark colours in tables in the Appendix.

#### 4. Findings and Discussions

The findings were analysed in two parts. In the first part, a frequency analysis was performed based on the number of studies and citations, distinguishing between year and database. Additionally, frequency analysis was conducted according to the field of study of the documents within each database. The discrepancy between the number of studies and the number of fields of study is that some studies pertain to multiple fields.

In the second part, the frequency of word occurrences in the article titles and abstracts of the studies and the co-occurrence of the keywords entered by the authors were analyzed using the VOSviewer application. This analysis aims to identify which technologies or processes the Industrial metaverse collaborates with the most.

##### 4.1. Frequency Analysis

The frequency analysis showing the total number of studies related to the analysed studies based on years is presented in Table 1. In contrast, the frequency analysis related to the total number of citations is presented in Table 2.

In Table 1 and Table 2, the colouring is done in white/black tones based on the number of studies in the databases. While the black tone indicates the highest number of studies or citations, the lightening of the tone colour shows a decrease in the number of studies.

Table 1. Frequency Analysis of Studies Published in WoS and Scopus Databases (01.01.2022-05.15.2024)

DATABASES	SCOPUS		WoS	
	Document Count	Percent	Document Count	Percent
2022	5	11,37	4	25
2023	18	40,91	8	50
2024	47	47,72	4	25
Total	44		16	

Table 1 shows that there are more studies in the Scopus database than in the WoS database. The document density in the Scopus database in 2024 is particularly notable. By mid-2024, the highest number of documents was reached compared to previous years.

Table 2. Citations Count by Years

YEARS	SCOPUS	WoS	Total Citations by years
2022		11	108
2023	34	55	89
2024 (Until 05.15.2024)	47	35	82
Total		101	279

The table shows that the Scopus database surpasses the WoS database in terms of the number of citations. Evaluating Table 1 and Table 2 together indicates that studies on the industrial metaverse are expected to intensify in 2024.



Figure 3. Frequency Analysis Graph of Studies Published in Scopus Database (01.01.2022-05.15.2024) by Study Area

Figure 3 shows that the physical sciences category in the Scopus database, with 31 studies, represents the highest number of works. Additionally, 5 studies in the social sciences indicate that the topic is also being explored within this field.

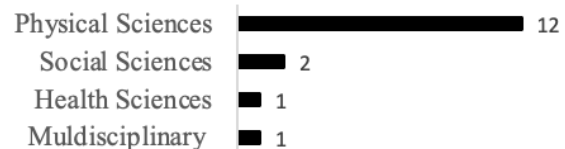


Figure 4. Frequency Analysis Graph of Studies Published in WoS Database (01.01.2022-05.15.2024) by Study Area

Figure 4 shows that the physical sciences category in the WoS database, with 12 studies, represents the highest number of works. Additionally, 2 studies in the social sciences indicate that the topic is also being explored within this field.

##### 4.2. Term Co-occurrence Analysis of Study Titles and Abstracts

To understand the relationship between the words in the titles and abstracts of the 44 studies examined from the

Scopus database, the relationship visualisation of 14 words that occur at least seven times in the titles and abstracts of all studies is shown in Figure 5. The lowest repetition frequency is for “artificial intelligence” with seven occurrences, while the highest is for the term “industrial metaverse” with 103 occurrences.

According to Figure 5, two distinct groups emerged from the analysis. The red cluster includes the terms “industrial metaverse”, “digital twin”, “system”, “technology”, “manufacturing”, “smart manufacturing”, and “artificial intelligence”. This group predominantly consists of technologies related to smart manufacturing processes.

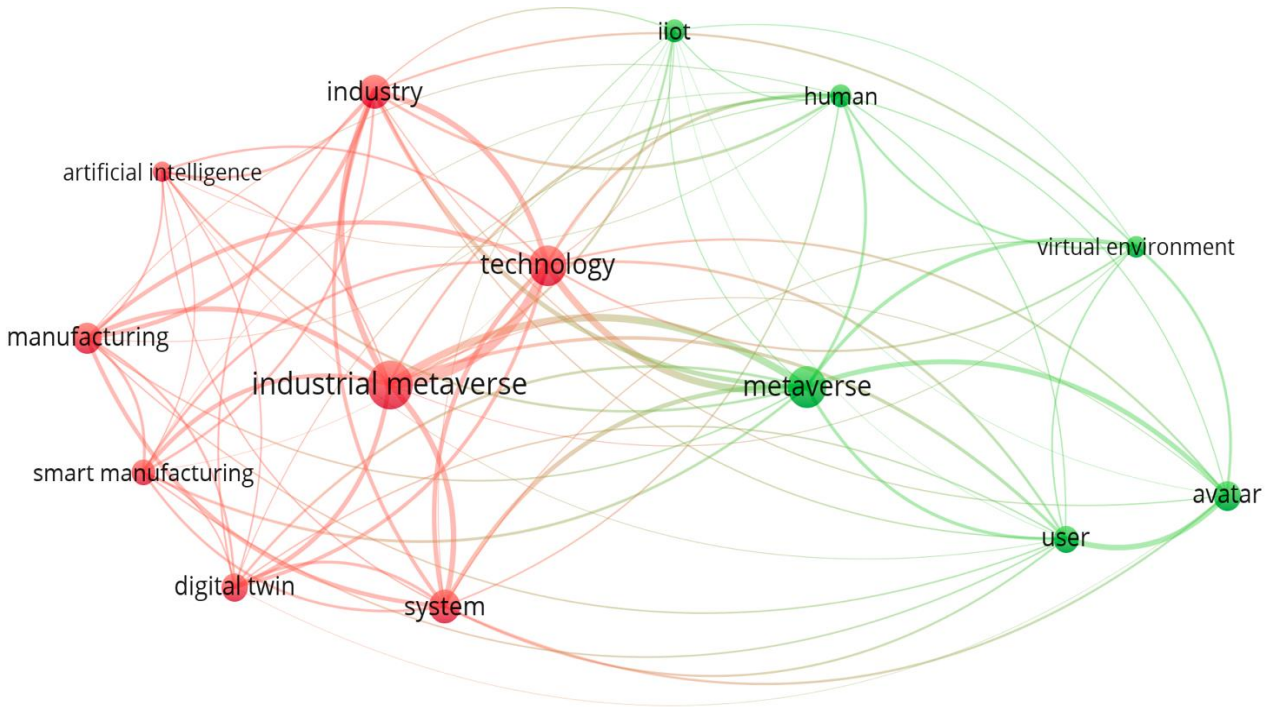


Figure 5. Network Visulation of the Co-occurrence of 14 Terms that Appear at Least 7 Times in the Titles and Abstracts of Documents Obtained from the Scopus Database.

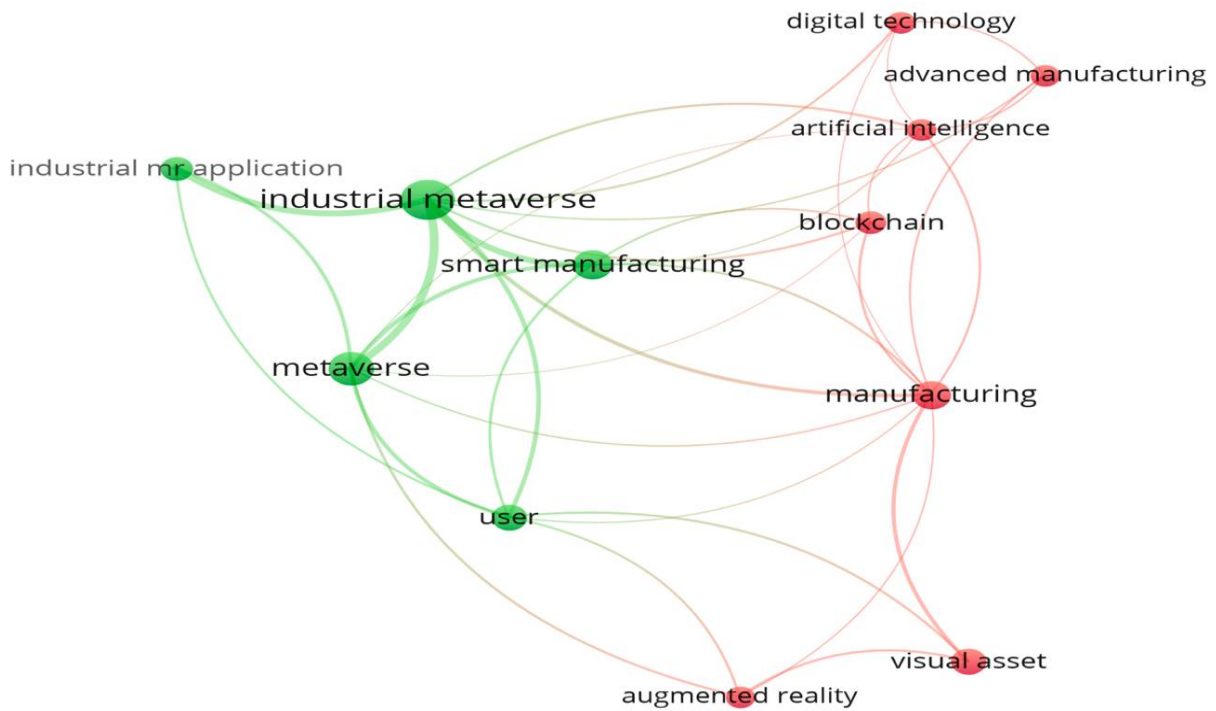


Figure 6. Network Visulation of the Co-occurrence of 14 Terms that Appear at Least 7 Times in the Titles and Abstracts of Documents Obtained from the WoS Database.

Digital twin technology models objects in the physical environment within a virtual environment, forming the infrastructure for industrial metaverse technology. The industrial metaverse environment is generated using these modelled objects, and AI technologies can further streamline this process, significantly contributing to the creation of smart production infrastructure.

The green cluster includes the terms “metaverse”, “human”, “avatar”, “user”, “IIOT”, and “virtual environment”. This group primarily consists of technologies related to the data collection infrastructure of industrial metaverse technology and the representation of employees in the virtual environment. In the industrial metaverse, an employee or person is represented by an avatar, a visual entity depicted as digital knowledge in the virtual environment. Both objects and workers need to have equivalent visual entities in the virtual environment, with their data and knowledge transferred via IIOT. When these technologies are applied to production processes, a twin of the production process is generated in the virtual environment. Together, the technologies in the red and green clusters form the smart manufacturing infrastructure. These findings are supported by the findings of Ren et al. [7], Xiang et al. [8], and Yang, Wang, and Zhao [9] investigate the link between industrial metaverse and smart manufacturing.

To understand the relationship between the words in the titles and abstracts of the 16 studies examined from the WoS database, the relationship visualisation of 14 words that occur at least seven times in the titles and abstracts of all studies is shown in Figure 5. The lowest repetition frequency is for the term “industrial mr application”, with 5 occurrences, while the highest is for “industrial metaverse”, with 35 occurrences.

According to Figure 6, two different clusters emerged from the analysis. The red cluster includes terms such as augmented reality, visual asset, manufacturing, blockchain, artificial intelligence, advanced manufacturing, and digital technology. This cluster generally comprises technologies that form the infrastructure of industrial metaverse technology. In the metaverse, objects in the physical environment are represented as visual assets. Similarly, the equivalent of an employee (user) in the metaverse is an avatar, also a visual entity. These visual entities are represented as digital data or knowledge in the virtual environment. Maintaining this information within a blockchain-based data storage infrastructure is crucial for security in the industrial metaverse. Additionally, the relationship between blockchain infrastructure, digital currency, and NFTs indicates the potential for commerce within the industrial metaverse infrastructure.

The green cluster includes terms such as industrial metaverse, metaverse, user, smart manufacturing, industrial MR (mixed reality) application, and IIOT. This cluster primarily consists of technologies related to the data collection infrastructure of industrial metaverse technology

and the representation of employees in a virtual environment. MR applications combining VR and AR are integral to the industrial metaverse infrastructure. A data and knowledge cycle-based infrastructure is necessary to use industrial metaverse technology through MR technology effectively. IIOT technology, which connects the physical environment to the virtual environment for data and knowledge acquisition, is a key component of this infrastructure. This technology enables the transfer of data and knowledge from the physical to the virtual environment, thus preparing the infrastructure for a smart production environment. These findings are consistent with those of Zheng et al. [12] and Kumar et al. [5]. However, it is important to note that industrial metaverse technology extends beyond smart manufacturing. According to Opperman et al., industrial metaverse also supports processes such as marketing, supply chain management, financial applications, and on-the-job employee training [13].

#### 4.3. Keyword Co-occurrence Analysis

Within the scope of the keyword co-occurrence analysis, the frequency of two keywords appearing together in the same article among 168 from the examined studies from the Scopus database was analysed. The density map resulting from the analysis is shown in Figure 7. The lowest co-occurrence frequency value is 3 for the keyword “smart manufacturing”, while the highest co-occurrence value is 44 for the keyword “industrial metaverse”.

When analysing the keyword co-occurrence in the WoS database, a relationship is observed between industrial metaverse technology and keywords such as digital twin, blockchain, metaverse, IIOT, VR, AR, XR (Extended Reality), and cyber-physical systems. Notably, these technologies are prominent in the Industry 4.0 process and form the infrastructure of the Industry 5.0 process [1]. Environmental sustainability, employee training, and smart manufacturing have emerged as processes related to industrial metaverse technology.

The more significant number of studies in the Scopus database has made the technologies and processes related to industrial metaverse technology more evident in the keyword co-occurrence analysis. These technologies constitute the infrastructure of industrial metaverse technology, and the processes interact with industrial metaverse technology. These findings are supported by Lyu and Fridenfalk [38], Ren et al. [7], Agalwar and Alathur [3], Xiang et al. [8], and Muller et al. [14].

Within the scope of the keyword co-occurrence analysis, the frequency of two keywords appearing together in the same article among 82 keywords from the examined studies was analysed. The density map resulting from the analysis is shown in Figure 8. The lowest co-occurrence frequency value is 2 for the keyword “digital twin”, while the highest is 16 for “industrial metaverse”.

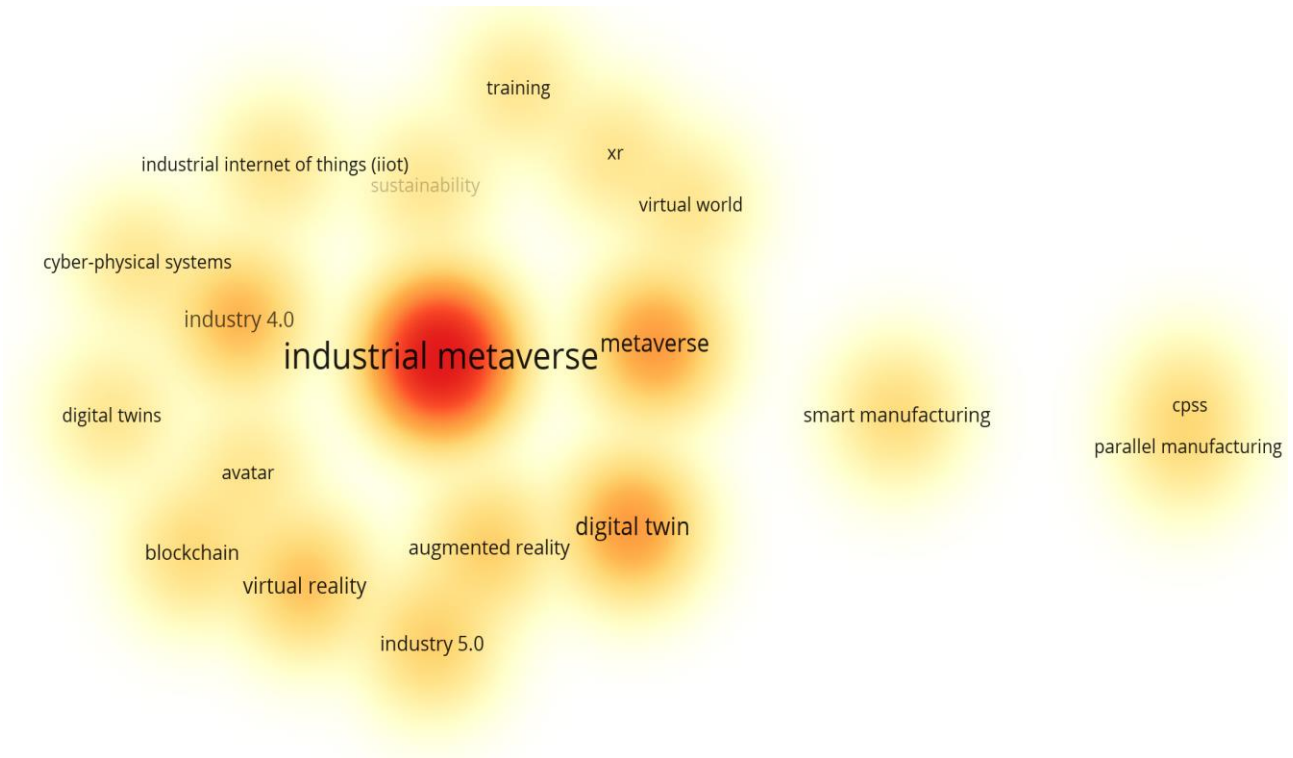


Figure 7. Keyword Co-Occurrence of Studies in the Scopus Database  
Note: cpss is an abbreviation of Cyber Physical System Security

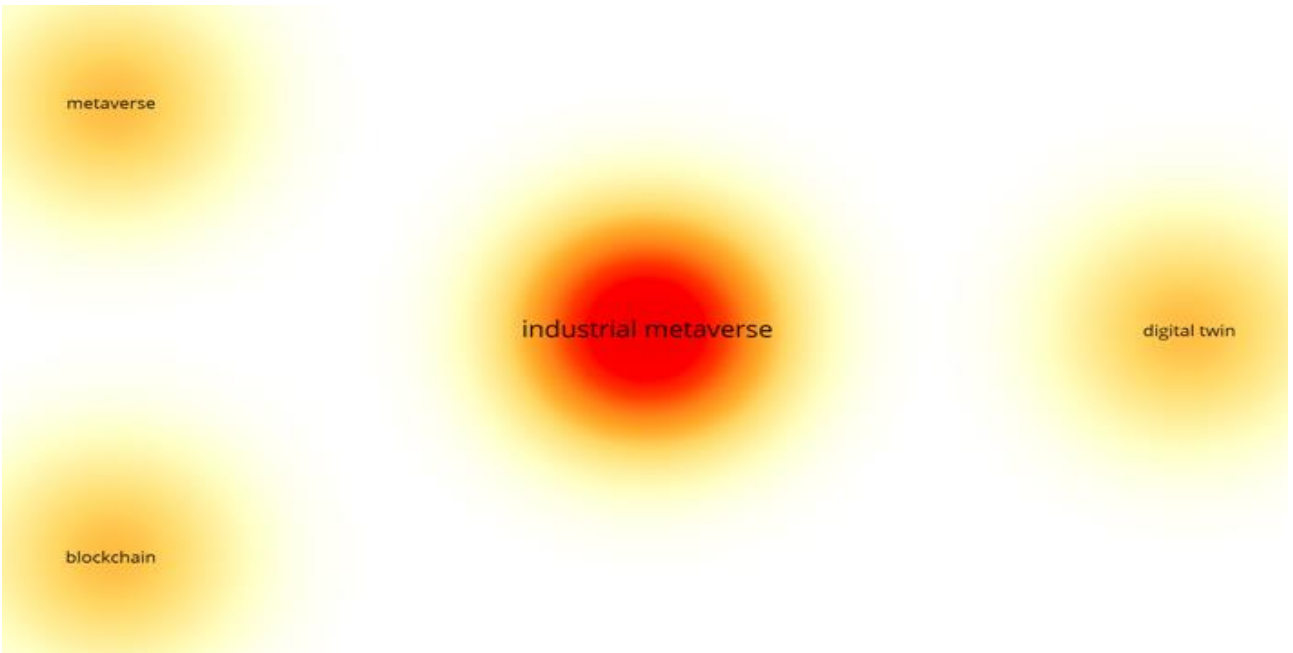


Figure Figure 8. Keyword Co-Occurrence of Studies in the WoS Database

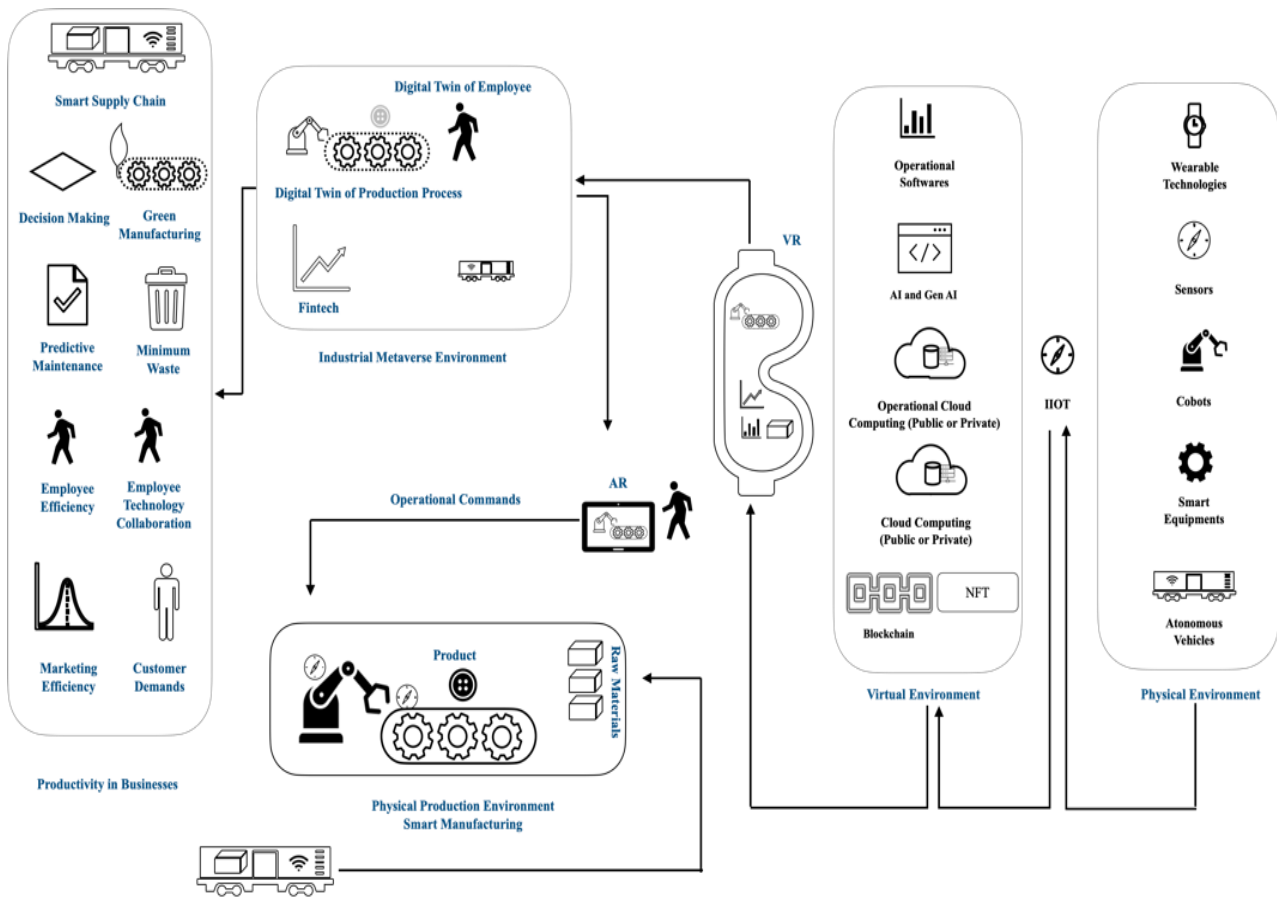


Figure 9. Industrial Metaverse Environment (Source Author Elaboration)

When examining the keyword co-occurrence in the WoS database, a relationship between industrial metaverse technology and the keywords digital twin, blockchain, and metaverse is observed. This relationship indicates that industrial metaverse technology leverages blockchain and digital twin technologies to bring the benefits of metaverse technology to the industrial sector. These findings are supported by the studies of Lyu and Fridenfalk [38] and Ren et al. [7]. Lyu and Fridenfalk examined the relationship between the industrial metaverse and the digital twin. Ren et al. focused on how the industrial metaverse infrastructure can be structured with digital twin, IIoT, and blockchain technologies and how smart manufacturing can be built.

The findings of this study highlight the intricate relationship between various technologies and processes associated with industrial metaverse technology. Industrial metaverse technology is deeply intertwined with smart manufacturing processes, digital twin technology, AI, IIoT, blockchain, and other advanced technologies. The visualisations of keyword relationships in the titles and abstracts of the studies reveal distinct clusters of technologies and concepts. These technologies are driving significant advancements not only in smart manufacturing but also in areas such as marketing, supply chain management, financial applications, and on-the-job employee training. These advancements underscore

industrial metaverse technology's versatility and potential impact across diverse sectors and business processes. Overall, this study contributes to the existing literature by providing a comprehensive overview of the technologies and processes associated with industrial metaverse technology. The findings offer valuable insights for researchers, businesses, and policymakers looking to understand and leverage the potential of industrial metaverse technology in the digital transformation era. As a result of these findings, Figure 9 is drawn to more clearly illustrate which technologies and processes industrial metaverse technology collaborates with, addressing RQ1.

## 5. CONCLUSION

The concept of the industrial metaverse encompasses the application of metaverse technology within the industrial sector. A synergy between various technologies, such as digital twins, VR, AR, blockchain, and IIoT, is required to harness industrial metaverse technology effectively. These technologies are pivotal for configuring enterprises' digital transformation processes in alignment with the Industry 5.0 framework. Industrial metaverse technology presents an opportunity to maximise the benefits derived from these technologies by enabling employees and business stakeholders to achieve optimal efficiency in the digital transformation process. This is accomplished by

visualising data and knowledge gathered during digital transformation using technologies like VR and AR.

This study conducted a bibliometric analysis of studies on the industrial metaverse added to the Scopus and WoS databases since 2021. This analysis aims to identify the technologies and processes associated with industrial metaverse technologies and provide insights for academics and businesses working in this field. The results indicate that the number of studies and citations in the Scopus database is higher than in the WoS database. The technologies associated with industrial metaverse technology include digital twins, IIoT, VR, AR, XR, blockchain, and AI. Industrial metaverse technology is related to smart manufacturing, employee training, the Industry 5.0 process, and environmental sustainability (green manufacturing) processes.

The findings align with existing literature, which has explored various dimensions of industrial metaverse technology. For instance, Kumar et al. have examined the relationship between the industrial metaverse and organisational agility and business performance, while Kshetri has analysed its economic returns. Ren et al., Xiang et al., and Yang, Wang, and Zhao have investigated the link between the industrial metaverse and smart manufacturing. Kshetri and Dwivedi have studied its relationship with environmental sustainability, and Gattullo et al. have focused on knowledge sharing. Furthermore, Opperman et al. have explored the relationship between the industrial metaverse and remote support and training, while Muller, Bohne, and Jamboula have examined its scope and components. Bharti and Sharma discussed the interaction between avatars and humans in decision-making processes, and Aromaa et al. explored the relationship between the industrial metaverse and human factors.

This study contributes to the literature by offering a comprehensive bibliometric analysis, thereby identifying the core technologies and processes associated with the industrial metaverse. As research on the industrial metaverse continues to grow, more use cases and potential benefits are expected to emerge. The current focus of studies from both databases is on the relationship between industrial metaverse applications and the smart manufacturing process. However, future research exploring the applications of the industrial metaverse in different areas and the technologies that can structure these applications will broaden the scope of this field. For instance, investigating how industrial metaverse technology can be leveraged to enhance remote collaboration, optimise supply chain management, or advance predictive maintenance practices could provide valuable insights and further solidify its role in digital transformation. The potential for real-time data processing and decision-making through like IIoT, digital twins, AI, VR AR, blockchain and NFT advanced technologies and smart processes, combined with the immersive and interactive capabilities of the metaverse, could reshape industrial operations.

A limitation of this study is that only the studies indexed in the Web of Science (WoS) and Scopus databases were examined. Future research encompassing different databases, as well as longitudinal studies tracking the evolution of industrial metaverse applications over time, could enhance the findings of this study. Additionally, case studies and real-world implementations of industrial metaverse technology could provide practical insights and validate the theoretical frameworks developed through bibliometric analysis.

## REFERENCES

- [1] K. Tantawi, I. Fidan, O. Huseynov, Y. Musa, and A. Tantawy, "Advances in industry 4.0: from intelligentization to the industrial metaverse," *International Journal Of Interactive Design And Manufacturing - Ijidem*, 2024.
- [2] P. Bhattacharya et al., "Towards Future Internet: The Metaverse Perspective for Diverse Industrial Applications," *Mathematics*, vol. 11, no. 4, Feb. 2023.
- [3] A. Agarwal and S. Alathur, "Metaverse revolution and the digital transformation: intersectional analysis of Industry 5.0," *Transforming Government: People, Process and Policy*, vol. 17, no. 4, pp. 688–707, Nov. 2023, doi: 10.1108/TG-03-2023-0036.
- [4] X. Yao, N. Ma, J. Zhang, K. Wang, E. Yang, and M. Faccio, "Enhancing wisdom manufacturing as industrial metaverse for industry and society 5.0," *J Intell Manuf*, vol. 35, no. 1, pp. 235–255, 2024, doi: 10.1007/s10845-022-02027-7.
- [5] A. Kumar, A. Shankar, A. Behl, B. B. Gupta, and S. Mavuri, "Lights, Camera, Metaverse! Eliciting Intention to Use Industrial Metaverse, Organizational Agility, and Firm Performance," *Journal of Global Information Management*, vol. 31, no. 8, 2023, doi: 10.4018/JGIM.333169.
- [6] N. Kshetri, "The Economics of the Industrial Metaverse," *IT Prof*, vol. 25, no. 1, pp. 84–88, 2023, doi: 10.1109/MITP.2023.3236494.
- [7] Ren L.; Dong J.; Zhang L.; Laili Y.; Wang X.; Qi Y.; Li B.H.; Wang L.; Yang L.T.; Deen M.J. "Industrial Metaverse for Smart Manufacturing: Model, Architecture, and Applications," *IEEE Trans Cybern*, vol. 54, no. 5, pp. 2683–2695, 2024, doi: 10.1109/TCYB.2024.3372591.
- [8] W. Xiang, K. Yu, F. Han, L. Fang, D. He, and Q.-L. Han, "Advanced Manufacturing in Industry 5.0: A Survey of Key Enabling Technologies and Future Trends," *IEEE Trans Industry Inform*, vol. 20, no. 2, pp. 1055–1068, 2024, doi: 10.1109/TII.2023.3274224.
- [9] J. Yang, X. Wang, and Y. Zhao, "Parallel Manufacturing for Industrial Metaverses: A New Paradigm in Smart Manufacturing," *IEEE/CAA Journal of Automatica Sinica*, vol. 9, no. 12, pp. 2063–2070, 2022, doi: 10.1109/JAS.2022.106097.
- [10] N. Kshetri and Y. K. Dwivedi, "Pollution-reducing and pollution-generating effects of the metaverse," *Int J Inf Manage*, vol. 69, 2023, doi: 10.1016/j.ijinfomgt.2023.102620.
- [11] Huang S.; Luo H.; Zheng P.; Ma N.; Chen J.; Mo G.; Jing S. "Towards Industrial Metaverse: Opportunities and challenges," *Journal of Advanced Manufacturing Science and Technology*, vol. 3, no. 4, 2023.

- [12] Zheng Z.; Li T.; Li B.; Chai X.; Song W.; Chen N.; Zhou Y.; Lin Y.; Li R. "Industrial Metaverse: Connotation, Features, Technologies, Applications and Challenges," in *Communications in Computer and Information Science*, vol. 1712 CCIS, pp. 239–263. 2022
- [13] Oppermann L.; Uzun Y.; Buchholz F.; Riedlinger U.; Fuchs S.; Stenzel H.; Odenthal L.; Altepost A.; Bau M. "Industrial Metaverse? Human-Centred Design for Collaborative Remote Maintenance and Training Using XR-Technologies," in *Springer Proceedings in Business and Economics*. pp. 195–209. doi: 10.1007/978-3-031-50559-1\_15. 2024
- [14] M. Muller, T. Bohne, and T. Jamboula, "Identifying key factors for the development of the Industrial Metaverse," 2023.
- [15] P. Bharti and V. K. Sharma, "Enhancing Decision-Making: The Significance of Humanized Avatars in Complex Scenarios," in *CEUR Workshop Proceedings*, vol. 3660. 2024
- [16] S. Aromaa, P. Heikkilä, E. Kaasinen, H. Lammi, A. Tammela, and K. Salminen, "Human factors and ergonomics considerations in the industrial metaverse," *International Journal of Human Factors and Ergonomics*, vol. 11, no. 1, pp. 4–27, 2024.
- [17] A. Alvarenga, F. Matos, R. Godina, and J. C. O. Matias, "Digital transformation and knowledge management in the public sector," *Sustainability (Switzerland)*, vol. 12, no. 14, Jul. 2020, doi: 10.3390/su12145824.
- [18] C. A. Schumann and C. Tittmann, "Digital business transformation in the context of knowledge management," in *Proceedings of the European Conference on Knowledge Management, ECKM, 2015*, vol. 0.
- [19] X. Zhang, Y. Y. Xu, and L. Ma, "Information technology investment and digital transformation: the roles of digital transformation strategy and top management," *Business Process Management Journal*, vol. 29, no. 2, 2023, doi: 10.1108/BPMJ-06-2022-0254.
- [20] P. Planing, W. Pfoertsch, and D. A. G., "The Digital Business Transformation Paths From Manufacturer To Digital Ecosystem Provider - Analyzing the Strategic Options of Large Corporations Towards Digitalization," *Allied Academies Summer Internet Conference*, vol. 18, no. 2, pp. 66–70, 2016.
- [21] K. Buntak, M. Kovačić, and I. Martinčević, "Impact of Digital Transformation on Knowledge Management in Organization," *Advances in Business Related Scientific Research Journal*, vol. 11, no. 1, pp. 36–47, 2020.
- [22] J. H. Ortiz, W. G. Marroquin, and L. Z. Cifuentes, "Industry 4.0: Current Status and Future Trends," in *Industry 4.0: Current Status and Future Trends*, 1st ed., J. H. Ortiz, Ed. Intechopen, 2020.
- [23] A. Majumdar, H. Garg, and R. Jain, "Managing the barriers of Industry 4.0 adoption and implementation in textile and clothing industry: Interpretive structural model and triple helix framework," *Comput Ind*, vol. 125, p. 103372, Feb. 2021.
- [24] S. Liu, J. Xie, and X. Wang, "QoE enhancement of the industrial metaverse based on Mixed Reality application optimization," *Displays*, vol. 79, Sep. 2023.
- [25] J. Lee and P. Kundu, "Integrated cyber-physical systems and industrial metaverse for remote manufacturing," *Manuf Lett*, vol. 34, pp. 12–15, Oct. 2022.
- [26] H. Wang et al., "A Survey on the Metaverse: The State-of-the-Art, Technologies, Applications, and Challenges," *IEEE Internet Things J*, vol. 10, no. 16, 2023.
- [27] S. X. Zhao, L. Qiao, R. J. Zhang, H. H. Zhang, and F. Y. Ye, "The Theoretical Principles and Practical Scenarios for Exploring Metaverse," *Journal of Library Science in China*, vol. 48, no. 6, pp. 6–15, 2022.
- [28] P. Khanna, R. Karim, and J. Kumari, "Issues and Challenges in Implementing the Metaverse in the Industrial Contexts from a Human-System Interaction Perspective," in *Lecture Notes in Mechanical Engineering*, pp. 303–318. 2024
- [29] M. Prummer, E. Regnath, S. Singh, and H. Kosch, "From Virtual Worlds to Real-World Impact: An Industrial Metaverse Survey," in *Lecture Notes in Networks and Systems*, vol. 919 LNNS, pp. 592–613. 2024.
- [30] A. Abbasi, "The Metaverse And NFTs: What Agencies And Brand Leaders Should Know," 2022. <https://www.forbes.com/sites/forbesagencycouncil/2022/01/06/the-metaverse-and-nfts-what-agencies-and-brand-leaders-should-know/?sh=2edfec54200b> (accessed 02/06/2022).
- [31] M. Wharton, "Effectively integrating physical security technology into the operational technology domain," *Cyber Security: A Peer-Reviewed Journal*, vol. 4, 2020.
- [32] C. Paulsen, "The Future of IT Operational Technology Supply Chains," *Computer (Long Beach Calif)*, vol. 53, no. 1, 2020.
- [33] E. Beulen and R. Bode, "An information technology and innovation committee to guide digital transformations," *Corporate Board role duties and composition*, vol. 17, no. 2, 2021.
- [34] Deloitte, "Industry 4.0: Are you ready?," *Deloitte Review*, no. 22, 2018.
- [35] C. J. Bartodziej, *The concept Industry 4.0. An Empirical Analysis of Technologies and Applications in Production Logistics 2017*. Springer Gabler Germany
- [36] K. Salminen and S. Aromaa, "Industrial metaverse - company perspectives," in *Procedia Computer Science*, vol. 232, pp. 2108–2116. 2024.
- [37] H. Endres, M. Indulska, and A. Ghosh, "Unlocking the potential of Industrial Internet of Things (IIOT) in the age of the industrial metaverse: Business models and challenges" *Industrial Marketing Management*, vol. 119 90-107. 2024
- [38] Z. Lyu and M. Fridenfalk, "Digital twins for building industrial metaverse," *Journal of Advanced Research*, 2023
- [39] D. Mourtzis, J. Angelopoulos, and N. Panopoulos, "Blockchain Integration in the Era of Industrial Metaverse," *Applied Sciences (Switzerland)*, vol. 13, no. 3. MDPI, Feb. 01, 2023.
- [40] S. Sri Amudha and R. Sevukan, "Application of author bibliographic coupling analysis and author keywords ranking in identifying research fronts of Indian Neurosciences research" (2019). *Library Philosophy and Practice*

## APPENDIX

Table 3. Study Lists in the WoS Database

Author	Title	Year
Kshetri, N	The Economics of the Industrial Metaverse	2023
<b>Kumar, A; Shankar, A; Behl, A; Gupta, BB; Mavuri, S</b>	<b>Lights, Camera, Metaverse! Eliciting Intention to Use Industrial Metaverse, Organizational Agility, and Firm Performance</b>	<b>2023</b>
Kshetri, N; Dwivedi, YK	Pollution-reducing and pollution-generating effects of the metaverse	2023
<b>Ren, L; Dong, JB; Zhang, L; Laili, Y; Wang, XK; Qi, Y; Li, BH; Wang, LH; Yang, LT; Deen, MJ</b>	<b>Industrial Metaverse for Smart Manufacturing: Model, Architecture, and Applications</b>	<b>2024</b>
Laviola, E; Gattullo, M; Manghisi, VM; Fiorentino, M; Uva, AE	Minimal AR: visual asset optimization for the authoring of augmented reality work instructions in manufacturing	2022
Tantawi, K; Fidan, I; Huseynov, O; Musa, Y; Tantawy, A	Advances in industry 4.0: from intelligentization to the industrial metaverse	2024
Jaimini, U; Zhang, TT; Brikis, GO; Sheth, A	iMetaverseKG: Industrial Metaverse Knowledge Graph to Promote Interoperability in Design and Engineering Applications	2022
Gattullo, M; Laviola, E; Evangelista, A; Fiorentino, M; Uva, AE	Towards the Evaluation of Augmented Reality in the Metaverse: Information Presentation Modes	2022
Liu, SG; Xie, JC; Wang, XW	QoE enhancement of the industrial metaverse based on Mixed Reality application optimization*	2023
<b>Hou, J; Chen, G; Li, ZJ; He, W; Gu, SD; Knoll, A; Jiang, CJ</b>	<b>Hybrid Residual Multiexpert Reinforcement Learning for Spatial Scheduling of High-Density Parking Lots</b>	<b>2024</b>
Patterson, EA	Engineering design and the impact of digital technology from computer-aid engineering to industrial metaverses: A perspective	2024
Siriweera, A; Naruse, K	QoS-Aware Federated Crosschain-Based Model-Driven Reference Architecture for IIoT Sensor Networks in Distributed Manufacturing	2023
<b>Xiang, W; Yu, K; Han, FL; Fang, L; He, DH; Han, QL</b>	<b>Advanced Manufacturing in Industry 5.0: A Survey of Key Enabling Technologies and Future Trends</b>	<b>2024</b>
Aromaa, S; Heikkilä, P; Kaasinen, E; Lammi, H; Tammela, A; Salminen, K	Human factors and ergonomics considerations in the industrial metaverse	2024
Yao, XF; Ma, NF; Zhang, JM; Wang, KS; Yang, EF; Faccio, M	Enhancing wisdom manufacturing as industrial metaverse for industry and society 5.0	2024
Camacho-Muñoz, GA; Franco, JCM; Nope-Rodríguez, SE; Loaliza-Correa, H; Gil-Parga, S; alvarez-Martínez, D	6D-ViCuT: Six degree-of-freedom visual cuboid tracking dataset for manual packing of cargo in warehouses	2023



Table 4. Study Lists in the Scopus Database

Author	Title	Year
Khanna P.; Karim R.; Kumari J.	Issues and Challenges in Implementing the Metaverse in the Industrial Contexts from a Human-System Interaction Perspective	2024
Lyu Z.; Fridenfalk M.	Digital twins for building industrial metaverse	2023
Liu S.; Xie J.; Wang X.	QoE enhancement of the industrial metaverse based on Mixed Reality application optimization	2023
Bellalouna F.; Puljiz D.	Use case for the Application of the Industrial Metaverse Approach for Engineering Design Review	2023
Martínez-Gutiérrez A.; Díez-González J.; Perez H.; Araújo M.	Towards industry 5.0 through metaverse	2024
Yao X.; Ma N.; Zhang J.; Wang K.; Yang E.; Faccio M.	Enhancing wisdom manufacturing as industrial metaverse for industry and society 5.0	2024
Bharti P.; Sharma V.K.	Avatars in the Metaverse: From Social Interaction to Collaborative Work and Beyond	2024
Endres H.; Indulska M.; Ghosh A.	Unlocking the potential of Industrial Internet of Things (IIOT) in the age of the industrial metaverse: Business models and challenges	2024
Kshetri N.; Dwivedi Y.K.	Pollution-reducing and pollution-generating effects of the metaverse	2023
<b>Ren L.; Dong J.; Zhang L.; Laili Y.; Wang X.; Qi Y.; Li B.H.; Wang L.; Yang L.T.; Deen M.J.</b>	<b>Industrial Metaverse for Smart Manufacturing: Model, Architecture, and Applications</b>	<b>2024</b>
Wenzheng L.	Conceptual Technology Features and System Architecture of Industrial Metaverse	2023
Aung N.; Dhelim S.; Chen L.; Ning H.; Atzori L.; Kechadi T.	Edge-Enabled Metaverse: The Convergence of Metaverse and Mobile Edge Computing	2024
Liu C.; Tang D.; Wang Z.	AR-Driven Industrial Metaverse for the Auxiliary Maintenance of Machine Tools in IoT-Enabled Manufacturing Workshop	2023
Wang X.; Wang Y.; Yang J.; Jia X.; Li L.; Ding W.; Wang F.-Y.	The survey on multi-source data fusion in cyber-physical-social systems: Foundational infrastructure for industrial metaverses and industries 5.0	2024
<b>Kumar A.; Shankar A.; Behl A.; Gupta B.B.; Mavuri S.</b>	<b>Lights, Camera, Metaverse! Eliciting Intention to Use Industrial Metaverse, Organizational Agility, and Firm Performance</b>	<b>2023</b>
Schultheiß A.; Polovoj E.; Dolanovic S.; Gutsche K.	Digital Service Twin - Design Criteria, Requirements and Scope for Service Management	2023
Kim D.B.; Bajestani M.S.; Shao G.; Jones A.; Noh S.D.	CONCEPTUAL ARCHITECTURE OF DIGITAL TWIN WITH HUMAN-IN-THE-LOOP - BASED SMART MANUFACTURING	2023
Ahr P.; Dreyer J.; Reski M.; Lipps C.; Tönjes R.; Schotten H.D.	Industry 4.0 Security Trust Anchors: Considering Supply Voltage Effects on SRAM-PUF Reliability	2023
Siriweera A.; Naruse K.	QoS-Aware Federated Crosschain-Based Model-Driven Reference Architecture for IIoT Sensor Networks in Distributed Manufacturing	2023
Nateghi A.; Mosharraf M.	Architecting the Future: A Model for Enterprise Integration in the Metaverse	2023
Bharti P.; Sharma V.K.	Enhancing Decision-Making: The Significance of Humanized Avatars in Complex Scenarios	2024
Rosilius M.; Wilhelm M.; von Eitzen I.; Decker S.; Damek S.; Braeutigam V.	Sustainable Solutions by the Use of Immersive Technologies for Repurposing Buildings	2023
<b>Hou J.; Chen G.; Li Z.; He W.; Gu S.; Knoll A.; Jiang C.</b>	<b>Hybrid Residual Multiexpert Reinforcement Learning for Spatial Scheduling of High-Density Parking Lots</b>	<b>2024</b>
Zhou Y.; Li T.; Li B.; Wu G.; Meng X.; Guo J.; Wan N.; Zhu J.; Li S.; Song W.; Su C.; Chen N.; Xing Y.; Wang Q.; Lin Y.; Li R.	Research on Intelligent Manufacturing Training System Based on Industrial Metaverse	2024
Salminen K.; Aromaa S.	Industrial metaverse - company perspectives	2024
Huang S.; Luo H.; Zheng P.; Ma N.; Chen J.; Mo G.; Jing S.	Towards Industrial Metaverse: Opportunities and challenges	2023
Ullrich M.; Thalappully R.; Heieck F.; Lüdemann-Ravit B.	Virtual Commissioning of Linked Cells Using Digital Models in an Industrial Metaverse	2024
Ullrich M.; Heller J.; Knüpper J.; Lüdemann-Ravit B.	A Reference Architecture for an Industrial Metaverse; [Eine Referenzarchitektur für das Industrial Metaverse]	2024
Jiang Y.; Kaynak O.; Yin S.; Luo H.; Liu M.	Industrial Metaverse: Solutions from a Higher-Dimensional World	2024
Laviola E.; Gattullo M.; Manghisi V.M.; Fiorentino M.; Uva A.E.	Minimal AR: visual asset optimization for the authoring of augmented reality work instructions in manufacturing	2022
Oppermann L.; Uzun Y.; Buchholz F.; Riedlinger U.; Fuchs S.; Stenzel H.; Odenthal L.; Altepost A.; Bau M.	Industrial Metaverse? Human-Centred Design for Collaborative Remote Maintenance and Training Using XR-Technologies	2024
Yang J.; Wang X.; Zhao Y.	Parallel Manufacturing for Industrial Metaverses: A New Paradigm in Smart Manufacturing	2022
Aromaa S.; Heikkilä P.; Kaasinen E.; Lammi H.; Tammela A.; Salminen K.	Human factors and ergonomics considerations in the industrial metaverse	2024
Muller M.; Bohne T.; Jamboula T.	Identifying key factors for the development of the Industrial Metaverse	2023

Tantawi K.; Fidan I.; Huseynov O.; Musa Y.; Tantawy A.	Advances in industry 4.0: from intelligentization to the industrial metaverse	2024
Huang J.	Control and Decision Theory in the Metaverse: A Survey; [元宇宙下的控制与决策综述]	2023
Gattullo M.; Laviola E.; Evangelista A.; Fiorentino M.; Uva A.E.	Towards the Evaluation of Augmented Reality in the Metaverse: Information Presentation Modes	2022
<b>Xiang W.; Yu K.; Han F.; Fang L.; He D.; Han Q.-L.</b>	<b>Advanced Manufacturing in Industry 5.0: A Survey of Key Enabling Technologies and Future Trends</b>	<b>2024</b>
Patterson E.A.	Engineering design and the impact of digital technology from computer-aided engineering to industrial metaverses: A perspective	2024
Xie J.; Zheng Z.; Wang X.; Meng H.; Liu S.; Li S.	Preliminary research on the operation mode of virtual-real integration in fully-mechanized mining face based on industrial metaverse; [基于工业元宇宙的综采工作面虚实融合运行模式初步探索]	2023
Zheng Z.; Li T.; Li B.; Chai X.; Song W.; Chen N.; Zhou Y.; Lin Y.; Li R.	Industrial Metaverse: Connotation, Features, Technologies, Applications and Challenges	2022
Zhao S.X.; Qiao L.; Zhang R.J.; Zhang H.H.; Ye F.Y.	The Theoretical Principles and Practical Scenarios for Exploring Metaverse	2022
Prummer M.; Regnath E.; Singh S.; Kosch H.	From Virtual Worlds to Real-World Impact: An Industrial Metaverse Survey	2024
Camacho-Muñoz G.A.; Franco J.C.M.; Nope-Rodríguez S.E.; Loaiza-Correa H.; Gil-Parga S.; Álvarez-Martínez D.	6D-ViCuT: Six degree-of-freedom visual cuboid tracking dataset for manual packing of cargo in warehouses	2023

# Dijital Araçlarla İnfomal Öğrenme Yeterlilik Ölçeğinin Geliştirilmesi

*Araştırma Makalesi/Research Article*

 Gürbüz OCAK<sup>1</sup>,  Gülin ÜTEBAY<sup>2\*</sup>,  Ramazan YURTSEVEN<sup>3</sup>

<sup>1</sup>Eğitim Bilimleri Bölümü, Afyon Kocatepe Üniversitesi, Afyonkarahisar, Türkiye

<sup>2</sup>Uzaktan Eğitim Uygulama ve Araştırma Merkezi, Afyon Kocatepe Üniversitesi, Afyonkarahisar, Türkiye

<sup>3</sup>Eğitim Bilimleri Bölümü, Afyon Kocatepe Üniversitesi, Afyonkarahisar, Türkiye

[gocak@aku.edu.tr](mailto:gocak@aku.edu.tr), [utebay@aku.edu.tr](mailto:utebay@aku.edu.tr), [ryurtseven@aku.edu.tr](mailto:ryurtseven@aku.edu.tr)

(Geliş/Received:06.02.2024; Kabul/Accepted:30.08.2024)

DOI: 10.17671/gazibtd.1431471

**Özet**— Bu çalışmada, üniversite öğrencilerinin dijital araçlarla infomal öğrenme yeterliliklerini ölçmek için kullanılacak bir ölçme aracı geliştirilmesi amaçlanmıştır. Bu kapsamda, "Dijital Araçlarla İnfomal Öğrenme Yeterlilik Ölçeği" (DAİÖYÖ) şeklinde adlandırılan likert tipi bir ölçek oluşturulmuştur. Çalışmada karma araştırma yöntemlerinden keşfedici sıralı deseni kullanılmıştır. Araştırmanın çalıma grubunu açımlayıcı faktör analizi (AFA) için 287, doğrulayıcı faktör analizi (DFA) için 200 üniversite öğrencisi oluşturmuştur. AFA öncesinde, veri uygunluğu için KMO değeri hesaplanmış ve .915 olarak bulunmuştur. AFA sonucunda ölçeğin dört faktörlü bir yapıya sahip olduğu, toplam varyansının %55'ini açıkladığı belirlenmiştir. AFA sonucunda, 22 madde, 4 faktörden (dijital öz yeterlilik, teknolojik isteklilik, bilişsel yeterlilik ve etkin kullanım) oluşan ölçek elde edilmiştir. Ardından doğrulayıcı faktör analizi (DFA) uygulanarak elde edilen ölçeğe ait yapının kabul edilebilir uyum indekslerine sahip olduğu bulunmuştur. Son olarak güvenilirlik için yapılan ölçeğin alt boyutlarına ilişkin cronbach alpha değerleri sırayla; 0.85, 0.86, 0.85 ve 0.83; toplam alpha değeri ise .927 olarak bulunmuştur. Analizler sonucunda, üniversite öğrencilerinin dijital araçlarla infomal öğrenme yeterliliklerini belirlemek amacıyla kullanılacak geçerli ve güvenilir bir ölçek geliştirilmiştir.

**Anahtar Kelimeler**— dijital araçlar, infomal öğrenme, üniversite öğrencileri, ölçek geliştirme.

## Development of Informal Learning Qualification Scale With Digital Tools

**Abstract**— In this study, it was aimed to develop a measurement tool to be used to measure the informal learning competencies of university students with digital tools. In this context, a Likert-type scale called "Informal Learning Competency Scale with Digital Tools" (ILCSDT) was created. Exploratory sequential design, one of the mixed research methods, was used in the study. The study group of the research consisted of 287 university students for exploratory factor analysis (EFA) and 200 university students for confirmatory factor analysis (CFA). Before EFA, the KMO value for data suitability was calculated and found to be .915. As a result of EFA, it was determined that the scale had a four-factor structure and explained 55% of the total variance. As a result of EFA, a scale consisting of 22 items and 4 factors (digital self-efficacy, technological willingness, cognitive competence and effective use) was obtained. Then, by applying confirmatory factor analysis (CFA), it was found that the structure of the scale obtained had acceptable fit indices. Finally, the Cronbach alpha values for the sub-dimensions of the scale for reliability are as follows; 0.85, 0.86, 0.85 and 0.83; The total alpha value was found to be .927. As a result of the analyses, a valid and reliable scale was developed that can be used to determine the informal learning proficiency of university students with digital tools.

**Keywords**— tools, informal learning, university students, scale development

## 1. GİRİŞ (INTRODUCTION)

21. yüzyılda hızlı bir şekilde gelişen teknolojiler ve dijital araçlar, bilgiye kolay erişim ve hızlı iletişim imkânları sunarak eğitim ve öğrenme sürecinde önemli ölçüde etkili olmaktadır. Dijital araçlarla öğrenme, gerek okullarda planlı olarak işe koşulmakta, gerekse bireysel öğrenme sürecinde etkin şekilde kullanılmaktadır. Bu noktada günümüzde farklı dijital araçların farklı amaçlar için kullanıldığı ve her birinin belirli öğrenme alanlarına katkı sağladığı görülmektedir. Dolayısıyla, eğitimde dijital araçlar çok çeşitlilik göstermekte ve bu araçlar öğrenenlerin farklı öğrenme ihtiyaçlarına hızlı ve kolay bir şekilde cevap verebilmektedir. Bunun yanında bu araçların kullanımı, öğrenenlerin çeşitli öğrenme stillerine ve tercihlerine uygun bir öğrenme deneyimi yaşamalarına da olanak sağlamaktadır. Örneğin öğrenciler, bulut bilişim gibi teknolojik ürünler sayesinde verilere kolayca erişebilmekte ve bu verileri internet üzerinde güvenli bir şekilde depolayabilmekte ve paylaşabilmektedir. Aynı şekilde, podcast gibi araçlar da farklı konularda bilgi edinmek isteyen öğrencilere, uzmanların deneyimlerini dinleme ve uzmanlık alanları hakkında bilgi sahibi olma imkânı sunmaktadır. Bu çeşitlilik öğrenenlerin öğrenme ihtiyaçlarını daha etkili şekilde karşılamasını sağlamakta ve bunun yanısıra eğitim süreçlerini kişiselleştirerek daha verimli öğrenmelerin gerçekleşmesine imkân sunmaktadır.

Öğrencilerin günlük yaşamlarının bir parçası haline gelen dijital araçlarla öğrenme sürecinde, “formal öğrenme”lerle birlikte kuşkusuz informal öğrenmeler de gerçekleşmektedir. Dijital araçlarla informal öğrenme sürecinde, dijital araçları yalnızca okulda değil, aynı zamanda okul dışında da kullanma ihtiyacı doğmaktadır [1]. İhtiyaçlar doğrultusunda informal öğrenmeler gerçekleşmektedir. İnfomal öğrenme; bireylerin, ev, sokak, tiyatro, sinema, park, teneffüs saatleri gibi farklı yerlerde ve farklı sosyal ortamlar içerisinde plansız, programsız yaşam boyu devam eden ve kendiliğinden gerçekleşen öğrenmelerini kapsayan bir süreç olarak tanımlanmaktadır[2]. Lohman [3]’a göre informal öğrenme; “diğer insanlarla etkileşime girme, bilgi ve deneyimleri paylaşma, insanların davranışlarını izleyerek bu davranışların sonuçlarını gözlemleyerek kendine uyarılama” sürecidir. Schugurensky [4] informal öğrenmeyi üç farklı biçim de tanımlamaktadır: (1) öğrencinin öğrenme hedeflerinin farkında olarak kendi kendine kasıtlı öğrendiği kazanımlar (2) öğrenmenin dışında gerçekleştiği tesadüfi öğrenmeler ve (3) kişinin bilgiyi kasıtsız olarak edindiği ve hatta öğrendiklerinin farkında olmadığı kazanımlardır.

İnfomal eğitimin bir parçası olan informal öğrenme, bireyin çevresiyle etkileşimiyle yaşamının her aşamasında her an gerçekleşebilmektedir. Küçükahmet

[5]’e göre informal eğitimin özellikleri şu şekilde özetlenebilir:

- a)Plansız ve programsız gerçekleşir.
- b)İnsan yaşamının süresince kendiliğinden meydana gelir.
- c)İnsanın çevresiyle etkileşimi ölçüsünde daha fazla öğrenir.
- d)Zaman ve mekân sınırı yoktur.
- e)İnfomal eğitimde bireyler hem olumlu hem de olumsuz davranış edinebilmektedirler.

Dijital araçların yaşamın birçok alanında kullanımının artmasıyla formal ya da informal öğrenmeler de meydana gelmektedir. Zira bireyler, dijital araçları bir amaç ve plan doğrultusunda bilinçli olarak kullanarak formal öğrenmeler gerçekleştirirken; dijital araçları kullanım sürecinde farkında olmadan amacı dışında da birçok işlemi ya da konuyu öğrenebilmektedir. Eraut [6] informal öğrenmenin öğrenenlere dijital ortamlardan daha fazla bağımsızlık ve esneklik sunduğunu, Moore ve Klein [7] dijital araçların, informal öğrenmeyi kolaylaştırdığını ve zenginleştirdiğini ifade etmişlerdir. Dijital araçlar, internet, bilgisayarlar, akıllı telefonlar, tabletler, e-kitaplar, video ve ses kayıtları gibi birçok farklı formda sunulabilen materyallerdir. Bu araçlar, bilgiye erişimi ve öğrenmeyi kolaylaştırarak, herkesin kendi öğrenme ihtiyaçlarına uygun olarak öğrenmesine olanak tanır [8]. Sanal sınıflar, çevrimiçi tartışma forumları, oyunlar ve simülasyonlar gibi dijital araçlar, insanların etkileşim kurmalarını ve bilgi paylaşımında bulunmalarını sağlayarak öğrenme deneyimlerini zenginleştiren interaktif öğrenme ortamları sunar [9]. Bunun yanı sıra öğrencilerin okul dışı öğrenme etkinliklerinde, bir araştırma sürecinde ya da günlük hayatta internet, bilgisayar, telefon gibi dijital ortamlarda bulunduğu süreçte dijital araçlar sayesinde farkında olmadan birçok şeyi öğrenebilir. Bir ödevi hazırlarken farklı bilgisayar programlarını kullanabilme ya da sunu doyası hazırlayabilme ya da bir sosyal medya aracını kullanırken bir uygulamayı indirme, yükleme, kurma gibi. Bu noktada dijital çağ, bilişim çağı gibi ifade edilen içinde yaşadığımız 21. yüzyıldada dijital informal öğrenmelerin de arttığı söylenebilir. Dijital çağda öğrenme araçlarının bazıları; kullanıcı profili, SCORM kapsayıcılar, özel ve paylaşılan ek açıklama alanı, sohbet, forum, wiki, blog, e-port folyol, kavram haritası vb. platformlardır. Bunlar eğitim materyallerinin ve etkileşimli içeriklerin çevrimiçi olarak sunulduğu, öğrencilerin katılımını ve etkileşimini sağlar ve çeşitli öğrenme araçlarını içerir ve eğitim sürecini destekler [10]. Ayrıca forumlar, hatırlatmalar, mesaj sistemleri ve diğer etkileşimli araçlar, öğrencilerin ders içeriğini anlama, öğrenme sürecine katılma ve öğretmenlerle ve diğer öğrencilerle iletişim kurma konusunda destek olur, böylece öğrencilerin öğrenme potansiyelini geliştirir [11]. Bu süreçte birey informal öğrenmeler de gerçekleştirir.

Alanyazında, dijital araçlarla informal öğrenmeye yönelik yakın dönemde farklı bağlamlarda gerçekleştirilmiş

mobil öğrenmeye ilişkin görüşlerini incelemiştir. Araştırmada eğitim fakültesi öğrencilerinin genel mobil kullanım davranışlarının sosyal amaçlı ve eğlence amaçlı kullanım şeklinde daha ön plana çıktığı belirlenmiştir. Karadeniz [13] araştırmada, müzelerin dijital çalışmaları örnekler sunularak kültürel, doğal ve teknolojik mirasın tanıtılması, sürdürülebilirliğinin sağlanması ve dijital öğrenme ortamlarında kullanılması üzerinde durulmuştur. Mehrvarz vd. [14], öğrencilerin dijital yeterlilikleri ile akademik performansları arasındaki ilişkiyi incelemiş öğrencilerin dijital yetkinliğinin, dijital informal öğrenme ve akademik performanslarına olumlu katkı sağladığı ifade etmişlerdir. He vd. [15] gerçekleştirdikleri çalışmada Çinli üniversite öğrencilerinin dijital informal öğrenmeye yönelik davranışlarını araştırmak için ayrıştırılmış planlı davranış teorisine dayalı bir model geliştirmişlerdir. Bu modelde bilişsel öğrenme, üst bilişsel öğrenme, sosyal öğrenme ve motivasyon dijital yeterlilik olarak modele entegre edilmiştir. Araştırma sonuçlarının öğrencilerin dijital yeterlilik ve uyumluluk gibi motivasyon faktörlerinin önemini destekler nitelikte olduğu belirlenmiştir. Akife ve Göçer [16], podcastlerin Türkçe öğretiminde mobil öğrenme aracı olarak kullanılabilirliği üzerine yapılan araştırmadaki ana başlıklar; “Podcastlerin Mobil Öğrenme ve Zaman-Mekân Bağımsızlığına Katkısı”, “Podcastlerin Mobil Öğrenme Ekosistemindeki Rolü”, “Eğitimde Podcastlerin Kullanımı” ve “Türkçe Öğretiminde Podcastlerin Kullanımı ve Pratik Örnekler” başlıklar altında toplanmıştır. Bu araştırma, mobil öğrenme ve podcast kullanımı konularını ele alarak, podcastlerin sağladığı avantajlar ve sınırlılıkları incelemiş, eğitim ve Türkçe öğretimi alanlarında podcastlerin işlevsel rolünü desteklemek için teorik bir çerçeve sunmuş ve bu konuda örnek uygulamalar verilmiştir. Sadovets vd.[17] yükseköğretimde informal öğrenme alanında eğitimin dijital dönüşümü bağlamında oyunlaştırma üzerine araştırma gerçekleştirmişlerdir. Çalışmada, üniversite öğrencilerinin dijital dönüşümü bağlamında bir öğrenme tekniği olarak oyunlaştırmanın etkin bir şekilde öğrencilerinin ders çalışmaya yönelik motivasyonlarını artırdığı, yaratıcı ve profesyonel yeteneklerini ortaya çıkardığı sonucuna ulaşmışlardır. Xianhan vd. [18] öğretmenlerin teknolojiye entegre olma amacıyla farklı türde informal öğrenmelerinin ilişkilendirilmesi üzerine çalışma gerçekleştirmişlerdir. Çalışmada medya, meslektaş, paydaş ve öğrenci etkileşimleri yoluyla yansımaları üzerine bir anket çalışması yapmışlardır. Bu üç etkileşim arasından, öğretmen-öğrenci etkileşimi ve meslektaş etkileşimi yoluyla öğrenmenin öğretmenlerin teknolojiyi benimsemesiyle önemli ölçüde ilişkili olduğu sonucuna ulaşılmıştır. Araştırmada, öğrenci etkileşimi yoluyla öğrenme üzerinde teknoloji entegrasyonunun etkisine algılanan fayda ve öz-yeterliliğin eşit oranda aracılık etmiş olduğu belirlenmiştir.

Kandemir ve Ulusoy [19] bir yapı veya şehirdeki kültürel değerlerin artırılmış gerçeklik teknolojileri aracılığıyla sunulması amacıyla yapılan araştırmada, artırılmış

çalışmalar göze çarpmaktadır [12-21]. Örneğin Yokuş [12], yaptığı çalışmada eğitim fakültesi öğrencilerinin

gerçeklik teknolojilerinin kullanılması sayesinde ziyaretçilere bilinmeyen yerleri etkileşimli, keyifli ve öğretici bir şekilde keşfetme fırsatı sunulmuştur. Bu sayede, artırılmış gerçeklik (AR) teknolojilerinin kültürel miras bölgelerinde nasıl kullanıldığını açıklamak amacıyla uygulama örnekleri verilmiştir. Ng vd. [20], farklı yaş grubu ve cinsiyetteki öğrencilerin informal öğrenme ve sosyal ağ sitelerini kullanma durumları üzerine bir araştırma gerçekleştirmişlerdir. Araştırmada öğrencilerin sosyal ağları informal öğrenme amaçları bağlamında kullanmalarına yönelik algıları belirlenmeye çalışılmıştır. Araştırma bulguları öğrencilerin genel olarak sosyal ağları alternatif bir öğrenme ortamı olarak kabul ettiklerini; bununla birlikte cinsiyet arasındaki kullanım farkının bir dereceye kadar önemli olduğunu ve yaşa göre daha az etkili olduğunu ortaya çıkarmıştır. Panagiotidis vd. [21], dijital araçların yabancı dil öğreniminde motivasyon faktörü olarak rolünü incelemişlerdir. Araştırma kapsamında web araçları ve hizmetlerinin, dijital oyunların, mobil uygulamaların ve iletişim araçlarının kullanımı ile dil öğrenme bağlamındaki motivasyon arasındaki ilişki, çok çeşitli yaklaşımlarla ve çeşitli dil öğrenme çerçevesinde kapsamlı bir şekilde incelenmiştir. Yapılan araştırmalar incelendiğinde, Türkiye’de üniversite öğrencilerinin dijital araçlarla informal öğrenme yeterliliklerini belirleyen araştırmaların sınırlı sayıda olduğu, ayrıca üniversite öğrencilerine yönelik bu konuda geliştirilmiş bir ölçeğin bulunmadığı görülmüştür. Bu nedenle, üniversite öğrencilerinin dijital araçlarla informal öğrenme yeterliliklerini ve düzeylerini belirlemede kullanılacak bir ölçme aracı ihtiyacı olduğu anlaşılmıştır. Bu bağlamda araştırmada, üniversite öğrencilerinin dijital araçlarla informal öğrenme yeterliliklerini ölçmek için kullanılacak bir ölçme aracı geliştirilmesi amaçlanmıştır.

## 2. YÖNTEM (METHOD)

### 2.1. Araştırmanın Modeli (Research Model)

Araştırmada, karma yöntemlerden biri olan keşfedici sıralı desen tercih edilmiştir. Bu desen yeni bir olguyu ortaya çıkarmak veya bir sorunu anlamak amacıyla ilk aşamada nitel verilerin toplandığı ve daha sonra bu nitel verilerin doğruluğunu test etmek için nicel verilerin toplandığı ve analiz edildiği bir araştırma desendir. Araştırmacı, ilk aşamada sınırlı sayıda bir örneklem grubuyla görüşmeler yaparak, gözlemler gerçekleştirerek ve benzer nitel veri toplama tekniklerini kullanarak bilgi toplar. Bu nitel verileri analiz ederek ana temaları belirler. Daha sonra, bu nitel bulguları açıklamak ve genellikle daha büyük bir örneklem grubunu temsil eden nicel verilerle desteklemek amacıyla daha fazla katılımcıdan nicel veriler toplar. Son aşamada, araştırmacı, nicel verilere dayalı bir ölçme aracını tasarlar ve geliştirirken, aynı zamanda nitel bulguları daha da zenginleştirmek amacıyla bu aracı test eder [22-24] Üniversite öğrencilerinin dijital araçlarla informal öğrenme yeterliliklerini ölçmek için kullanılacak bir ölçme aracı geliştirilmesini amaçlayan bu araştırmada

da ilk olarak öğrencilerin dijital araçlarla informal öğrenme hakkında literatür taraması yapılmış ve kullanılan araçların ne olduğunu anlamak amacıyla konuyla ilgili sınırlı bir örneklem grubuyla nitel veriler toplanmıştır. Daha sonra literatüre ve bu nitel verilere dayanarak bir madde havuzu oluşturulmuş ve daha büyük bir örneklem üzerinde uygulanarak öğrencilerin dijital araçlarla informal öğrenme yeterliliklerini belirlemeye yönelik bir ölçek hazırlanmıştır. Bu nedenle bu araştırmada keşfedici sıralı bir araştırma deseni kullanılmıştır.

## 2.2. Çalışma Grubu (Working Group)

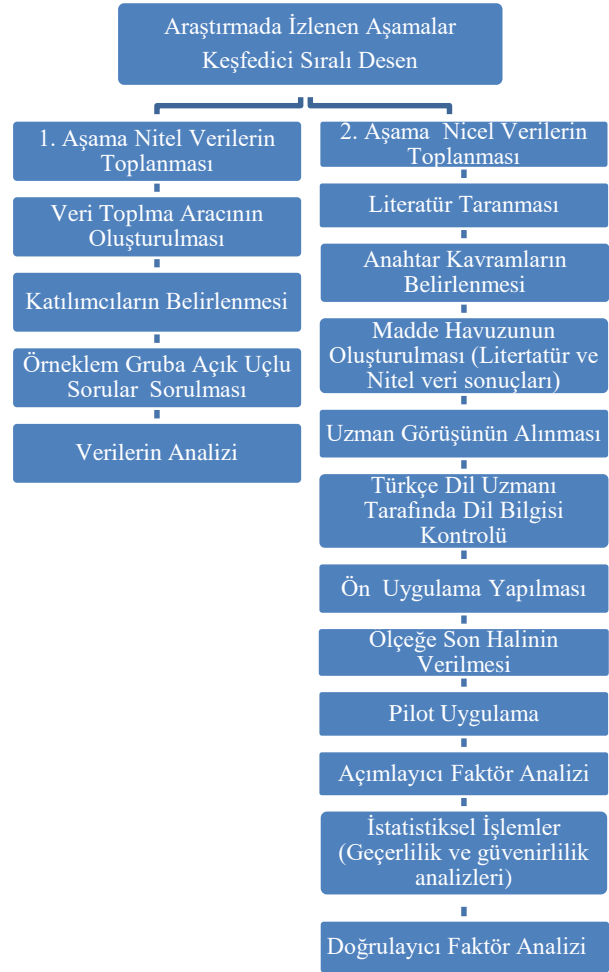
Araştırma, üç ana aşamadan oluştuğu için her aşamada farklı çalışma grupları ile farklı çalışmalar gerçekleştirilmiştir. Bu bağlamda, ilk aşamada; geliştirilecek olan ölçeğin uygulanacağı kitle olan üniversite öğrencilerinden açık uçlu sorularla görüş almak üzere 10 öğrenci çalışma grubuna alınmıştır. Ayrıca bu aşamada görüşlerini almak üzere çalışma grubuna 3 uzman dâhil edilmiştir.

İkinci aşamada yapılan araştırma çalışma grubu, 2022-2023 akademik yılında bir devlet üniversitesi'nde kayıtlı olan 117 erkek (%41) ve 170 kız (%59) olmak üzere toplam 287 üniversite öğrencisinden oluşmaktadır. Bu öğrencilerin dağılımı şu şekildedir: 16 doktora (%5,57), 33 lisansüstü (%11,49), 159 lisans (%55,40), 79 ön lisans (%27,52). Araştırmaya dâhil olan öğrencilerin çoğunluğu lisans düzeyinde eğitim alırken, diğerleri ise ön lisans, lisansüstü ve doktora seviyesinde eğitim almaktadır. Ayrıca öğrencilerin yaş ortalaması 24.6 olarak hesaplanmıştır. Ölçek geliştirme çalışmalarında, yeterli örneklem büyüklüğü için testin güvenilirliğini artırmak ve faktör analizinde madde sayısının 5 ila 10 katı kadar katılımcının olabileceği belirtilmektedir [25][26]. Bundan dolayı bu çalışmada madde sayısının beş katı civarında örneklem sayısı göz önüne alınmıştır. Bu bağlamda 65 maddelik bir ölçek için bu aşamada (Açımlayıcı Faktör Analizi Aşaması) 287 öğrenci çalışma grubuna alınmıştır.

Üçüncü aşamada, çalışma grubu, doğrulayıcı faktör analizi için toplamda 200 üniversite öğrencisinden 81'i erkek (%40,5), 119'u ise kız (%59,5) oluşmaktadır. Bu öğrencilerin dağılımı şu şekildedir: 13 doktora (%6,5), 32 lisansüstü (%16), 153 lisans (%76,5), 2 ön lisans (%1) olmak üzere toplam 200 üniversite öğrencisinden veri toplanmıştır. Ayrıca öğrencilerin yaş ortalaması 26 olarak hesaplanmıştır.

## 2.3. Ölçek Geliştirirken Yapılan İşlemler (Procedures Performed in Scale Development)

DAİÖYÖ'yü geliştirme aşamasında, aşağıdaki adımlar takip edilmiştir [27-29].



Şekil 1. Ölçek geliştirirken yapılan işlemler

Şekil 1' de sunulan aşamalar bağlamında, öncelikle amaç belirlenmiş ve dijital araçlarla informal öğrenme eğilimleri ile ilgili mevcut literatür taranarak teorik bir temel oluşturma amacıyla gerçekleştirilmiştir. Daha sonra ölçek geliştirmeye ilgili çalışmalar taranmıştır. Maddeler yazılırken dijital araçlarla informal öğrenmeye yönelik yapılan araştırmalar incelenmiş ve anahtar kavramlar belirlenmiştir. Ancak bu amaçla hazırlanan herhangi bir Türkçe ölçeğe rastlanmamıştır. Öncelikle, alanda bir literatür taraması yapılmış ve bu tarama sonucunda 7 adet açık uçlu soru hazırlanmıştır. Bu soruların kapsam geçerliliğini sağlamak amacıyla, ölçek hazırlama konusunda deneyime sahip olan üç uzmandan uzman görüşleri alındı. Uzmanların önerilerine dayalı olarak, bir soruda içerik düzenlemesi yapıldı ve iki soru formdan çıkarıldı. Uzman görüşlerinin ardından, son hali 5 açık uçlu soru içeren forma dönüştürüldü ve bu form 10 katılımcıya yönlendirilmiştir. Bu sorular dijital ortamların (Yenilikçi teknolojiler, anlık iletişim uygulamaları, mobil telefonlar, çevrimiçi tartışma forumları, bilgisayar oyunları ve sosyal medya) informal öğrenme deneyimlerinize etkileri nelerdir? Dijital informal öğrenme ortamlarını kullanmadan önce nasıl bir planlama

yapıyorsunuz? Nelere dikkat ediyorsunuz? İnfomal öğrenme ortamlarını kullanırken, süreçte neler yapıyorsunuz? Bu süreçte rolünüz nedir? Dijital araçlarla infomal öğrenme ortamlarını kullanmaya yönelik sizi motive eden şeyler/durumlar nelerdir? İnfomal öğrenme sürecinizde, en çok hangi çevrim içi teknolojik araçlar veya uygulamalar size yardımcı oluyor? Bu araçların her birinin nasıl yardımcı olduğunu açıklayınız, şeklindedir.

Ölçek maddeleri, beşli likert tipi olup "Tamamen Katılıyorum (5)", "Katılıyorum (4)", "Biraz Katılıyorum (3)", "Katılmıyorum (2)", "Tamamen Katılmıyorum (1)" seçenekleriyle derecelendirilmiştir. Ölçek maddeleri, "Tamamen Katılıyorum" ile "Tamamen Katılmıyorum" arasında değerlendirilir ve verilen cevapların puanı arttıkça, dijital araçlarla infomal öğrenme becerilerinin düzeyinin arttığı sonucuna varılmıştır. Böylece üniversite öğrencilerinin dijital araçlarla infomal öğrenme yeterliliklerini ifade edebilecek 65 tane madde yazılmıştır. Ölçek maddelerinin kapsam geçerliğini doğrulamak için, dijital öğrenmeler konusunda ve ölçek geliştirme konusunda çalışmaları olan üç uzmandan değerlendirme istenmiştir. Üç öğretim üyesi maddeleri kapsam, açıklık, doğruluk gibi hususlar bakımından değerlendirmiştir. Gerekli düzenlemeler yapıldıktan sonra bir öğretim görevlisi Türkçe dil uzmanı maddeleri dil bakımından incelemiş ve araştırmacı gerekli düzenlemeleri gerçekleştirmiştir. Dil ve anlatım yönünden problem olup olmadığını tespit edebilmek için 13 kişilik öğrenci grubuna pilot olarak uygulanmıştır. Neticede maddelerin anlaşılır olduğu sonucuna ulaşılmıştır. Daha sonra 65 maddelik dijital araçlarla infomal öğrenme ölçeği taslağı toplam 287 üniversite öğrencisine uygulanmıştır. Açımlayıcı faktör analizi uygulandıktan sonra doğrulayıcı faktör analizi (DFA) yapmak amacıyla 200 üniversite öğrencisinden veri toplanmış ve dfa analizleri yapılmıştır.

### 2.3. Verilerin Analizi (Analysing Data)

Üniversite öğrencilerinin infomal öğrenme sürecinde ne tür dijital araçlardan faydalandıklarını belirlemek amacıyla, geçerlik çalışmaları için açımlayıcı faktör analizi yöntemi kullanılarak maddeler arasındaki faktörleri tespit etmek hedeflenmiştir. Öncelikle dijital araçlarla infomal öğrenme sürecinin açımlayıcı faktör analizi ile uygun olup olmadığını kontrol etmek için Kaiser-Meyer-Olkin (KMO) katsayısına, örneklem büyüklüğünü tespit etmek için ise Barlett testi sonucuna bakılmıştır. Ardından ölçeğin faktör yapısını araştırmak için temel bileşenler analizi kullanılmış ve birbirleriyle yüksek ilişkiye sahip maddelerin bir araya toplanması için Varimax Döndürme yöntemi işleme alınmıştır [30]. Ölçeğin güvenilirliğini değerlendirmek amacıyla Cronbach alfa güvenilirlik katsayısı, madde toplamı ve madde kalanı korelasyon analizlerinin yanı sıra madde alt-üst grup ortalamaları arasındaki farkları değerlendirmek için iç tutarlılık ölçütüne dayalı t testi kullanılarak gerçekleştirilmiştir. Son olarak, gerçek verilerle faktörler arasındaki uyumu değerlendirmek ve bu faktörlerin modeli yeterince açıklayıp açıklamadığını incelemek için birinci ve ikinci düzey doğrulayıcı faktör analizi (DFA) yapılmıştır [31][32].

### 3. BULGULAR (FINDING)

#### 3.1. Geçerlik Çalışmalarına İlişkin Bulgular (Findings Related to Validity Studies)

Dijital araçlarla infomal öğrenme yeterlilik ölçeği (DAİÖYÖ)'nin yapı geçerliği için madde aralarındaki faktörleri tanımlamak amacıyla faktör analizi (AFA) yöntemi kullanılmıştır. AFA birçok değişkenli aynı yapıyı ölçen, daha az sayıda ve belirgin özelliklere sahip anlamlı değişkenleri belirlemek için çoklu değişken analizi yapan istatistiksel bir yöntemdir [33]. Dijital araçlarla infomal öğrenme yeterlilik ölçeğinin faktör analizi sürecinde, aynı faktörde güçlü ilişkilere sahip maddeleri toplamak için Varimax Döndürme işlemi yapılmıştır. Varimax döndürme, faktörler arasındaki ilişkileri daha açık bir şekilde ortaya koymak için kullanılan istatistiksel bir tekniktir. Bu adım, temel bileşen analizi sonuçlarının veri koordinatlarını yeniden düzenlemeyi içerir. Bu düzenleme veya döndürme işlemi, veri noktaları arasındaki ortak varyansı en üst düzeye çıkarmayı hedefler. Bu yöntem sayesinde, verilerin her bir temel bileşen ile ne şekilde ilişkilendiği net bir biçimde gösterilir ve bu şekilde paylaşılan varyans en üst düzeye çıkarılmış sonuçlar elde edilir [34]. Temel bileşenler analizi (Principal Components Analysis), başlangıçta karmaşık olan bir veri setini daha az sayıda ve daha anlamlı bağımsız değişkenler içeren bir veri kümesine dönüştüren bir istatistiksel yöntemdir. Bu analiz, orijinal veri setindeki karmaşıklığı azaltarak verileri daha anlaşılır ve daha özlü bir şekilde ifade etmeye yardımcı olur [35]. Bu bağlamda maddelerden faktör elde edebilmek için temel bileşenler analizi (principal components) uygulanmıştır. Faktör sayısını belirleme aşamasında, Kaiser'in [36] özdeğer  $\geq 1$  kuralı göz önünde bulundurulmuştur. Bu kurala göre, faktörlerin özdeğerleri 1 veya daha büyük olan maddeler kullanılmıştır. Verilerin faktör analizi için uygunluğunu değerlendirmek için hesaplanan Kaiser-Meyer-Olkin (KMO) değeri 0.915 olarak bulunmuştur. KMO değeri, verilerin faktör analizi için uygun bir yapıya sahip olduğunu göstermektedir. Barlett testi (Bartlett Bütünlük Testi) sonucu, .00 ( $p < .05$ ) olarak hesaplandığından verilerin normal dağılımda olduğu ve istatistiksel olarak anlamlı sonuçlar elde edildiğini göstermektedir. Analiz sonuçları Tablo 1' de görülmektedir.

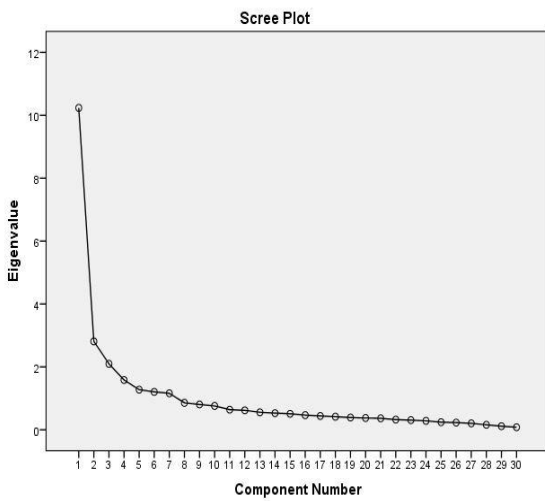
Tablo 1. DAİÖYÖ'ye ilişkin kmo ve bartlet testi sonuçları(kmo and bartlet's test results for elses)

KMO Bartlet testi	-	.915
	<i>Ki-kare</i>	11983,229
	<i>df</i>	2080
	<i>p</i>	,000

#### 3.2. Faktör Sayısının Belirlenmesi (Determination of the Number of Factors)

65 madde ile gerçekleştirilen faktör analizi sonucunda, 13 faktörün toplam varyansının %65.758 olduğu tespit edilmiştir. Varimax Döndürme tekniği, bir faktör içinde birbirine yüksek ilişkili olan maddeleri belirlemek

amacıyla kullanılmıştır. Bu yöntemle faktör yükleri arasındaki farklar, birbirine paralel olan, 0.10'dan az ve 0.30'dan düşük olan (6,9,14,15,17,18,20,26,28,29,30,31,32,33,34,38,41,42,43,44,45,46,47,48,51,52,53,54,55,60,61,62,63,64,65) 35 madde ölçekten tek tek çıkarılmış ve analizler tekrar yapılmış ve en sonunda faktör sayısı dört, toplam varyans %55.747 bulunmuştur. Faktör sayısına göre maddeler 4 grupta toplanmış 1.Grup (35,36,37,39,40,56,57,58,59) 2.Grup (21,22,23,24,25,27,49,50) 3.Grup (7,8,10,11,12,13,16,19) 4.Grup (1,2,3,4,5) şeklindedir. Açımlayıcı faktör analizinde baskın faktörleri belirlemek için Cattell [37] tarafından geliştirilen yamaç eğim grafiğinden yararlanılmıştır. Ana kırılma noktalarına bağlı olarak, ölçeğin faktör sayısı Şekil 2 'de görüldüğü gibi 4 olarak belirlenmiştir.



Şekil 2. Dijital araçlarla informal öğrenme ölçeği yamaç birikinti grafiği (Scale of informal learning with digital tools slope accumulation graph)

Analiz sonuçlarına göre, toplam varyansın %55.747 olduğu ve 30 maddeden meydana gelen dört faktörlü bir ölçme aracı elde edilmiştir. Bu ölçme aracının öz değerleri, varyans ve toplam varyanslarına ait yüzdeleri Tablo 2'de sunulmuştur.

Tablo 2. DAİÖYÖ'nün toplam varyans değerleri (total variance values of else)

Faktör	Başlangıç özdeğerleri			Karesi alınan yüklerin toplam çıkarımı			Karesi alınan yüklerin döndürme toplamı		
	Top.	%Varyans	%Kümüls	Top.	%Varyans	%Kümüls	Top.	%Varyans	%Kümüls
1	10,24	34,12	34,116	10,24	34,12	34,12	4,84	16,142	16,142
2	2,81	9,37	43,484	2,81	9,37	43,48	4,44	14,805	30,947
3	2,09	6,99	50,473	2,09	6,99	50,47	4,16	13,857	44,804
4	1,58	5,27	55,747	1,58	5,27	55,75	3,28	10,943	55,747

Faktör analizi sonucunda Tablo 2 'de görüldüğü gibi, dört faktör tespit edilmiştir. Faktörlerin varyans açıklama

yüzdeleri sırasıyla %16.142, %14.805, %13.857 ve %10.943 olarak görülmektedir. Bu dört faktörün, toplam varyansı %55.747 olarak açıklanmaktadır.

Çoklu değişkenli desenlerde, genellikle kabul edilebilir bir seviyede, varyansın %40 ile %60 arasında açıklandığı gözlenir [38]. Bu sonuçlara göre DAİÖYÖ'nün toplam varyans değeri (%55.747), yeterli seviyededir ve kabul edilebilir bir düzeydedir. Sonuç olarak, dijital araçlarla gerçekleştirilen informal öğrenme ölçeği üzerindeki madde faktör yükleri dağılımı, Tablo 3'te gösterildiği gibidir.

Tablo 3. DAİÖYÖ'nün döndürülmüş bileşenlerine ait değerler (Values belonging to the rotated components of the ILCS-DT)

Maddeler	Faktörler			
	1	2	3	4
M35	,631			
M36	,733			
M37	,700			
M39	,787			
M40	,686			
M56	,609			
M57	,637			
M58	,691			
M59	,546			
M21		,810		
M22		,844		
M23		,823		
M24		,817		
M25		,331		
M27		,420		
M49		,576		
M50		,494		
M7			,664	
M8			,589	
M10			,598	
M11			,756	
M12			,683	
M13			,505	
M16			,573	
M19			,666	
M1				,723
M2				,797
M3				,701
M4				,711
M5				,691

Faktör analizi uygulanırken, faktörlerin birbirinden bağımsız olmasını sağlamak ve daha anlaşılır ve anlamlı yorumlar yapabilmek amacıyla eksen döndürmesi (factor rotation) yöntemi kullanılır. Bu yöntemle, bir madde bir faktördeki yükü arttığında, diğer faktördeki yükü azalmaktadır. Varimax Döndürme yöntemi kullanılarak, faktörler arasında yüksek ilişkiye sahip olan maddeleri bir araya toplanmıştır. Bu yöntemle, faktörlerin yapısı daha anlaşılır hale getirilmiş ve maddelerin faktörler üzerindeki yükleri daha bağımsız hale getirilmiştir [30]. Bu bağlamda, Tablo 3'te verildiği üzere, 30 maddeye ait faktör yükleri, .30'dan büyüktür ve .33 - .84 arasındadır. Bu değerler, 30 maddenin ölçekte kullanılmak için uygun niteliklere sahip olduğunu göstermektedir.

### 3.3. Güvenirlilik Çalışmalarına İlişkin Bulgular (Findings Related to Reliability Studies)



Ölçeğin iç tutarlılığını değerlendirmek için, ölçekteki maddelerin veya alt boyutların toplam puanları ile ölçeğin toplam puanları arasındaki korelasyon katsayıları kullanılır. Madde-toplam puan korelasyonları, her bir maddenin puanı ile ölçeğin toplam puanı arasındaki ilişkiyi ölçen ve ölçeğin iç tutarlılığını gösteren bir ölçüdür. Pozitif ve 0.30'dan yüksek olan madde-toplam korelasyonları, bu maddelerin ayırt edici olduğunu göstermektedir[39]. Maddelerin her biri için, madde puanları ile ölçeğin toplam puanı arasındaki korelasyon katsayıları hesaplanarak, ölçeğin tutarlılık derecesi incelenmiştir [26]. Madde analizi için başka bir yaklaşım, testin toplam puanlarına göre oluşturulan alt %27 ve üst %27 grupların, madde puan ortalamaları arasındaki farkları bağımsız örneklem t-testi ile incelemektir. Ölçeğin iç tutarlılığını belirten bir gösterge olarak kabul edilen gruplar arasındaki anlamlı farklar, ölçeğin güvenilirliği hakkında bilgi verir [39]. Bu sebeple %27'lik alt ve üst gruplar arasındaki madde puanlarında gözlenen farklar için, t değerleri hesaplanmış ve analiz edilmiştir. Alt ve üst gruplar arasında maddelerin ortalama puanlarındaki farkın anlamlılığını değerlendirmek amacıyla yapılan t-testine ait sonuçlar, aşağıdaki Tablo 4'te verilmiştir.

Tablo 4. DAİÖYÖ'ye ilişkin t-testi sonuçları (t-test results related to ILCS-DT)

M. No	t değ.	p değ.	r değ.	M. No	t değ.	p değ.	r değ.
M1	9,668	0,00	,542	M23	12,939	0,00	,660
M2	8,811	0,00	,529	M24	11,782	0,00	,617
M3	8,360	0,00	,547	M25	8,127	0,00	,442
M4	9,203	0,00	,494	M27	7,578	0,00	,444
M5	5,728	0,00	,370	M35	9,003	0,00	,482
M7	12,635	0,00	,654	M36	9,961	0,00	,529
M8	10,599	0,00	,596	M37	10,916	0,00	,568
M10	11,465	0,00	,589	M39	10,511	0,00	,591
M11	10,170	0,00	,595	M40	9,617	0,00	,507
M12	11,326	0,00	,612	M49	6,826	0,00	,376
M13	10,815	0,00	,594	M50	9,278	0,00	,526
M16	10,966	0,00	,613	M56	11,467	0,00	,594
M19	9,858	0,00	,571	M57	9,905	0,00	,546
M21	11,701	0,00	,608	M58	6,907	0,00	,417
M22	12,690	0,00	,622	M59	4,646	0,00	,254

Tablo 4'te verilen analiz sonuçlarına göre, her bir madde için üst ve alt gruplar arasındaki puanlar arasında istatistiksel açıdan anlamlı bir farklılık olduğu bulunmuştur ( $p=0,00$  olup  $p<0,05$ ). DAİÖYÖ'nün güvenilirlik düzeyini belirlemek amacıyla, iç tutarlılık güvenilirlik katsayılarını hesaplamak için Cronbach Alfa formülü kullanılmıştır. Tablo 5'de güvenilirlik analizi sonucu verilmiştir. Dijital araçlarla informal öğrenme yeterlilik ölçeğine ait iç tutarlılık kat sayısı 0,927 olarak tespit edilmiştir. Faktörlerin güvenilirliğini değerlendirmek için yapılan Cronbach Alfa ( $\alpha$ ) analizi sonuçlarına göre, tüm faktörlerin yüksek güvenilirlik düzeyine sahip olduğu gözlenmektedir.

Tablo 5. DAİÖYÖ'nün güvenilirlik analizi sonuçları(Reliability analysis results of ILCS-DT)

Faktörler	Madde Sayısı	Cronbach alfa ( $\alpha$ )
<i>Faktör1</i>	9	.878
<i>Faktör2</i>	8	.855
<i>Faktör3</i>	8	.877
<i>Faktör4</i>	5	.836
<i>Toplam</i>	30	.927

Tablo 5'te dijital araçlarla informal öğrenme ölçeği için Cronbach alfa güvenilirlik katsayısı .927; faktör 1 .878; faktör 2 .855 faktör 3 .877 ve faktör 4 .836 olarak verilmiştir. Güvenirlik katsayısı (Cronbach  $\alpha$ ), maddelerin puanları ile toplam test puanı arasındaki tutarlılığı ölçen bir ölçüdür. Bu katsayının değerleri, eğer (0.00-0.40) aralığında ise, ölçeğin güvenilir olmadığını; (0.40-0.60) aralığında ise, düşük güvenilirliğe sahip olduğunu; (0.60-0.80) aralığında ise, oldukça güvenilir olduğunu ve (0.80-1.00) aralığında ise, yüksek derecede güvenilir olduğunu ifade etmektedir[40][41]. Buna göre ölçeğin dört faktörünün de yüksek derecede güvenilir olduğu görülmektedir.

Tablo 6. Faktörler arası pearson korelasyon analizi sonuçları (Pearson correlation analysis results between factors)

	Faktör1	Faktör2	Faktör3	Faktör4	Toplam
<i>Faktör1</i>	1	.406(**)	.558(**)	.381(**)	.748(**)
	287	.000	.000	.000	.000
		287	287	287	287
<i>Faktör2</i>		1	.614(**)	.427(**)	.820(**)
			.000	.000	.000
			287	287	287
<i>Faktör3</i>			1	.523(**)	.862(**)
				.000	.000
				287	287
<i>Faktör4</i>				1	.692(**)
					.000
					287
<i>Toplam</i>					

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

Tablo 6'da belirtildiği üzere, faktörler arasındaki ilişkileri değerlendirmek için gerçekleştirilen pearson korelasyon analizi sonuçları, anlamlı bir şekilde pozitif yönde ilişkilerin olduğunu göstermektedir. Bu bağlamda, faktör1 ile faktör3 arasında ( $r=.558$ ;  $p<0,001$ ) en yüksek ilişki bulunurken, aynı zamanda faktör1 ile faktör4 arasında ( $r=.381$ ;  $p<0,01$ ) daha düşük bir ilişki tespit edilmektedir. Ek olarak, toplam puan ile faktör1 ( $r=.748$ ), faktör2 ( $r=.820$ ), faktör3 ( $r=.862$ ) ve faktör4 ( $r=.692$ ) puanları arasında yüksek derecede pozitif ve anlamlı bir ilişkinin olduğu ve tüm faktörlerin aynı yapı içinde olduğu görülmektedir.

Gerçekleştirilen geçerlik ve güvenilirlik analizleri neticesinde dijital araçlarla informal öğrenme ölçeğinde kalan 30 maddenin içeriklerine göre dört faktörlü (dijital öz yeterlilik, teknolojik isteklilik, bilişsel yeterlilik ve etkin kullanım) bir ölçme aracı olduğu ortaya çıkmıştır. DAİÖYÖ için yapılan açımlayıcı faktör analizi sonuçlarından sonra 1. ve 2. düzey doğrulayıcı faktör analiz sonuçları aşağıda sunulmuştur.

### 3.3. Doğrulayıcı Faktör Analizlerine (DFA) Bulgular (Confirmatory Factor Analyses (CFA) Findings)

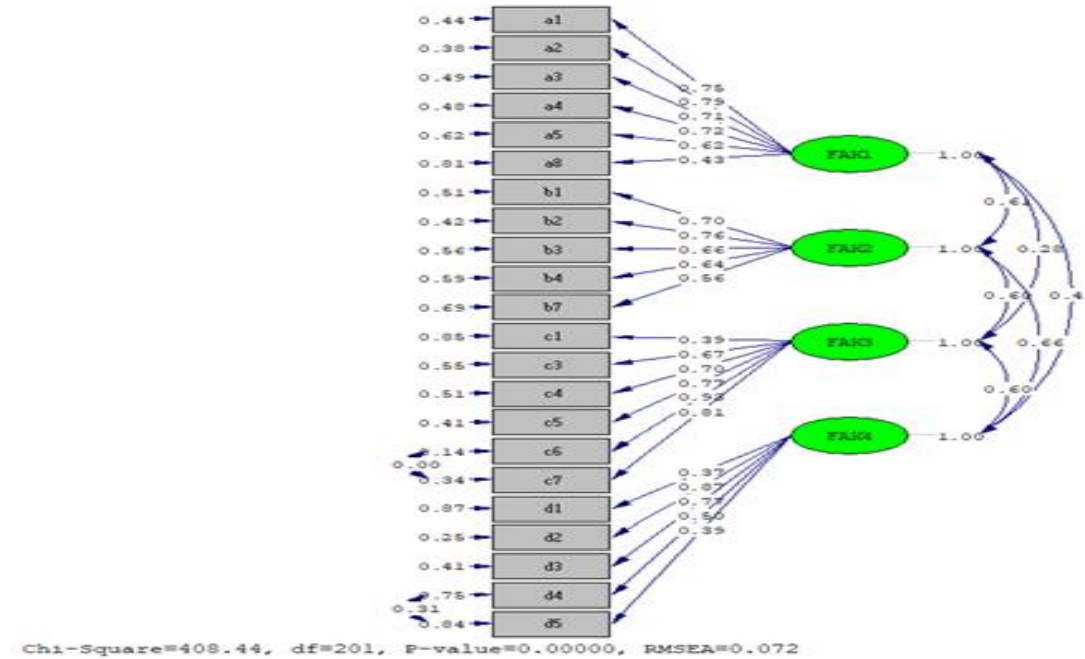
#### 3.3.1 Birinci Sıralı Doğrulayıcı Faktör Analizi (First Order Confirmatory Factor Analysis)

Doğrulayıcı Faktör Analizi (DFA), araştırmacıların mevcut teorik çerçeveyi desteklemek ve önceki bulguları doğrulamak amacıyla kullanılır [42]. Amacı, oluşturulan modelin teorik yapıyı iyi bir şekilde temsil edip etmediğini ve gözlenen verilerle uyumlu olup olmadığını değerlendirmektir. Eğer model verilerle uyumluysa, bu teorik yapıyı destekleyen bir bulgu olarak kabul edilir ve yapı geçerli kabul edilir [43]. Bu bağlamda DAİÖYÖ'ye ait DFA sonucunda ulaşılan dört boyuta göre bir model oluşturulmuş ve bu model DFA ile test edilmiştir.

DFA aşamasında, dijital araçlarla informal öğrenme yeterlilik ölçeğine ilişkin uyum indeksleri incelenmiştir. Oluşturulan modelin teori ile uyumlu olup olmadığına karar verirken, genel kabul görmüş standart uyum indekslerinden bazıları şunlardır:  $\chi^2$ ,  $\chi^2/sd$ , GFI, CFI, TLI, RMSEA, RMR ve SRMR. Uyum indeksleri, modelin

veriyle ne kadar iyi eşleştiğini ve ne kadar iyi uyum sağladığını değerlendirmek için kullanılır [44]. Bu bağlamda informal öğrenme yeterlilik ölçeğine ait ilişkin uyum indeksleri; CMIN=408,44; DF=201;  $p<0,00$ ; CMIN/DF=2,03; RMSEA=0,072; CFI=0,95; GFI=0,84; IFI =0,95; RFI= 0,90; SRMR= 0,063 olarak belirtilen sınırlar içerisinde bulunmuştur. Ki-kare/serbestlik derecesi ( $\chi^2/sd$ ) değeri, 5'ten küçük olduğunda, modelin iyi bir uyum sağladığı, 3'ten daha küçük olduğunda ise, modelin çok iyi bir uyuma sahip olduğu kabul edilir.

Modelin hata (uyumsuzluk) indeksi olan RMSEA için kabul edilebilir bir uyumu göstermesi için 0.08 değeri yeterlidir. Ancak, bu değer 0.05'ten daha küçükse, mükemmel bir uyumu ifade etmektedir. Ayrıca, SRMR değerlerinin de 0.08' den düşük olması kabul edilebilir uyumu ifade eder. Bu değer, modelin veriyle iyi bir uyum sergilediğini gösterir. Yani, SRMR değeri 0.08 'in altındaysa, modelin veriyle kabul edilebilir uyum düzeyine sahip olduğu söylenebilir. CFI ve IFI değerleri, 0.95 ile 1.00 aralığında olduğunda, mükemmel bir uyumu gösterirken, 0.90 ile 0.95 arasında olduğunda ise kabul edilebilir bir uyumu ifade etmektedir. [38][45-49]. Bu sonuçlara göre; dijital araçlarla informal öğrenme yeterlilik ölçeğinin Ki-kare/serbestlik derecesi ( $\chi^2/sd$ ) değerinin iyi bir uyuma, RMSEA, SRMR, CFI değerlerinin kabul edilebilir uyuma sahip olduğu, görülmektedir. Dijital araçlarla informal öğrenme yeterlilik ölçeğine ilişkin oluşturulan modele ait Path Diyagramı aşağıda Şekil 2'deki gibidir.

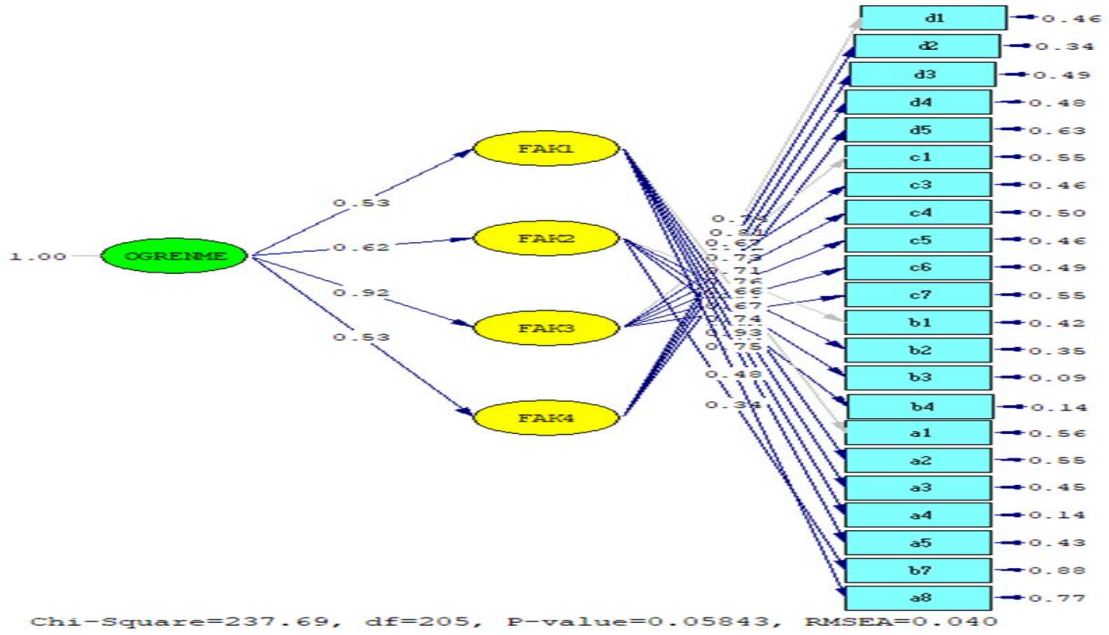


Şekil 2. DAİÖYÖ birinci düzey path diyagramı (First level path diagram of ILCS-DT)

### 3.3.2 İkinci Sıralı DFA ve Faktöriyel Geçerlik (Second Order CFA and Factorial Validity)

Birinci düzeyde alt boyutlarda bulunan maddelerin doğrulaması yapıldıktan sonra, ikinci seviyedeki sıralı DFA sonuçlarına bakıldığında, ölçekte bulunan 22 madde için faktör yüklerinin 0.34 ile 0.93 arasında değiştiği gözlenmektedir. Grafikteki t değerleri analiz edildi ve

anlamlı bir uyumsuz değer bulunamadı. Sonuç olarak,  $\chi^2$  değerinin düşük olması ve p anlamlılık düzeyinin 0.05'ten büyük olması, uygun bir sonuç olarak kabul edilebilir [50]. Bu da verilerin ölçekle iyi bir uyum sağladığının göstergesidir. Dijital araçlarla informal öğrenme ölçeği belirlemek için ikinci düzey DFA diyagramı şekil 3 deki gibidir.



Şekil 3. DAİÖYÖ ikinci düzey path diyagramı ( ILCS-DT second level path diagram)

Şekil 3' deki gibi ikinci düzey doğrulayıcı faktör analizi sonucunda Chi Square/df değerinin  $237.69 / 205 = 1.15$ , RMSEA değeri 0,040 olması mükemmel bir uyuma sahip olduğunu göstermektedir. Bu nedenle, doğrulayıcı faktör analiz sonuçlarına ait uyum değerleri Tablo 7'de gösterildiği gibidir.

Tablo 7. DAİÖYÖ ait uyum indeks sonuçları (Fit index results for the scale of ILCS-DT)

Ölçüt	Mükemmel uyum değeri	Kabul edilebilir uyum değeri	Ölçek uyum değerleri
$\chi^2/df$	$0 \leq \chi^2/df \leq 3$	$3 \leq \chi^2/df \leq 5$	1,15
RMSEA	$0 \leq RMSEA \leq 0,05$	$0,05 \leq RMSEA \leq 0,08$	0,04
NNFI	$\geq 0,95$	$\geq 0,90$	0,97
AGFI	$0,90 \leq AGFI \leq 1$	$0,85 \leq AGFI$	0,78
CFI	$\geq 0,95$	$CFI \geq 0,90$	0,98

NFI	$0,95 \leq NFI \leq 1$	$NFI \geq 0,90$	0,97
RFI	$\geq 0,95$	$\geq 0,90$	0,89
IFI	$\leq 0,05$	$\geq 0,90$	0,98
SRMR	$0,00 \leq SRMR \leq 0,05$	$0,05 \leq SRMR \leq 0,08$	0,061
PNFI	$\geq 0,95$	$\geq 0,50$	0,80
PGFI	$\geq 0,95$	$\geq 0,50$	0,67

Kaynak: [38][45-49].

Tablo 7' deki gibi doğrulayıcı faktör analizinden elde edilen uyum değerleri ( $\chi^2/df$ , RMSEA, NNFI, AGFI, CFI, NFI, RFI, IFI, SRMR, PNFI ve PGFI) yapılan karşılaştırmalarda, ölçek maddelerinin dört alt boyutu ve ölçeğin geneli arasındaki ilişkiyi doğrulayan bir model tespit edilmiştir.

### 3.3.3 Kalan Maddeler İçin Güvenirlilik Analizi (Reliability Analysis for the Remaining Items)

DAİÖYÖ'nün DFA aşamasında madde yükleri düşük olmasından dolayı 8 madde ölçekten çıkarılmıştır. Bu bağlamda, geriye kalan maddelerin güvenirliliğini değerlendirmek için Cronbach Alpha güvenirlilik analizi gerçekleştirilmiş ve ölçeğin güvenirliliği 0.898 olarak belirlenmiştir. Bunun yanı sıra, dijital öz yeterlilik alt boyutunun güvenirliliği %0.855, teknolojik isteklilik alt boyutunun güvenirliliği %0.867, bilişsel yeterlilik alt boyutunun güvenirliliği %0.852 ve etkili kullanım alt boyutunun güvenirliliği %0.839 olarak tespit edilmiştir. Bu bağlamda, dijital araçlarla informal öğrenme yeterlilik ölçeği dahil olmak üzere alt boyutlarıyla birlikte güvenilir bir ölçme aracı olarak kabul edilmiştir. Bu nedenle, Cronbach  $\alpha$ , maddelerin toplam test puanlarıyla uyumunu ölçer. Bu katsayının değerleri, (0.00 - 0.40) arasında ise ölçeğin güvenilir olmadığını, (0.40 - 0.60) arasında ise düşük güvenirliliğe sahip olduğunu, (0.60 - 0.80) arasında olduğunda oldukça güvenilir olduğunu ve (0.80 - 1.00) arasında ise yüksek derecede güvenilir olduğunu ifade etmektedir [40] [41].

Son olarak 200 üniversite öğrencisi üzerinde gerçekleştirilen pilot bir uygulama için kullanılan verilerin faktör analizine uygunluğunu yeniden değerlendirmek amacıyla hesaplanan Kaiser Meyer Olkin (KMO) değeri 0.890 bulunmuştur. Bu değer, verilerin faktör analizinde kullanılmak üzere uygun bir yapıya sahip olduğunu gösterdiği söylenebilir. Ayrıca Barlett testi (Bartlett Bütünlük Testi) sonucu, anlamlılık değeri .00 ( $p < .05$ ) olarak hesaplanmış, bu da verilerin normal dağılımda olduğu ve istatistiksel olarak anlamlı sonuçlar elde edildiğini göstermektedir.

## 4. SONUÇ (CONCLUSION)

Araştırmanın temel amacı, üniversite öğrencilerinin informal öğrenme süreçlerinde kullandıkları dijital araçları belirlemeyi ve kullanımın sonuçlarını gözlemlemeyi amaçlayan geçerli ve güvenilir bir ölçek geliştirmektir. Bu amaç doğrultusunda ölçek geliştirme sürecinde izlenen adımlar sonucunda, dijital araçlarla informal öğrenme yeterliliğini 5'li likert olarak "Dijital Araçlarla İnfomal Öğrenme Yeterlilik Ölçeği" geliştirilmiştir. Bu kapsamda; literatür araştırması ve yarı yapılandırılmış görüşmelerle elde edilen verilerin analizi sonucunda, 65 maddeyi içeren bir madde havuzu oluşturulmuştur. 287 üniversite öğrencisine uygulanan bu ölçek çalışması sonucunda, faktör analizi ve DFA analizi sonuçlarına göre faktör yükleri düşük olan 43 madde ölçekten çıkarılmıştır. Sonuç olarak, dört farklı boyutu içeren, 22 olumlu maddeden oluşan bir ölçek geliştirilmiştir. Geriye kalan maddelerin güvenirliliğini değerlendirmek için Cronbach Alpha güvenirlilik analizi yapılmış ve ölçeğin güvenirliliği 0.898 olarak hesaplanmıştır. Ölçeğin alt boyutları, sırasıyla dijital öz yeterlilik, teknolojik isteklilik, bilişsel yeterlilik ve etkin kullanım olarak tanımlanmıştır. Birinci faktör (6 madde): Bu kategorideki maddeler, bireylerin dijital teknolojileri etkin bir şekilde kullanabilme yeteneğini

temsil ettiği için bu boyut "dijital öz yeterlilik" olarak adlandırılmıştır. Bu alt boyut, dijital eğitim programlarının etkisini değerlendirmek için kullanılabilir. Öğrencilerin dijital öz yeterlilikleri ve teknolojik isteklilikleri ölçülerek, eğitim programlarının bu becerileri geliştirmekte ne kadar etkili olduğu değerlendirilebilir. İkinci faktör (5 madde): yeni teknolojilere hızla uyum sağlama veya mevcut teknolojileri kullanma konusundaki tutumlarına "teknolojik isteklilik" olarak adlandırılmıştır. Katılımcılara dijital araçları ne sıklıkla kullandıklarını, bu araçlarla ne kadar rahat olduklarını ve dijital teknolojilere karşı duydukları güveni değerlendiren anketler veya öz-değerlendirme formları sunabiliriz. Bu, katılımcıların kendi teknolojik isteklilik düzeylerini belirlemelerine olanak tanır. Üçüncü faktör (6 madde), bilgiyi algılama, anlama, işleme, hatırlama, problem çözme, karar verme ve yargılama gibi bilişsel yeteneklerini ifade ettiği için bu boyut: "bilişsel yeterlilik" olarak adlandırılmıştır. Eğitimciler, dijital öz yeterliliği artırmayı amaçlayan eğitim programlarını tasarlarken bu alt boyutları kullanabilirler. Bu, programların öğrencilere daha iyi teknolojik beceriler kazandırmak için nasıl geliştirilebileceğini anlamalarına yardımcı olabilir. Dördüncü faktör (5 madde): dijital teknolojileri bilinçli, verimli ve güvenli bir şekilde kullanılabilmesi adına "etkin kullanım" olarak adlandırılmıştır. Bu alt boyutlar, akademik çalışmalarda dijital yeteneklerin etkisini araştırmak için kullanılabilir. Ölçeğin alt boyutların belirlenmesinden sonra hesaplanan DFA sonuçlarına göre,  $X^2/df$  değerinin 1.15 olması mükemmel uyum değeri göstermektedir. Ayrıca, RMSEA değerinin 0.040 olması da mükemmel uyum değeri bir uyuma işaret etmektedir. Diğer uyum değerleri göz önüne alındığında, modelin doğrulandığı sonucuna ulaşılabilir.

Analizlerin sonuçlarına göre, DAİÖYÖ'nün geçerli ve güvenilir bir ölçüm aracı olduğu bulunmuştur. Dört alt boyut, 22 madden oluşan DAİÖYÖ'den en düşük 22, en yüksek 110 puan alınabilir ve tüm maddeler olumlu ifadelerden oluşmaktadır. Yüksek puanlar, öğrencilerin dijital araçlarla informal öğrenme yeterliliklerinin yüksek düzeyde olduğunu gösterirken, düşük puanlar da bu yeterliliklerin düşük düzeyde olduğunu ifade etmektedir.

Türkçe alan yazın kapsamında dijital araçlarla informal öğrenme yeterliliği üzerine geliştirilmiş ya da uluslararası literatürden uyarlanmış herhangi bir ölçeğe ulaşılamamıştır. Ancak uluslararası alanyazında He ve Zhu [51] 'nin üniversite öğrencilerine yönelik informal öğrenmede kişisel faktörler ve dijitalin etkileri bağlamında geliştirdikleri ölçekte "dijital informal öğrenmeye karşı tutum, dijital yeterlilik, dijital informal öğrenme, kişisel yenilikçilik" alt boyutlarının yer aldığı görülmektedir. Başka bir çalışmada Mehrvarz vd. [14] tarafından COVID-19 salgını sırasında öğrencilerin dijital yeterlilikleri ile akademik katılımları arasındaki ilişkide informal dijital öğrenmenin rolü üzerine geliştirilen ölçeğin "dijital informal öğrenme, akademik katılım, dijital yetkinlik" alt boyutlarından oluşmuş olduğu görülmüştür. Diğer bir çalışmada Liu vd. [52]. Dijital informal İngilizce öğrenimi ve kültürlerarası yetkinlik üzerine geliştirdikleri ölçekte

“Kendini Tanıma, Başkalarını Tanıma, Tutum, Farkındalık ve Kültürlerarası İletişim Becerilerinin” Çin yükseköğretim bağlamında kültürlere arası yetkinliği ölçmenin temel boyutları olarak öne çıktığını göstermiştir. Yakın dönemde benzer şekilde gerçekleştirilmiş diğer bir çalışmada Liu ve Ma [53] Teknoloji kabul modeline dayalı olarak İngilizce öğrenenlerin, informal dijital İngilizce öğreniminde ChatGPT'yi kullanımının ölçülmesi ölçeğinde “Algılanan Fayda, Algılanan Kullanım Kolaylığı ve Kullanıcı Bilgi Teknolojisinin Kabulü” alt boyutlarından oluşmaktadır.

Elde edilen sonuçlara bakıldığında geliştirilen “DAİÖYÖ”nün ilgili konuyla benzer araştırmalar için gerekli olan geçerlik ve güvenilirliği karşıladığı görülmektedir. Bu amaçla geliştirilen bu ölçek öncelikle üniversite öğrencilerinin dijital becerilerinin belirlenmesi ve geliştirilmesi konularında hangi noktalara odaklanılacağı ile ilgili eğitmenler ve karar vericilere yararlı bilgiler sağlayabilir. Ayrıca işverenlere adayların dijital beceri düzeylerini değerlendirme imkânı sunabilir ve doğru adayları seçmelerine yardımcı olabilir.

#### KAYNAKLAR (REFERENCES)

- [1] N. Dabbagh and A. Kitsantas, “Personal learning environments, social media, and self-regulated learning: a natural formula for connecting formal and informal learning”, *The Internet and Higher Education*, 15(1), 3–8,2012.
- [2] N. Senemoğlu, *Gelişim öğrenme ve öğretim kuramdan uygulamaya*. Pegem Akademi Yayıncılık,2011.
- [3] M. C. Lohman, “A survey of factors influencing the engagement of two professional groups in informal workplace learning activities”, *Human Resource Development Quarterly*, 16(4), 501–527,2005.
- [4] D. Schugurensky, “The forms of informal learning: Towards a conceptualization of the field (Working paper)”. *Centre for the Study of Education and Work*, OISE/UT. <https://tspace.library.utoronto.ca/handle/1807/2733>,2000.
- [5] L. Küçükahmet, *Eğitim Bilimine Giriş*,Ankara: Nobel Akademik Yayıncılık, 2015.
- [6] M. Eraut, “Informal learning in the workplace”, *Studies in Continuing Education*, 26,2004.
- [7] A.L. Moore and J.D. Klein, “Facilitating Informal Learning at Work”, *TechTrends* 64, 219– 228. <https://doi.org/10.1007/s11528-019-00458-3>,2020
- [8] A. Martin and J. Grudziecki, “DigEuLit: Concepts and tools for digital literacy development. *Innovation in teaching and learning in information and computer sciences*”, 5(4), 249–267,2006.
- [9] B. Sözer, N. Özdamar, N. and H. Pilanci,“Yabancı dil öğrenimi için hazırlanan e-öğrenme ortamlarına ilişkin kullanılabilirlik araştırmalarının incelenmesi”, *Açıköğretim Uygulamaları ve Araştırmaları Dergisi*, 6(4), 174-207,2020.
- [10] J.A González-Martínez., M.L Bote-Lorenzo, E. Gómez-Sánchez and R. CanoParra ,” Cloud computing and education: A state-of-the-art survey”, *Comput. Educ* 80, 132–151,2015.
- [11] E. Gilman, I.S Milara, M. Cortés and J.Riecki, “Towards User Supportin Ubiquitous Learning Systems”, *IEEE Trans. Learn. Technol.* 8, 55–68,2015.
- [12] G. Yokus., “Eğitim Fakültesi öğrencilerinin mobil öğrenmeye ilişkin görüşlerinin incelenmesi ve eğitim bilimler alanına yönelik mobil uygulama geliştirme çalışması”, *Mersin Üniversitesi, Eğitim Bilimleri Enstitüsü*,2016
- [13] C. Karadeniz, “Müzedede Dijital Teknolojilerin Kullanımı ve Salgın Sürecinde Dijital Katılım”. *İdil Sanat ve Dil Dergisi*, 9(70), 975-984,2020.
- [14] M. Mehrvarz, E. Heidari, M. Farrokhnia, and O. Noroozi, “The mediating role of digital informal learning in the relationship between students' digital competence and their academic performance”, *Computers ve Education*, 167, 104184,2021.
- [15] T. He, Q. Huang, X. Yu and S. Li, “Exploring students' digital informal learning: the roles of digital competence and DTPB factors” *Behaviour ve Information Technology*, 40(13), 1406-1416,2021.
- [16] K. Akife and A. Göçer, “Mobil Öğrenme Aracı Olarak Podcastin Türkçe Öğretiminde Kullanılabilirliği Üzerine: Kuramsal Bir Çalışma” *Uluslararası Eğitim Bilim ve Teknoloji Dergisi*, 7(2), 77-97,2021.
- [17] O.Sadovets, O. Martynyuk, O. Orlovska, H. Lysak, S. Korol and M. Zembytska, “Gamification in the informal learning space of higher education (in the context of the digital transformation of education)”,*Postmodern Openings*, 13(1), 330-350,2022.
- [18] H. Xianhan, L. Chun S. Mingyao and S. Caixia, “Associations of different types of informal teacher learning with teachers' technology integration intention”, *Computers ve Education*, 190, 104604,2022.
- [19] Ö. Kandemir and G. Ulusoy, “Bir Yapı'dan Bir Kent'e Kültürel Değerlerin Artırılmış Gerçeklik Teknolojileriyle Ortaya Konulması”,*Yakın Doğu Üniversitesi Yakın Mimarlık Dergisi*, 7(1), 90-114, 2023.
- [20] L.S. Ng, S. M. Thang and N. M. Noor, “The usage of social networking sites for informal learning: A comparative study between Malaysia students of different gender and age group”, In *Research Anthology on Applying Social Networking Strategies to Classrooms and Libraries* (pp. 1467-1481). Global, 2023.
- [21] P. Panagiotidis, P. Krystalli and P. Arvanitis, “Technology as a motivational factor in foreign language learning”, *European Journal of Education*, 6(1), 69-84,2023.
- [22] J. W. Creswell, and V. L Plano Clark, “Designing and conducting mixed methods research”, Thousand Oaks, CA: Sage,2011.
- [23] J. W. Creswell, “Educational research: Planning, conducting, and evaluating quantitative and qualitative research (4th ed.)”. MA: Pearson Education, Inc.,2011”
- [24] J. R. Fraenkel, and N. E Wallen,“How to design and evaluate research in education” (7th ed.). McGraw-Hill,2011.
- [25] İ. Seçer, SPSS ve Lisrel ile pratik veri analizi: Analiz ve raporlaştırma. Ankara:Anı,2015.
- [26] E. Tavşancıl. Tutumların ölçülmesi ve SPSS veri analizi. Ankara: Nobel Yayın Dağıtım,2010.




- [27] A. A. Tezbaşaran, Likert tipi ölçek geliştirme kılavuzu, 2. Baskı.: Türk Psikologlar Derneği Yayınları,1997.
- [28] S.,Tekindal, “Duyuşsal özelliklerin ölçülmesi için araç oluşturma”, Kocaeli Üniversitesi, 2009.
- [29] A. Erkuş, Psikolojide ölçme ve ölçek geliştirme. Ankara: Pegem Yayınları,2012.
- [30] S. Gürbüz and F. Şahin, Sosyal bilimlerde araştırma yöntemleri, Ankara: Seçkin Yayıncılık.,2014.
- [31] N. Sümer, “Yapısal Eşitlik Modelleri. Türk Psikoloji Yazıları”, 3(6), 49-74,2000.
- [32] K. Özdamar, Tabloların Oluşturulması, Güvenirlik ve Soru Analizi. Paket Programlarla İstatistiksel Veri Analizi-1. (5. Baskı). Eskişehir: Kaan Kitabevi,2004.
- [33] Ş. Büyüköztürk, Faktör analizi: Temel kavramlar ve ölçek geliştirmede kullanımı. Kuram ve uygulamada eğitim yönetimi, 32(32), 470-483,2002.
- [34] M. Allen, The sage encyclopedia of communication research methods (Vols. 1-4), 2017.
- [35] G.H. Dunteman, Principal components analysis, Sage, 69,1989.
- [36] H.F. Kaiser, “The Application of Electronic Computers to Factor Analysis, Educational and Psychological Measurement”, 20, 141-151,1960.
- [37] R. B Cattell, The scientific use of factor analysis in behavioral and life sciences. New,1978.
- [38] Ö. Çokluk, G. Şekercioğlu and Ş. Büyüköztürk, Sosyal bilimler için çok değişkenli istatistik: SPSS ve LISREL uygulamaları. Ankara: Pegem Akademi, 2012.
- [39] Ş. Büyüköztürk, Sosyal bilimler için veri analizi el kitabı. Ankara: Pegem Yayıncılık,2010.
- [40] A. Akgül and O. Çevik, İstatistiksel analiz teknikleri SPSS’te işletme yönetimi uygulamaları. Ankara: Emek Ofset,2003.
- [41] Ş. Kalaycı, SPSS uygulamalı çok değişkenli istatistik teknikleri. Ankara: Asil Yayın Dağıtım,2008.
- [42] R. L. Worthington and T. A. Whittaker, “Scale development research: A content analysis and recommendations for best practices”, *The counseling psychologist*, 34(6), 806-838,2006.
- [43] R. E. Schumacker and R. G.Lomax, A Beginner's Guide to Structural Equation Modeling, New York: Taylor and Francis Group,2010.
- [44] C. Çapık,“Geçerlik ve güvenilirlik çalışmalarında doğrulayıcı faktör analizinin kullanımı”,*Journal of Anatolia Nursing and Health Sciences*, 17(3), 196-205,2014.
- [45] B. M. Byrne, “Structural equation modeling with lisrel, prelis, and simplis: Basic concepts, applications and programming”, Mahwah, New Jersey: Lawrence Erlbaum Associates,1998.
- [46] L. T. Hu and P. M. Bentler, “Cutoff criteria for fit indexes in covariance structure analysis: Conventional criteria versus new alternatives. Structural equation modeling”: *a multidisciplinary journal*, 6(1), 1-55,1999.
- [47] H. W. Marsh, K. T. Hau, C. Artelt, J. Baumert and , J. L. Peschar ,”OECD’s brief self-report measure of educational psychology’s most useful affective constructs: Cross-cultural, psychometric comparisons across 25 countries”, *International Journal of Testing*, 6(4), 311-360,2006.
- [48] R. B. Kline, “Principles and practice of structural equation modeling”, New York, NY: Guilford Press,2011.
- [49] M. M. Yaşhoğlu, “Sosyal bilimlerde faktör analizi ve geçerlilik: Keşfedici ve doğrulayıcı faktör analizlerinin kullanılması”, *İstanbul Üniversitesi İşletme Fakültesi Dergisi*, 46, 74-85,2017.
- [50] R. P. Bagozzi, Evaluating structural equation models with unobservable variables and measurement error: a comment,1981.
- [51] T. He and C. Zhu, C. “Digital informal learning among Chinese university students: the effects of digital competence and personal factors”, *International Journal of Educational Technology in Higher Education*, 14(1), 1-19,2017.
- [52] G. Liu, C. Ma, J. Bao, J. And Z. Liu,” Toward a model of informal digital learning of English and intercultural competence: a large-scale structural equation modeling approach” *Computer Assisted Language Learning*, 1-25,2023.
- [53] G. Liu and C. Ma, “Measuring EFL learners’ use of ChatGPT in informal digital learning of English based on the technology acceptance model”, *Innovation in Language Learning and Teaching*, 1-14,2023.

## DİJİTAL ARAÇLARLA İNFORMAL ÖĞRENME YETERLİLİK ÖLÇEĞİ

	Kesinlikle Katılıyorum	Katılıyorum	Biraz Katılıyorum	Katılmıyorum	Hiç Katılmıyorum
1-Dijital ortamlarda öğrenirken, eksiklerimi vakit kaybetmeden belirlerim.	5	4	3	2	1
2-Dijital ortamdaki öğrenme sürecinde kendi hedeflerimi oluştururum.	5	4	3	2	1
3-Dijital ortamlarda çalışma planımı kendim belirlerim.	5	4	3	2	1
4- Dijital ortamlardaki öğrenme sürecinde çalışma planımı uygulamada sorun yaşamam.	5	4	3	2	1
5- Dijital ortamlarda bir konuyu öğrenirken öğrenmeye ayracağım zamanı etkili kullanabilirim.	5	4	3	2	1
6- Mesleğimi icra etmede gerekli olan yeni bilgilere ulaşmak için dijital teknolojileri kullanırım.	5	4	3	2	1
7- Farklı öğrenme stratejileri geliştirmek için gerektiğinde dijital teknolojileri kullanırım.	5	4	3	2	1
8- Dijital teknolojileri öğrenmeye yardımcı olacak ipuçlarını aramak için kullanırım.	5	4	3	2	1
9- Dijital teknolojileri teknik yardım almak için kullanırım.	5	4	3	2	1
10- Arkadaşlarımla sorunlarımla çözebilmeleri için onlara uygun dijital araçları önerebilirim.	5	4	3	2	1
11- İnfomal öğrenme ile ilgili olduğumu düşündüğüm dijital teknolojilerde meydana gelen yenilikleri takip ederim.	5	4	3	2	1
12- İnfomal öğrenmemi desteklemek için resim düzenleme araçları kullanabilirim.	5	4	3	2	1
13- İnfomal öğrenmemi desteklemek için resim düzenleme programları kullanabilirim.	5	4	3	2	1
14- İnfomal öğrenmemi desteklemek için video düzenleme programlarını kullanabilirim.	5	4	3	2	1
15- İnfomal öğrenmemi desteklemek için video düzenleme araçları kullanabilirim.	5	4	3	2	1
16- Çevrimiçi kitap tartışma kulüpleriyle öğrenme konusunda işbirliği yaparım.	5	4	3	2	1
17- İnfomal öğrenme sürecinde farklı iletişim araçlarını (ör: Whatsapp, telegram, Skype vb.) kullanabilirim.	5	4	3	2	1
18- İnfomal öğrenme sürecinde web tarayıcılarını (ör: Edge, Google Chrome vb.) kullanabilirim.	5	4	3	2	1
19- İnfomal öğrenme sürecinde çevrimiçi çeviri araçları (ör: Google Çeviri, educeviri vb.) kullanabilirim.	5	4	3	2	1
20- Bir konuyu araştırırken farklı dijital araçları (masaüstü bilgisayar, mobil cihaz, tablet vb.) kullanırım.	5	4	3	2	1
21- İnfomal öğrenme sürecinde arama motorlarını (ör: Google, Yandex vb.) kullanabilirim.	5	4	3	2	1
22- Beni güvensiz hissettiren çevrimiçi içerik gördüğümde, onu kapatırım.	5	4	3	2	1

# Blok Zincir (Blockchain) Üzerine Yazılan Yüksek Lisans ve Doktora Tezlerinin Bibliyometrik Analizi

*Araştırma Makalesi / Research Article*

 Hacı Mehmet ALAKAŞ\*,  Erdal KÜPÇÜOĞLU,  Tamer EREN

Kırıkkale Üniversitesi, Mühendislik ve Doğa Bilimleri Fakültesi, Endüstri Mühendisliği Bölümü, Kırıkkale, Türkiye

[hmalagas@kku.edu.tr](mailto:hmalagas@kku.edu.tr), [erdal.kupcuoglu@gmail.com](mailto:erdal.kupcuoglu@gmail.com), [teren@kku.edu.tr](mailto:teren@kku.edu.tr)

(Geliş/Received:15.03.2024; Kabul/Accepted:03.09.2024)

DOI: 10.17671/gazibtd.1453335

**Özet**—Bu çalışmada blok zincir üzerine Türkiye’de yazılan lisansüstü tezlerinin bibliyometrik analizi yapılmıştır. Ulusal tez merkezinde bulunan lisansüstü (yüksek lisans - doktora) tezleri taranmıştır. “blokzincir”, “blok zincir” ve “blockchain” terimleri kullanılarak aramalar gerçekleştirilmiştir. Yapılan arama sonucu elde edilen 476 tezin analizi gerçekleştirilmiştir. Tezlere ait veriler danışman, üniversite, tür, yazıldığı yıl, enstitü, ana bilim dalı, yazıldığı konulara, kullanılan anahtar kelimelere ve diline göre analiz edilmişlerdir. Özellikle Fen Bilimleri Enstitülerinde, Bilgisayar Mühendisliği ana bilim dalında yapılan tez çalışmalarının blok zincir yapısı, çalışma sistemi, güvenilirliği, akıllı sözleşmeler ile gerçekleştirilen teknik altyapıları incelediği görülmüştür. Sosyal Bilimler Enstitüsü işletme, hukuk, ekonomi, bankacılık alanlarında blok zincir alt yapısını kullanan kripto paraların da kullanıldığı, farklı alanlara uygulanan blok zincir yapılarının verimliliği, etkinliği, kullanılabilirliği, hukuksal etkileri üzerine tezler yazılmıştır. Blok zincir teknolojisi, güvenilirlik, merkezi olmayan yapısı ve hızlı sonuçlar alınabilmesi nedeniyle, blok zincir teknolojisi üzerine yetişmiş insan kaynağının az olmasına rağmen günümüzde daha da yaygınlaşmaya başlamıştır. Bu çalışmamızda Türkiye’de blok zincir alanında yapılan akademik çalışmalara kaynaklık yapabilecek analizler gerçekleştirilmeye çalışıldı.

**Anahtar Kelimeler**—blok zincir, bibliyometrik analiz, lisans üstü tezler

## Bibliometric Analysis of Master's and Doctoral Theses on Blockchain

**Abstract**—In this study, a bibliometric analysis of postgraduate theses written on blockchain in Turkey was conducted. Graduate (master's - doctoral) theses in the national thesis centre were scanned. Searches were carried out using the terms "blockchain", "block chain" and "blockchain". The 476 theses obtained as a result of the search were analyzed. The data of the theses were analyzed according to supervisor, university, type, year of writing, institute, main branch of science, subjects, keywords used and language. In particular, it was observed that the thesis studies carried out in the Institutes of Natural Sciences, Computer Engineering Department examined the block chain structure, working system, reliability, technical infrastructures realized with smart contracts. Theses were written on the efficiency, effectiveness, usability, and legal effects of block chain structures applied to different fields, including cryptocurrencies using block chain infrastructure in the fields of business, law, economy, and banking in the Institute of Social Sciences. Block chain technology has become more widespread today due to its reliability, decentralized structure and fast results. The weakness of block chain technology in terms of human resources reduces the number of such studies. In this study, we tried to carry out analyses that can serve as a source for academic studies in the field of block chain in Turkey.

**Keywords**—blockchain, bibliometric analysis, graduate theses



## 1. GİRİŞ (INTRODUCTION)

Çok kısa bir sürede çok fazla verinin ortaya çıkmaya başladığı, verilerin kullanılabilir hale getirebilmesi üzerine yapay zeka çalışmalarının günden güne hızlandığı 21. yüzyılda ortaya çıkan teknolojiler hayatımızda yerini almaya başladı. Pandemi gibi dünyayı komple etkisi altına alan olumsuz durumlar ekonomileri ve teknolojileri bu yeni durumlara karşı hazırlıklı hale getirmek ihtiyacını ortaya çıkardı.

Sosyal medya, finans uygulamaları ve diğer teknolojiler büyük miktarda veri üretmektedir. Verilerin kontrol, yönetim, güvenli paylaşılması daha da önem kazanmaktadır. Veri depolama ve veri yönetimine her gün yeni bir teknoloji katılmaktadır. Blok zinciri (Blockchain) bu teknolojilerden en dikkat çekenidir [1].

Blok zincir, blokzincir veya blockchain, dağıtık bir veri tabanı olarak tanımlanabilir. Bu veri tabanı, ağa bağlı olan tüm katılımcılar tarafından paylaşılır ve güncellenir. Blok zincir teknolojisi, kripto paralar, dijital işletme, uluslararası ödeme sistemleri gibi çeşitli alanlarda kullanılmaktadır [2].

Blok zincir, değiştirilmeye karşı korumalı ve kullanıcılar arasında teyitli işlemler sağlayan merkezi olmayan bir veri tabanı teknolojisi olarak tanımlanmaktadır. Araştırmalar, blok zincirin doğrulanmış işlemleri tüm katılımcılara tamamen açık bir şekilde sunarak güvensizlik ve şüphe sorunlarını azalttığını göstermiştir. Ayrıca, bir veri erişim sistemi ve fikir birliğinin sağlanması ile kullanıcıların ağ üzerindeki işlemlerini yönetmesine, izlemesine imkan tanımaktadır. Blok zincir birçok alanda (endüstri, finans, sağlık gibi) yenilikler sunarak, değiştirilemeyen yapısı ile popülerliğini artırmıştır [3].

Bu çalışmada Ulusal tez merkezinde bulunan, tez adında "blok zincir", "blokzincir" ve "blockchain" kelimelerini içeren yüksek lisans ve doktora tezleri bibliyometrik olarak araştırılmışlardır. Tezler iki aşamada incelenmişlerdir. Tezler tez danışmanı, yazılan üniversite, tezin türü, tezin yazıldığı yıl, tezin yazıldığı enstitü, tezin türüne ve kullanılan anahtar kelimelere göre sınıflandırma bilgileri olarak birinci aşamada incelenmişlerdir. İkinci aşamada, tezlerin yazıldığı ana bilim dalı ile birlikte tezlerin hangi konularda yazıldığı, tezlerde kullanılan blok zincir teknolojileri, blok zincir uygulama alanları ve uygulama alanlarına göre hangi blok zincir teknolojilerinin kullanıldığı, yeni araştırmalara kaynaklık edebilmesi açısından incelenmiştir.

Çalışmanın ikinci bölümünde blok zincir ve blok zincirin yapısı araştırılmış, üçüncü bölümde bibliyometrik analiz metodunu kullanarak blok zincir ver blockchain üzerine yazılan tezlerden oluşan literatür araştırması yapılmış, dördüncü bölümde makalede kullanılan yöntem anlatılmış, beşinci bölümde araştırma sonucunda elde edilen bulgulara yer verilmiş ve sonuçlar değerlendirilmiştir.

## 2. BLOK ZİNCİR VE BLOK ZİNCİRİN YAPISI (BLOCKCHAIN AND BLOCKCHAIN STRUCTURE)

Blok zincir kavramı ilk olarak 2008 yılında Bitcoin'in arkasındaki teknoloji olarak ortaya çıkmıştır. İlk ortaya atıldığı makalede çevrimiçi ödemelerin aracı bir finans kuruluşu olmadan taraflar arasında gönderilmesini mümkün kılan eşler arası bir elektronik nakit ödeme sistemi olarak tanıtılmıştır [4]. Ancak süreç içerisinde blok zincir teknolojisinin Bitcoinlerin ötesinde dijital bir araç olduğu kanıtlanmıştır. Hatta blok zincir teknolojisinin dördüncü sanayi devriminin temel bir teknolojisi olduğu belirtilmiştir [5].

Bilgi teknolojisi olan blok zincir, 4.Sanayi Devriminin (Endüstri 4.0) yapı taşı sayılan bir yenilik olarak tanımlanmaktadır [6]. Blok zincir teknolojisinin amacı, üçüncü taraf kuruluşların verileri ve işlemleri kontrol altına alamadığı merkezileşmemiş bir ortam oluşturmaktır [7]. Veriler, kriptografik yöntemlerle yapılması muhtemel sahtekarlıklara karşı korunmaktadır [8]. Blok zincir, içinde yer alan düğümler arasındaki etkileşimlerden oluşmaktadır [9]. Bu anonim düğümler tarafından onaylanan, gün geçtikçe büyümeye devam eden, veri kayıtlarını tutan veri tabanı sistemidir [7].

Blok zincir teknolojisinin denetlenebilirlik, merkezi bir otoriteye bağlı olmayan yapısı, izlenebilirlik, şeffaflık, anonimlik, güvenlik özellikleridir. Bunlar sisteme fayda veya avantaj sağlayabilmektedir [10].

*Denetlenebilirlik (Traceability):* Blok zincir teknolojisi ağdaki bütün işlemleri doğrulamak ve kaydetmek için dijital zaman damgası ve dijital dağıtılmış defter kullanmakta böylelikle ağdaki düğümlerden birine erişim sağlandığında önceki kayıtlar denetlenebilmektedir.

Blok zincirde yapılan işlemler doğrulandığından ve sonrasında sisteme kaydedildiğinden dolayı kullanıcılar daha önce kayıtlı olan verilere ağdaki düğümler sayesinde erişim sağlayarak kolaylıkla düğümü izleyebilir ve doğrulayabilir [11]. Kaynak takibi sağlayan blokzincir teknolojisi kullanıcılarına denetlenebilir hizmet imkanı sunar [10].

*Merkeziyetsizlik (Decentralized):* Blok zincir teknolojisinin merkeziyetçi olmayan yapısı karar verme ve planlama sürecinde güvenilir bir yetkili olarak adlandırılan üçüncü tarafa gerek olmadan işlemlerin tutarlılığını ve bütünlüğünü sağlamaktadır [12] [13]. Blok zincirde bulunan veriler bir sunucuda toplanmamış farklı olan düğümler arasında dağıtılmıştır [14]. Merkezi çözümler maliyetlidir ve güven problemi oluşturur bunun aksine ademi merkeziyette sistem esnek, güvenilir ve verimlidir [15].

*Azaltılmış İşlem Maliyetleri (Reduced Transaction Costs):* Blokzincir teknolojisi sayesinde işlemler gerçekleşirken geleneksel uygulamaların aksine ilgili gider maliyetlerinde önemli ölçüde düşüş meydana gelmektedir [16].

*İzlenebilirlik (Transparency):* Ürünlerin kaynağına yönelik ayrıntılara ulaşma imkanı tanımakta ve her bir blok izlenebilmektedir. Bu teknoloji veritabanı güvenliği kullanarak, devamlı izlenebilirlik sağlamakta, böylelikle ürünler hakkında karar alınırken kalitelerinin değerlendirilmesine ve bilgilerinin doğruluğunun test edilmesine yardımcı olmaktadır [17].

*Anonimlik (Anonymity):* Blokzincir teknolojisi üçüncü bir tarafa ihtiyaç duymadığından ötürü merkezi otorite kullanıcıların bilgilerini kaydedemez ve izleyemez [18]. Kullanıcıların gizliliği önemli bir husustur ve korunması amaçlandığından anonim bir kimlik oluşturulmaktadır. Sistem kriptografik özel anahtar oluşturularak anonimlik sağlanmakta ve kullanıcılar etkileşim sağlarken özel bilgilerini kullanmadan ağ içerisinde karşılıklı etkileşimde bulunabilmektedir [17].

*Gizlilik (Privacy):* Blokzincir teknolojisi kullanıcılara ait verilerin gizliliğinden ödün vermeyen bir çerçeveye sahiptir [19]. Kullanıcılar verilerini kontrol edebilmekte ve ağ üzerinde işlem yapabilmektedir.

*Güvenlik (Security):* Blok zincir teknolojisi mevcut çözümlere kıyasla verilerin güvenliğini kriptografik olarak korumaktadır. Cihazlar arasında yapılan bütün etkileşimler güvence altına alınmıştır ve bu sayede blokta kayıtlı olan veriler manipüle edilememektedir [20]. Ayrıca, koruma gücü daha yüksek olan ortak anahtar yapısının kullanılması ve merkezi bir sunucunun var olmaması ağa karşı yapılması muhtemel saldırıları zorlaştırmakta ve güvenilir bir çerçeveye sunmaktadır [21].

*Güven (Trust):* Yapılan işlemlerin uygunluğu ve geçerliliği sisteme kayıt yaptıran kullanıcılara ifade edilmektedir [10]. Blokzincir teknolojisi değişmez ve doğru veri sağlandığından yönetim sistemine güven duyulmaktadır.

*Şeffaflık (Transparency):* Bütün erişim ve verilerin kontrolünün merkezi sunucuya ait olan merkezi modellere karşın, blok zincir teknolojisiyle işlemlerin açıklamalı bilgilerine, ortak veri kümelerine ulaşma imkanı sağlanan şeffaflık sunmaktadır [21]. Bu şeffaflık durumu blok zincirde gerçekleşen dijital işlemlerin kurcalanmaya karşı korumaktadır [22]. Blok zincir teknolojisi düğüm noktalarında zincirin bir yedeğini saklamakta ve depolamakta böylelikle veriler kullanıcılar tarafından izlenebilir kılınmaktadır [20].

*Esneklik (Flexibility):* Blok zincir teknolojisi, akıllı sözleşmeler ile esnekliği iyileştirmektedir. Bu sayede, katılımcılar ile iş birliği yaparak stratejiler ve kaynaklar yeniden yapılandırılabilir [17].

*Değişmezlik (Immutability):* Her bir blok için şifre mevcuttur ve işlem kaydı tutulmaktadır [20]. İşlem başladıktan ve veriler deftere eklendikten sonra düzenleme işlemi yapılmamakta ve kayıtlar silinmemekte böylelikle kalıcı ve değiştirilmez bir yapı oluşmaktadır. Değişmeyen defterler sayesinde işlemlerin bütünlüğü sağlanmaktadır [21].

### 3. BİBLİYOMETRİK ANALİZ (BIBLIOMETRIC ANALYSIS)

Bibliyometrik analiz, bilimsel yayınların sayısal olarak değerlendirilmesi ve görselleştirilmesi yöntemidir. Bibliyometrik analiz, bir araştırma alanının gelişimini, etkisini, trendlerini ve ilişkilerini ortaya çıkarmak için kullanılabilir [23]. [2], blok zincir ile alakalı 2017-2022 yıllarına ait 134 yüksek lisans ve doktora tezini bibliyometrik olarak incelemişlerdir. [24], uyguladıkları çözümleme hakkında verileri ortaya koymayı amaçlamışlardır. [23], (WoS) Web of Science' da yayınlanan 2015-2020 seneleri arasında "makale" çalışmalarından 1106'ını incelemişlerdir. Bu verileri excel ve vosviewer programı ile görselleştirmişlerdir. [25], yaptıkları araştırma ile blok zincir üzerine yazılan 2018-2022 senelerini kapsayan 57 tez çalışmasını incelemişlerdir. [26], Türkiye'de kripto para temalı doktora tezleri Yükseköğretim Kurulu Başkanlığı Tez Merkezi'nden alınıp çalışma evrenine dâhil edilerek analiz gerçekleştirmişlerdir. Yükseköğretim Kurulu Başkanlığı Tez Merkezi'nde "Bitcoin", "Kripto para", "Ethereum", "blokzincir", "blok zincir", "blokzinciri", "blok zinciri" ve "cryptocurrency" arama terimleri kullanılarak aramalar gerçekleştirmişler. Yapılan arama sonucu elde edilen 29 doktora tezinin analizini gerçekleştirmişlerdir. [27], çalışmasında Scopus' da "turizm" ve "blok zincir" konusunun beraber bulunduğu 108 makale R programı kullanılarak incelenmiştir.

### 4. YÖNTEM (METHOD)

Bu çalışmada blok zincir üzerine Türkiye'de yazılan lisansüstü (yüksek lisans – doktora) tezlerinin bibliyometrik analizi gerçekleştirilmiştir. Ulusal tez merkezinde bulunan lisansüstü tezler incelenmiştir. "blokzincir", "blok zincir" ve "blockchain" terimleri kullanılarak aramalar gerçekleştirilmiştir. Yapılan arama sonucu elde edilen 476 tezin analizi gerçekleştirilmiştir. Tezlere ait veriler danışman, üniversite, tür, yazıldığı yıl, enstitü, ana bilim dalı, yazıldığı konulara ve diline göre analiz edilmişlerdir. Tezler iki aşamada incelenmişlerdir. Tezler tez danışmanı, yazılan üniversite, tezin türü, tezin yazıldığı yıl, tezin yazıldığı enstitü, tezin türüne ve kullanılan anahtar kelimelere göre sınıflandırma bilgileri olarak birinci aşamada incelenmişlerdir. İkinci aşamada, tezlerin yazıldığı ana bilim dalı ile birlikte tezlerin hangi konularda yazıldığı, tezlerde kullanılan blok zincir teknolojileri, blok zincir uygulama alanları ve uygulama alanlarına göre hangi blok zincir teknolojilerinin kullanıldığı incelenmiştir.

### 5. BULGULAR (RESULTS)

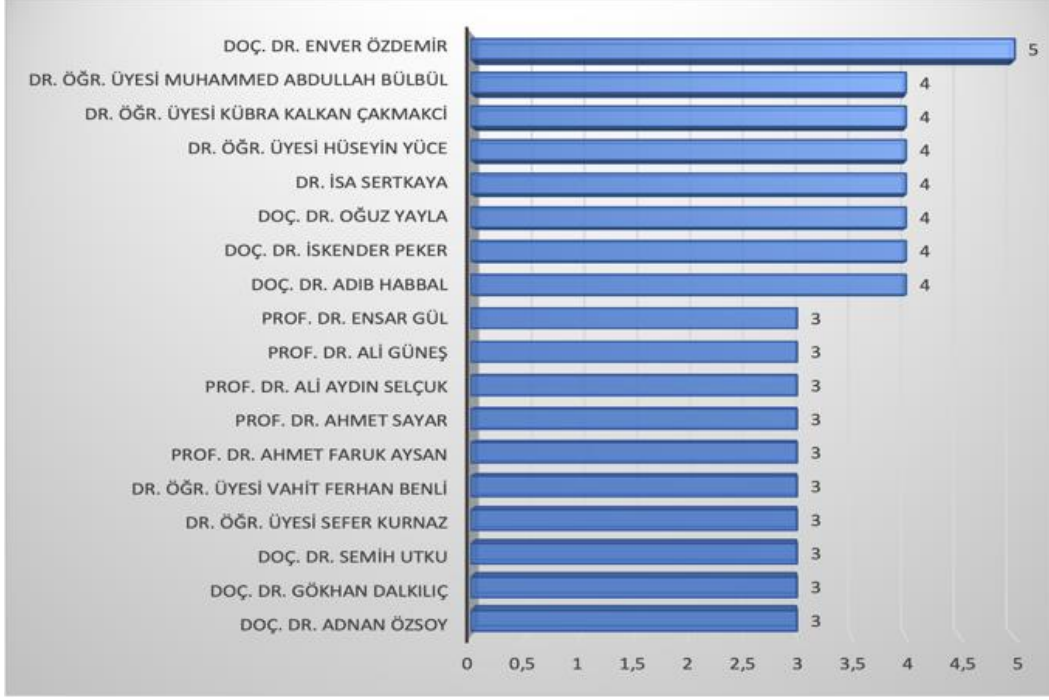
Çalışmada ele alınan tezler bibliyometrik olarak incelenirken danışman, üniversite, enstitü, türü, yıllına, yazım dili ve kullanılan anahtar kelimelerle ilgili dağılımlarına göre "Tezlerin Sınıflandırma Bilgileri" ve ana bilim dalı, konusu, blok zincir teknolojileri, uygulama alanları ile ilgili olarak da "Tezlerin Alanlarına ve İçeriğine

Göre İncelenmesi” olarak iki ana başlık altında yorumlanmıştır.

### 5.1 Tezlerin Sınıflandırma Bilgileri (Classification Information of Theses)

2024 Ocak ayı itibariyle Ulusal tez merkezinde bulunan lisansüstü (yüksek lisans – doktora) tezlerin bibliyometrik analizi ile;

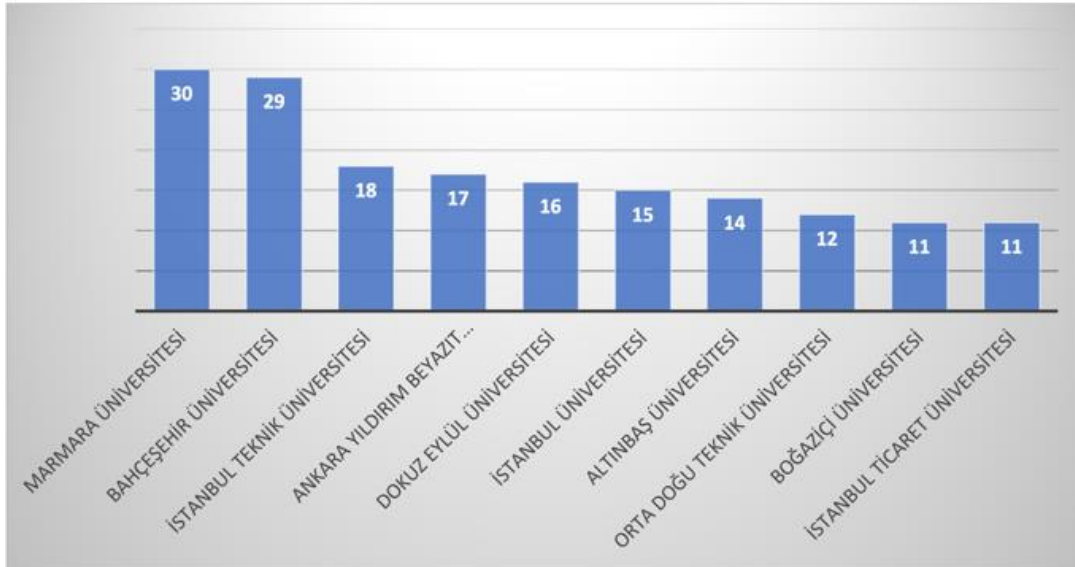
Marmara üniversitesinde en fazla sayıda tez yazılırken, İstanbul Teknik Üniversitesi Bilişim Uygulamaları Ana Bilim Dalı - Bilgi Güvenliği Mühendisliği ve Kriptografi Bilim Dalı öğretim üyesi Doç. Dr. Enver ÖZDEMİR bu alanda en çok tez danışmanlığı yapmıştır (Şekil 1). Üçten aşağı olanlar grafikte gösterilmemiştir.



Şekil 1– Danışmanlarına Göre Tez Sayıları (Number of Theses According to Advisors)

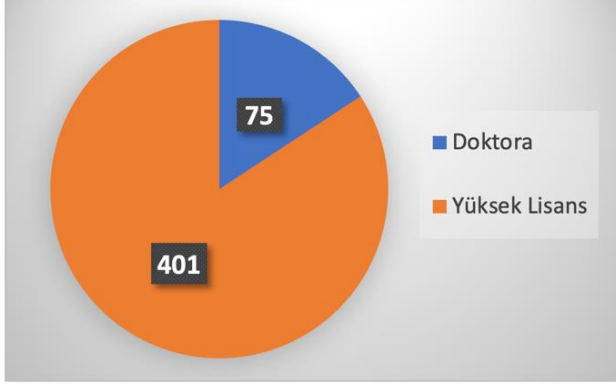
İncelenen tezler içerisinde Marmara üniversitesi (30) ve Bahçeşehir üniversitesi (29) tez ile en çok çalışma yapılan üniversiteler olmuştur. İstanbul Teknik üniversitesi (18), Ankara Yıldırım Beyazıt üniversitesi (17), Dokuz Eylül üniversitesi (16), İstanbul üniversitesi (15), Altınbaş

üniversitesi (14) Ortadoğu Teknik üniversitesi (ODTÜ) (12) ve Boğaziçi üniversitesi ile İstanbul Ticaret üniversiteleri (11) tez ile ilk on üniversite içerisinde girmişlerdir (Şekil 2).



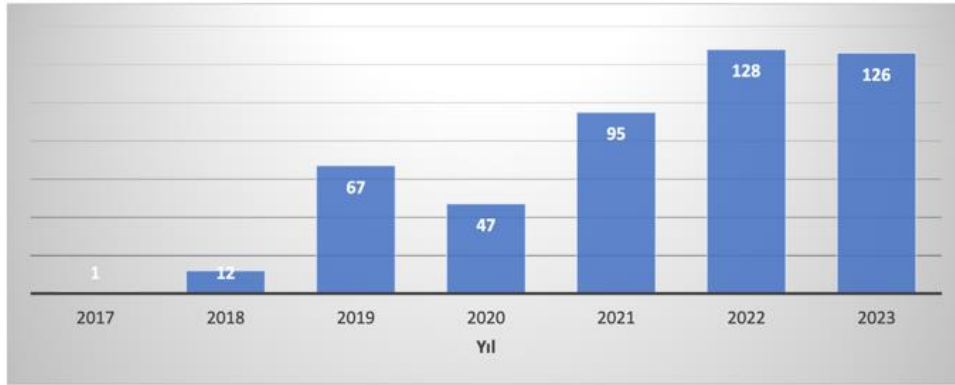
Şekil 1 - Üniversiteye Göre Tez Sayıları (Number of Theses by University)

2024 Ocak ayı sonuna kadar, incelenen toplam 476 tezde; 401 yüksek lisans ve 75 doktora tezi çalışması yapılmıştır (Şekil 3).



Şekil 2 - Türlerine Göre Tez Sayıları (Number of Theses by Type)

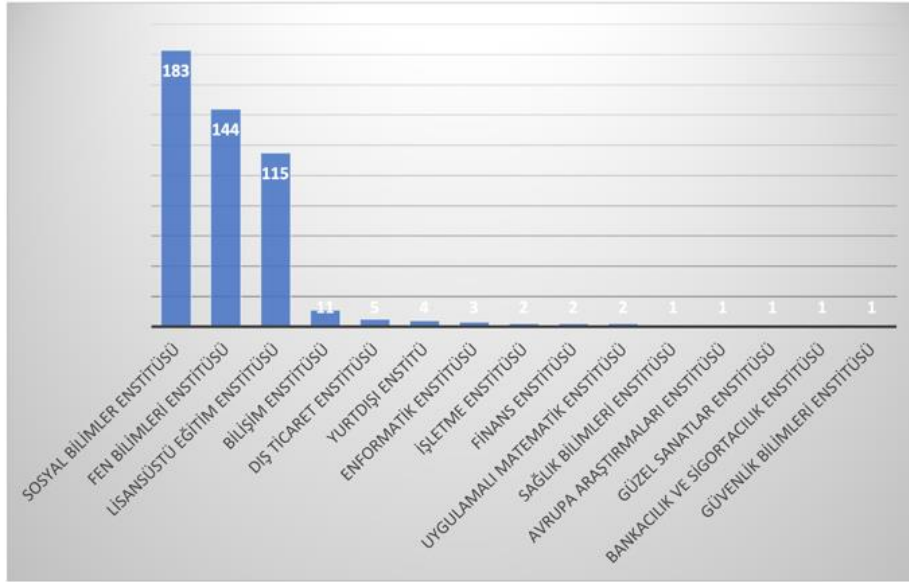
Blok zincir (blockchain) kavramının ilk olarak 2008 de tanıtılmasıyla başlayan araştırmalarla 2017 yılında ilk tez çalışması yapılmıştır (Şekil 4). Daha sonraki yıllarda 2018 (12), 2019 (67), 2020 (47), 2021 (95), 2022 (128) ve 2023 yılında (126) te bulunmaktadır. Özellikle pandemi (Covid-19) sonrası blok zincir alanında çalışmaların daha arttığı görülmüştü. 11 Mart 2020'de Covid-19 pandemi ilanıyla, sağlık alanının yanında finansal etkiler üzerindeki çalışmalar hızlanmıştır. Ekonomi ve finans alanında covid-19 etkileri görülmüştür. Ekonomi alanında araştırmalar çoğalmıştır [28].



Şekil 3 - Yıllara Göre Tez Sayıları (Number of Theses by Years)

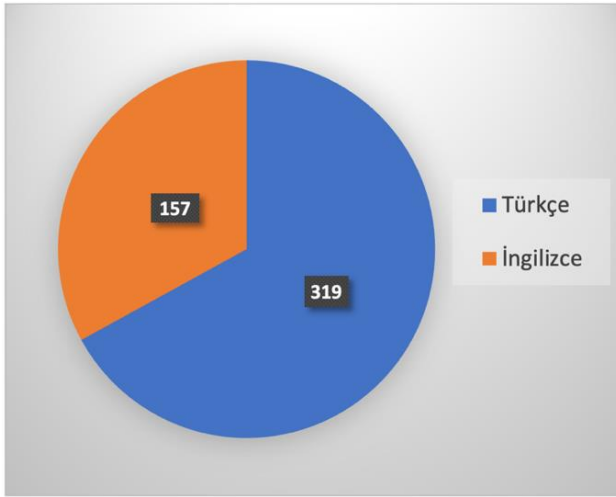
Tezlerin yapıldığı enstitüler incelendiğinde (Şekil 5), 183 adet tez ile Sosyal Bilimleri enstitüsü ilk sıradadır. Sonra 144 tez ile Fen Bilimleri Enstitüsü ve 115 tez ile Lisansüstü

Eğitim Enstitüsü gelmektedir. Lisansüstü Eğitim enstitüsünde hazırlanan tezler aslında diğer enstitülere girmesine rağmen bu enstitü adı altında kabul edilmiştir.



Şekil 4 - Enstitülere Göre Tez Sayıları (Number of Theses by Institutes)

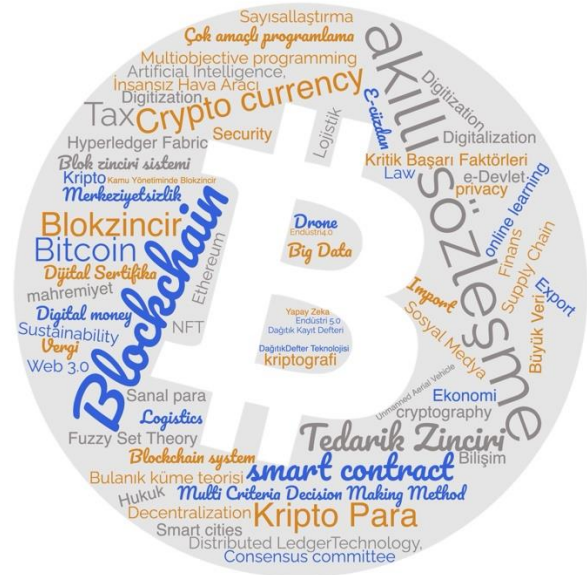
İncelenen 476 tezin 319 tanesi Türkçe, 157 tanesine İngilizce diliyle yazılmıştır (Şekil 6).



Şekil 6- Yazıldığı Dile Göre Tez Sayıları (Number of Theses According to the Language They Are Written in)

Anahtar kelime bulutu, incelenen tezlerde kullanılan anahtar kelimelerden oluşturulmuştur. Bu işlem için, kullanım kolaylığı, özelleştirilebilirlik, veri analizi yetenekleri ve internet üzerinden erişilebilirlik gibi avantajları nedeniyle WordClouds.com tercih edilmiştir. Bu özellikler, anahtar kelimelerin görsel olarak etkili bir şekilde sunulmasını sağlamaktadır. İncelenen lisansüstü tezlerde kullanılan anahtar kelimeler ile oluşturulan kelime bulutu şekil 7’de verilmiştir.

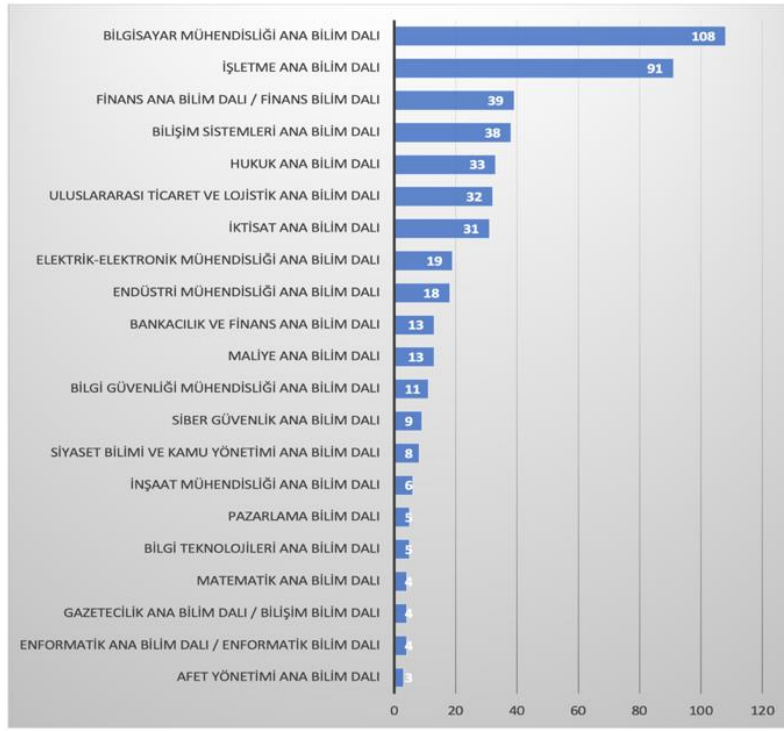
“Blok zincir, blockchain, kripto para, akıllı sözleşmeler, tedarik zinciri, smart contracts” anahtar kelimeler arasında en fazla kullanılanlardır.



Şekil 7- Kullanılan Anahtar Kelime Bulutu (Used Keyword cloud)

## 5.2 Tezlerin Alanlarına ve İçeriğine Göre İncelenmesi (Examination of Theses According to Their Fields and Content)

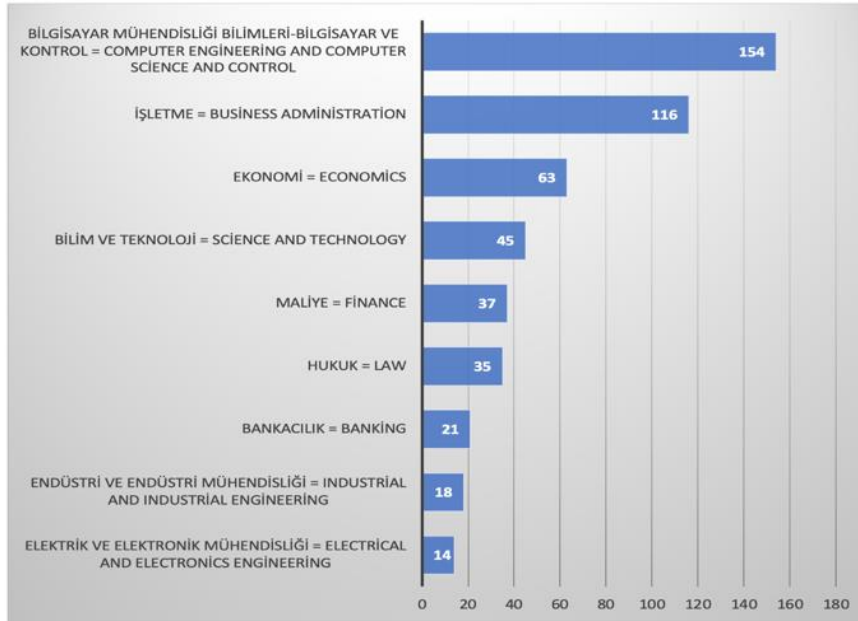
Bilgisayar Mühendisliği ana bilim dalı (108), Bilişim Sistemleri ana bilim dalı (38), Siber Güvenlik ana bilim dalı (9), Bilgi Güvenliği Mühendisliği ana bilim dalı (11), Bilgi Teknolojileri ana bilim dalı (5), Enformatik ana bilim dalı (4) alanlarında tezler yazılmıştır. Özellikle kripto para, akıllı sözleşmeler konularında İşletme ana bilim dalı (91), Finans ana bilim dalı (39), İktisat ana bilim dalı (31), Hukuk ana bilim dalı (33), Maliye ana bilim dalı (13), Bankacılık ve Finans ana bilim dalı (13) ana bilim dallarında tezler yazılmıştır (Şekil 8)



Şekil 8- Ana Bilim Dalına Göre Tez Sayıları (Number of Theses by Department)

Bilgisayar Mühendisliği Bilimleri ve Bilgisayar ve Kontrol (154), İşletme (116), Ekonomi (63), Bilim ve Teknoloji (45), Maliye (37), Hukuk (35), Bankacılık (21), Endüstri

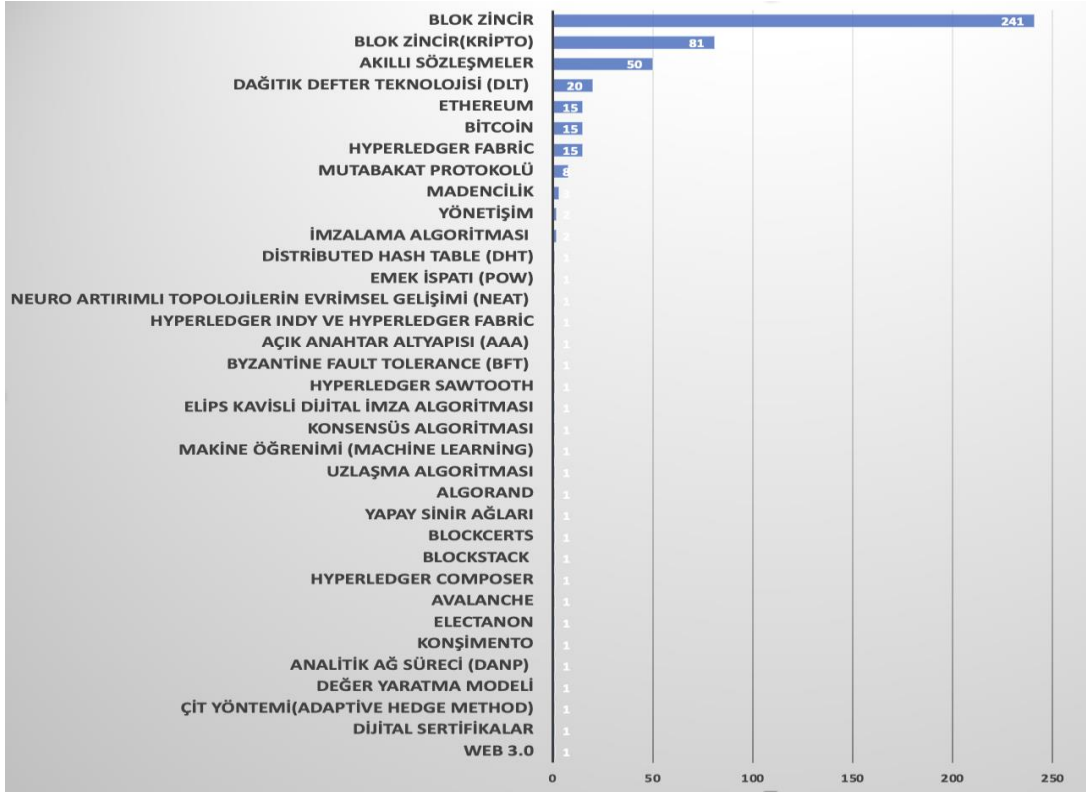
ve Endüstri Mühendisliği (18), Elektrik ve Elektronik Mühendisliği (14) konularında çoğunluk tezler yazılmıştır (Şekil 9).



Şekil 9- Konusuna Göre Tez Sayıları (Number of Theses by Subject)

İncelenen tezler içerisinde blok zincir teknolojisini bütün olarak(241) ve kripto para teknolojileri(81) üzerine çalışılan tez sayısı oldukça fazladır. Akıllı sözleşmeler (Smart Contracts) (50), Dağıtık defter teknolojisi (DLT)

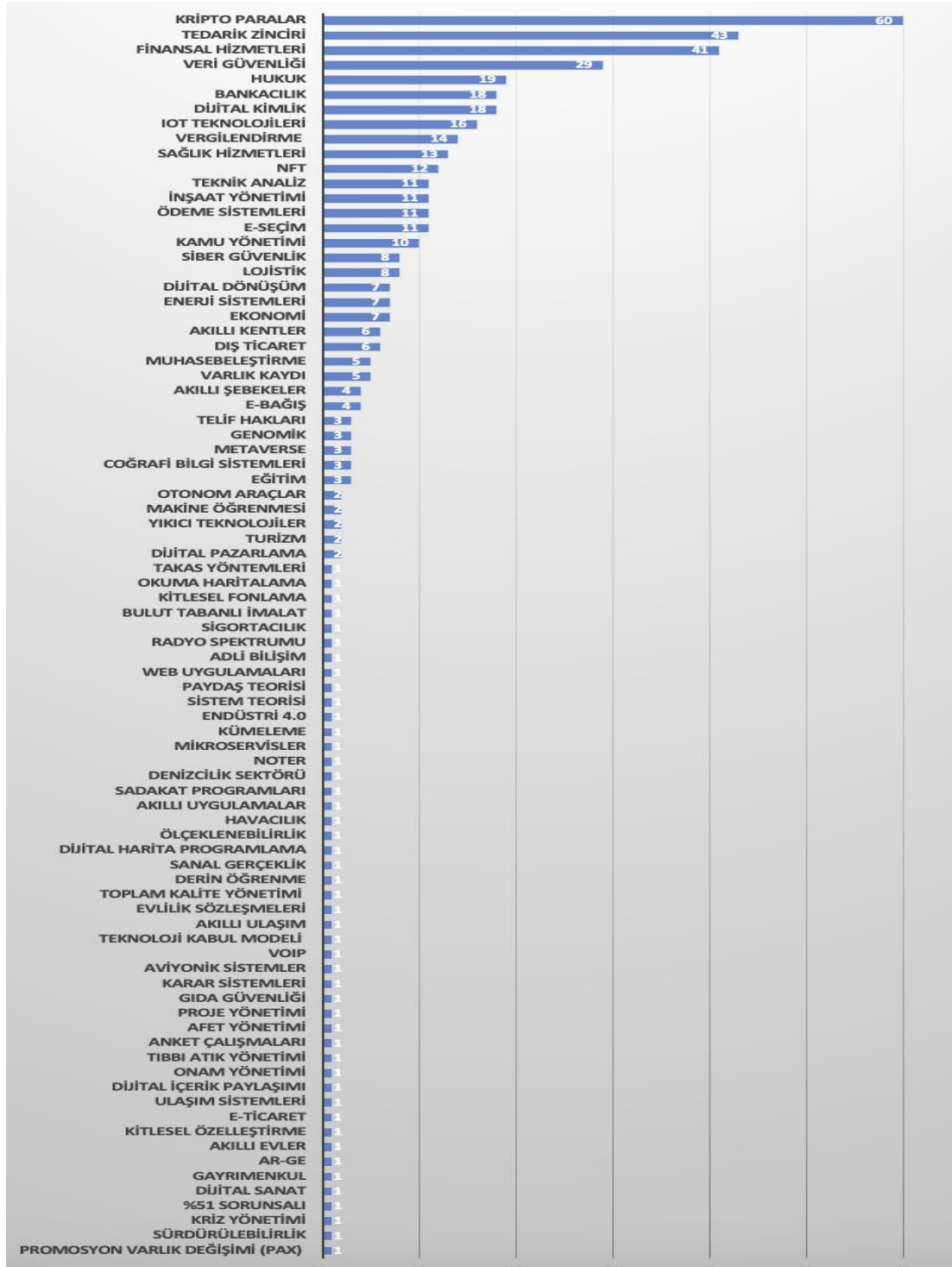
(20), Ethereum, Bitcoin, Hyperledger Fabric (15), ve Mutabakat protokolü (8) tezlerde diğer en çok kullanılan teknolojilerdir (Şekil 10).



Şekil 10- Kullanılan Blok Zincir Teknolojilerine Göre Tez Sayıları (Number of Theses by Blockchain Technologies Used)

476 tez içerisinde blok zincir uygulamaları arasında Kripto Paralar (60) en fazla çalışma yapılan uygulama olsa da, Tedarik Zinciri (43), Finansal Hizmetler (41), Veri Güvenliği (29), Hukuk (19), Bankacılık, Dijital Kimlik (18), IoT Teknolojileri (16), Vergilendirme (14), Sağlık Hizmetleri (13), NFT (Non-Fungible Token) (12), Teknik Analiz, İnşaat Yönetimi, Ödeme Sistemleri, E-Seçim (oylama) (11), Kamu Yönetimi (10), Siber Güvenlik, Lojistik (8), Dijital Dönüşüm, Enerji Sistemleri, Ekonomi

(7), Akıllı Kentler, Dış (uluslararası) Ticaret (6), Muhasebeleştirme, Varlık Kaydı (5), Akıllı Şebekeler, E-Bağış (4), Telif Hakları, Genomik, Metaverse, Coğrafi Bilgi Sistemleri, Eğitim (3), Otonom Araçlar, Elektronik Oylama, Makine Öğrenmesi, Yıkıcı Teknolojiler, Turizm, Dijital Pazarlama (2) ve diğer uygulama alanlarının toplamı Kripto Paralar üzerine yapılan çalışmalardan oldukça fazla olduğu görülmektedir (Şekil 11).

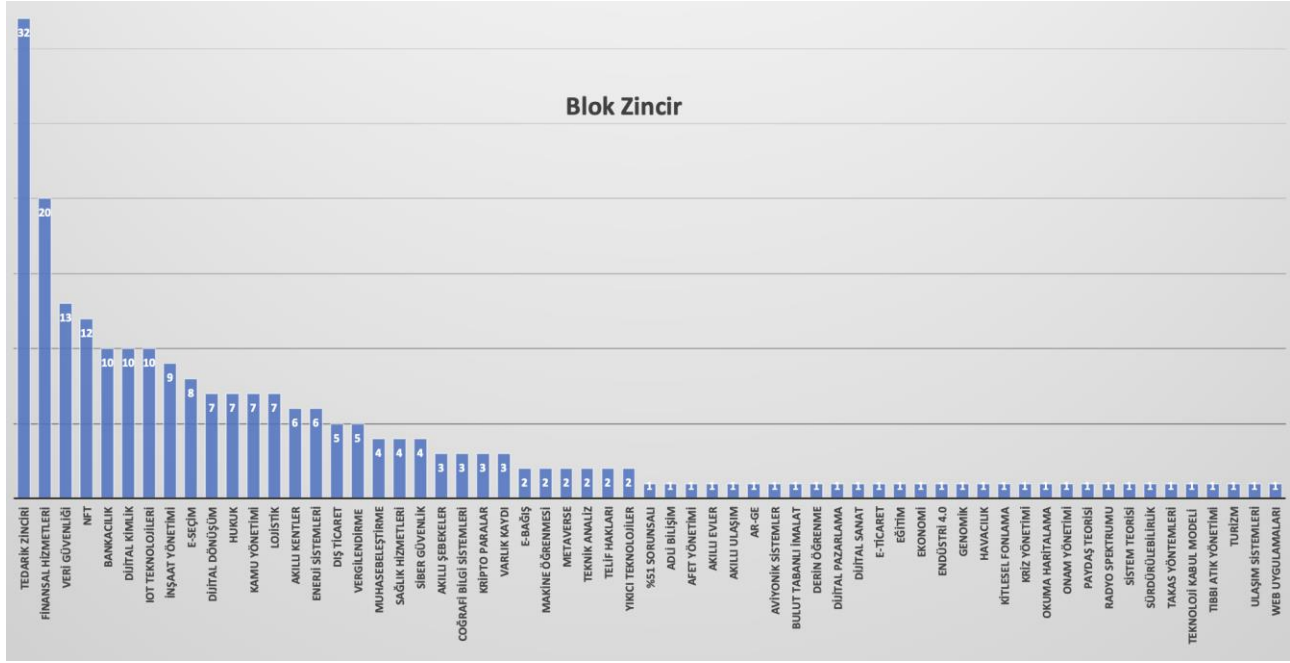


Şekil 11- Uygulama Alanlarına Göre Tez Sayıları (Number of Theses by Application Fields)

Çalışmamızın bu bölümünde incelenen tezlerde en çok araştırılan blok zincir teknolojilerinin uygulama alanları bulunmaktadır. Blok zincir teknolojisinin bütün olarak ele alındığı tez çalışmalarının uygulama alanları şekil 12'de

gösterilmiştir. Tedarik zinciri (32), Finansal Hizmetler (20), Veri Güvenliği (13) ve diğer alanlarda blok zincir teknolojileri kullanılarak araştırmalar yapılmıştır.

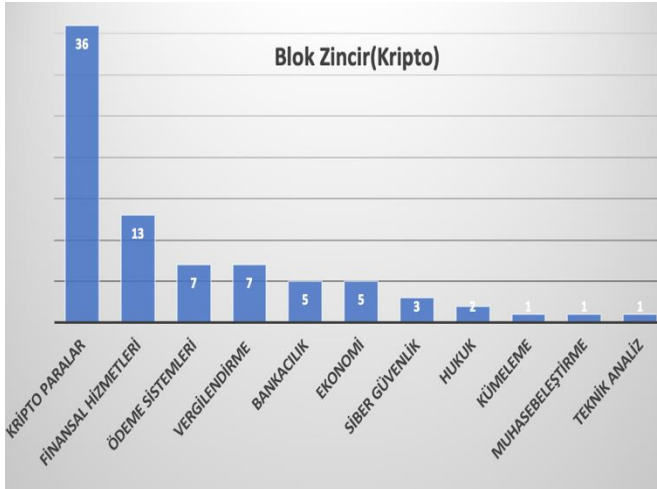




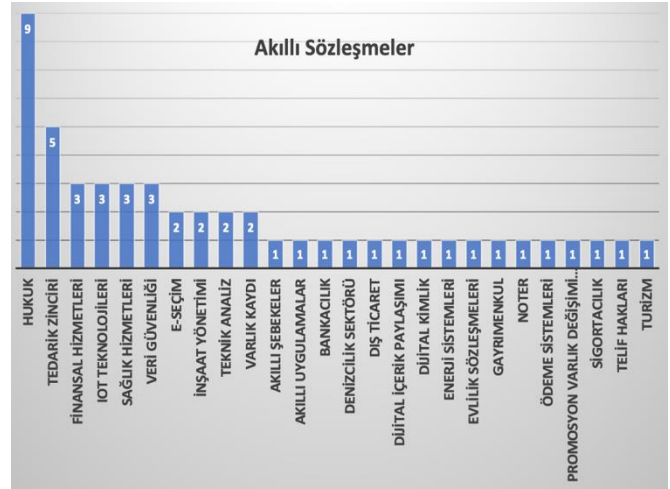
Şekil 12- Blok Zincir Teknolojisinin Uygulama Alanları Sayıları (Number of Application Areas of Blockchain Technology)

Blok zincir teknolojisinin kripto paralarla ilgili yapılan araştırmalar, Finansal Hizmetler, Ödeme Sistemleri, Vergilendirme, Bankacılık, Ekonomi, Siber Güvenlik, Hukuk, Kümeleme, Muhasebeleştirme ve Teknik Analizler

olarak tespit edilmiştir (Şekil 13). Akıllı Sözleşmelerin araştırma sayısı az olmasına rağmen kripto paralardan daha fazla alanda uygulandığı görülmektedir (Şekil 14).



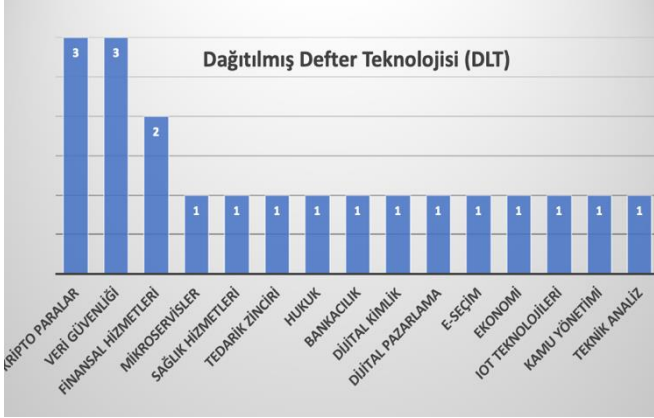
Şekil 13- Blok Zincir Kripto Teknolojisinin Uygulama Alanları Sayıları (Number of Application Areas of Blockchain Crypto Technology)



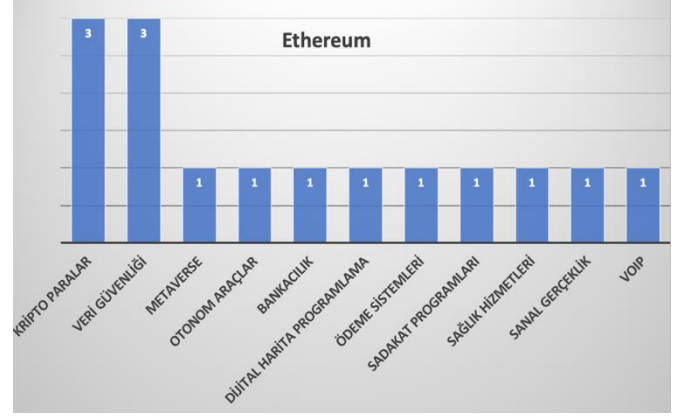
Şekil 14- Akıllı Sözleşmeler Teknolojisinin Uygulama Alanları Sayıları (Number of Application Areas of Smart Contracts Technology)

Şekil 15'te Dağıtılmış Defter Teknolojisi (DLT) ile ilgili, şekil 16'da Ethereum ile ilgili, şekil 17'de Bitcoin ile ilgili ve şekil 18'de Hyperledger Fabric teknolojileri ile ilgili

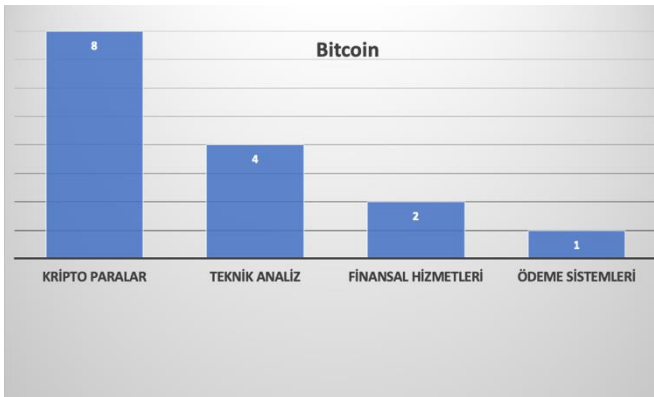
yapılan tez çalışmalarının alanlarına göre dağılımları gösterilmiştir.



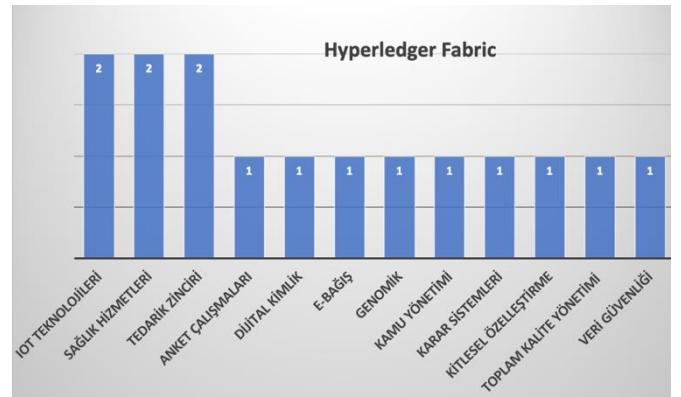
Şekil 15- Dağıtılmış Defter Teknolojisinin (DLT) Uygulama Alanları Sayıları (Number of Distributed Ledger Technology (DLT) Application Areas)



Şekil 16- Ethereum Teknolojisinin Uygulama Alanları Sayıları (Number of Application Areas of Ethereum Technology)



Şekil 17- Bitcoin Teknolojisinin Uygulama Alanları Sayıları (Number of Application Areas of Bitcoin Technology)



Şekil 18- Hyperledger Fabric Teknolojisinin Uygulama Alanları Sayıları (Number of Application Areas of Hyperledger Fabric Technology)

Elde edilen sonuçlara göre;

- İstanbul Teknik Üniversitesinden Doç. Dr. Enver Özdemir blok zincir alanında en fazla tez danışmanlığında (5) bulunmuştur.
- 2017 ve 2023 yılları arasında 401 yüksek lisans tezi, 75 doktora tezi yazılmıştır. 2021 de 95 tez yazılırken, 2022 senesinde 128 tez ve 2023 senesinde 126 tez yazılarak iki senede hemen hemen aynı oranda tez yazıldığı anlaşılmıştır.
- Marmara üniversitesi (30) ve Bahçeşehir üniversitesi (29) tezle en çok tez çalışması yapan üniversiteler olmuştur.
- Sosyal bilimler enstitüsü 183, Fen bilimleri enstitüsü 144 tez ile en çok tez çalışması yapılan enstitüler olmuşturlar.
- Tezlerin 319'u Türkçe yazılırken 157'si İngilizce yazılmıştır.
- Bilgisayar Mühendisliği 108 tez ile en çok tez yazılan ana bilim dalı olurken, onu 91 tez ile İşletme ana bilim dalı, Finans (39), Bilişim Sistemleri (38), Hukuk (33), Uluslararası Ticaret ve Lojistik (32), İktisat (31), Elektrik ve Elektronik Mühendisliği (19), Endüstri

mühendisliği (18) ana bilim dallarında tez çalışmaları olduğu görülmüştür.

- Tezlerin 319'u Türkçe dilinde yazılırken, 157 tanesi İngilizce dilinde hazırlanmıştır.
- Anahtar kelime bulutu içerisinde "Blok zincir, blockchain, kripto para, akıllı sözleşmeler, tedarik zinciri, smart contracts" en fazla kullanılanlardır.

## 6. SONUÇ VE DEĞERLENDİRME (CONCLUSION AND EVALUATION)

Çalışmada blok zincir üzerine yazılan 476 adet yüksek lisans ve doktora tezi incelenmiştir. Tezler sınıflandırma bilgileri ile tez alanlarına ve içeriğine göre olmak üzere 2 başlık altında incelenmiştir.

Sınıflandırma bilgilerine göre; İstanbul Teknik Üniversitesi öğretim üyesi Enver Özdemir yüksek lisans çalışmaları ile öne çıkmaktadır. Enver Özdemir çalışmalarında veri transferinde, mobil cihazlarda, video akışı sırasında, eşler arası medya dağıtımında, IoT ortamlarında ve sağlık sisteminde blockchain uygulamaları ve bunların güvenilirliği üzerine çalışmalar yapmıştır.

Marmara üniversitesi ve Bahçeşehir üniversitesinde yüksek sayıda tez çalışması yapılmıştır. Marmara, Maltepe ve Sakarya üniversitelerinde 4'er adet doktora tezi çalışması yapılmıştır. Marmara Üniversitesinde Endüstri Mühendisliği, işletme, bankacılık ile gazetecilik, Maltepe Üniversitesinde disiplinlerarası iletişim bilimleri, bilgisayar mühendisliği, halkla ilişkiler ve tanıtım, uluslararası ticaret ve lojistik, Sakarya üniversitesinde endüstri mühendisliği ve bilgisayar ve bilişim anabilim dallarında ikiye tane olmak üzere doktora tezleri yazılmıştır. Marmara üniversitesinde 26 yüksek lisans 4 doktora tezi çalışması bulunmaktadır. Bahçeşehir üniversitesinde 28 yüksek lisans 1 doktora tezi çalışması yapılmıştır.

Tezlerin 319'u Türkçe 157'si İngilizce dilinde yazılırken 401 yüksek lisans 75 doktora tezi hazırlanmıştır. Tez çalışmalarında 2019 yılında hızlı bir yükselme olurken 2020 de bir düşüş görülmüştür ancak pandeminin de etkisiyle 2021 2022 2023 yıllarında tez sayılarında artış olmuştur. Son yıllarda sadece dijital para tarafı değil, diğer alanlarda da blok zincirin uygulanmasıyla yeni çalışmalar yapılmaya başlandığı görülmüştür. Fen bilimleri enstitüsünde 2018 (6), 2019 (27), 2020 (19), 2021 (28), 2022 (33), 2023 (31) tez çalışması yapılırken, Sosyal bilimler enstitüsünde pandemiye kadar 2017 (1), 2018 (4), 2019 (24), 2020 (23), tez çalışması yapılırken pandemiden sonra 2021 (32), 2022 (45), 2023 (54) tez çalışması yapılarak blok zincirin değişik alanlarda uygulanmaya başladığı görülmüştür. "Blok zincir, blockchain, kripto para, akıllı sözleşmeler, tedarik zinciri, smart contracts" en fazla kullanılan anahtar kelimelerdir.

Tez alanlarına ve içeriğine göre baktığımızda; Bilgisayar Mühendisliği Ana Bilim Dalı en çok tez çalışması yapılan Ana Bilim Dalı olmuştur. Bilişim sistemleri, Elektrik Elektronik Mühendisliği, Bilgi Güvenliği Mühendisliği, Siber Güvenlik, Bilgi Teknolojileri, Enformatik ana bilim dallarında Blok zincir genel yapısı, akıllı sözleşmeler Dağıtık Defter Teknolojisi (DLT), Hyperledger Fabric, Ethereum teknolojilerini kullanan ve Veri Güvenliği, IoT Teknolojileri, Kripto Paralar, Dijital Kimlik, Sağlık Hizmetleri, E-Seçim(oylama), Finansal Hizmetleri uygulama alanlarında hazırlanmış tezler öne çıkmaktadır. İşletme, Finans, Hukuk, Uluslararası Ticaret ve Lojistik, İktisat, Endüstri Mühendisliği, Bankacılık, Maliye, Siyaset Bilimi ve Kamu Yönetimi, İnşaat Mühendisliği, Pazarlama, Matematik, Gazetecilik, ve Afet Yönetimi ana bilim dallarında blok zincir, kripto, akıllı sözleşmeler teknolojileri kullanılarak Kripto Paralar, Tedarik Zinciri, Finansal Hizmetleri, Hukuk, Vergilendirme, Bankacılık, İnşaat Yönetimi, Dijital Dönüşüm alanlarında etkilerinin araştırıldığı tezler bulunmaktadır.

Bu çalışmamız ile blok zincir üzerine yapılacak olan yeni çalışmalar için bir referans hazırlanmaya çalışıldı. Yükseköğretim kurulu tez tez merkezinde bulunan yüksek lisans ve doktora tezleri incelenerek, blok zincir üzerine yapılan çalışmalarda danışmanlar, üniversiteler, tez türleri, yazıldığı yıllar, enstitüler, yazıldığı dil ile ilgili sınıflandırma bilgileri verilmiştir. İkinci bölümde ise tezlerin yazıldığı ana bilim dalları ve bu ana bilim

dallarında yazılan tezlerin hangi konuları içerdiğini gösteren bir analiz yapılmıştır.

Blok zincir kripto paralar gibi sadece bir alanda uygulanıyormuş gibi düşünülse de aslında birçok alanda çalışmalar yapılmaya başlanmıştır. Özellikle akıllı sözleşmeler ile blok zinciri alt yapısı kullanılarak daha güvenilir yapılar kurulmaktadır. Böylelikle uzun süren işlemler kısa sürelerde gerçekleştirilmektedir. Blok zincirinin sağladığı kolaylıklar birçok alanda kullanımının yaygınlaşacağını göstermektedir.

#### KAYNAKLAR (REFERENCES)

- [1] Ü. Gökhan ve Ç. Uluyol, "Blok Zinciri Teknolojisi", *Bilişim Teknolojileri Dergisi*, c. 13, sy 2, 2020.
- [2] Y. Doğan, T. Gürdal, ve U. Uygun, "Blok Zincir Alanında Yazılan Lisansüstü Tezlerin Bibliyometrik Analizi: 2017-2022 Dönemi.", *Uluslararası Ekonomik Araştırmalar Dergisi*, c. 9, sy 1, ss. 9-26, 2023.
- [3] A. Gorkhali, L. Li, ve A. Shrestha, "Blockchain: a literature review", *Journal of Management Analytics*, c. 7, sy 3, ss. 321-343, 2020.
- [4] S. Nakamoto, & Bitcoin, A. "A peer-to-peer electronic cash system. Bitcoin".-URL: <https://bitcoin.org/bitcoin.pdf>, 4(2), 15, 2008.
- [5] C. Herweijer, D. Waughray, ve S. Warren, "Building block (chain) s for a better planet.", [http://www3.weforum.org/docs/WEF\\_Building-Blokzincirs.pdf](http://www3.weforum.org/docs/WEF_Building-Blokzincirs.pdf).
- [6] S. Kaya ve M. Turgut, "Blockchain Technology in Supply Chain", *The Journal of International Scientific Researches*, c. 4, sy 2, ss. 121-134, 2019.
- [7] J. Yli-Huumo, D. Ko, S. Choi, S. Park, ve K. Smolander, "Where Is Current Research on Blockchain Technology?—A Systematic Review", *PLoS One*, c. 11, sy 10, s. e0163477, 2016.
- [8] S. Tönnissen ve F. Teuteberg, "Analysing the impact of blockchain-technology for operations and supply chain management: An explanatory model drawn from multiple case studies", *Int J Inf Manage*, c. 52, s. 101953, 2020.
- [9] X. Xu vd., "The blockchain as a software connector", *Proceedings - 2016 13th Working IEEE/IFIP Conference on Software Architecture, WICSA 2016*, ss. 182-191, Tem. 2016.
- [10] A. I. Özdemir vd., "The role of blockchain in reducing the impact of barriers to humanitarian supply chain management", *International Journal of Logistics Management*, c. 32, sy 2, ss. 454-478, 2020.
- [11] Z. Zheng, S. Xie, H. N. Dai, X. Chen, ve H. Wang, "Blockchain challenges and opportunities: A survey", *International Journal of Web and Grid Services*, c. 14, sy 4, ss. 352-375, 2018.
- [12] A. Anjum, M. Sporny, ve A. Sill, "Blockchain Standards for Compliance and Trust", *IEEE Cloud Computing*, c. 4, sy 4, ss. 84-90, 2017.
- [13] A. S. Rajasekaran, M. Azees, ve F. Al-Turjman, "A comprehensive survey on blockchain technology", *Sustainable Energy Technologies and Assessments*, c. 52, s. 102039, 2022.
- [14] S. S. Kamble, A. Gunasekaran, ve R. Sharma, "Modeling the blockchain enabled traceability in agriculture supply chain", *Int J Inf Manage*, c. 52, s. 101967, 2020.
- [15] T. J. MacDonald, D. W. E. Allen, ve J. Potts, "Blockchains and the boundaries of self-organized economies: Predictions for the future of banking", *New Economic Windows*, ss. 279-296, 2016.
- [16] M. Iansiti, K. R. Lakhani, ve H. Mohamed, "The Truth about Blockchain", *Harvard Business School*, 2017.

- [17] S. Yousefi ve B. Mohamadpour Tosarkani, "An analytical approach for evaluating the impact of blockchain technology on sustainable supply chain performance", *Int J Prod Econ*, c. 246, s. 108429, 2022.
- [18] A. G. Gad, D. T. Mosa, L. Abualigah, ve A. A. Abohany, "Emerging Trends in Blockchain Technology and Applications: A Review and Outlook", *Journal of King Saud University - Computer and Information Sciences*, c. 34, sy 9, ss. 6719-6742, 2022.
- [19] X. Yue, H. Wang, D. Jin, M. Li, ve W. Jiang, "Healthcare Data Gateways: Found Healthcare Intelligence on Blockchain with Novel Privacy Risk Control", *J Med Syst*, c. 40, sy 10, ss. 1-8, 2016.
- [20] H. Pavlić Skender ve P. A. Zaninović, "*Perspectives of Blockchain Technology for Sustainable Supply Chains*", ss. 77-92, 2020.
- [21] H. F. Atlam, M. A. Azad, A. G. Alzahrani, ve G. Wills, "A Review of Blockchain in Internet of Things and AI", *Big Data and Cognitive Computing*, Vol. 4, Page 28, c. 4, sy 4, s. 28, 2020.
- [22] A. Aoun, A. Ilinca, M. Ghandour, ve H. Ibrahim, "A review of Industry 4.0 characteristics and challenges, with potential improvements using blockchain technology", *Comput Ind Eng*, c. 162, s. 107746, 2021.
- [23] S. Çizmecioglu ve A. Z. Akman, "Blok zincir ve kripto para konularının bibliyometrik bir analizi: 2015-2020 Dönemi", *Business Economics and Management Research Journal*, c. 4, sy 1, ss. 1-16, 2021.
- [24] M. Weber vd., "Blockchain Kripto Varlık Bitcoin Üzerine Bibliyometrik Bir Analiz Web Of Science Uygulaması", *Uluslararası Muhasebe ve Finans Araştırmaları Dergisi*, c. 5, sy 1, ss. 44-72, 2023.
- [25] K. Karakılıç ve S. Apak, "Blokzincir Teknolojisi Alanında Yayımlanan Lisansüstü Tez Çalışmalarının Bibliyometrik Analizi: Ulusal Tez Merkezi Örneği", *Bitlis Eren Üniversitesi Sosyal Bilimler Dergisi*, c. 11, sy 2, ss. 324-339, 2022.
- [26] T. T. Galip, "Türkiye'de Yazılan Kripto Para Temalı Doktora Tezlerinin Bibliyometrik Analizi", *Article in Journal of History School*, 2023.
- [27] B. Güvenol vd., "Turizm Ve Blok Zinciri Kavramlarına Yönelik Bibliyometrik Bir Analiz", *Journal of Business in The Digital Age*, c. 6, sy Özel Sayı, ss. 66-81, 2023.
- [28] Ş. Arslan ve S. Çankaya, "COVİD-19 Pandemisinin Bireysel Yatırımcı Davranışlarına Etkisi", *Research of Financial Economic and Social Studies*, c. 7, sy 4, ss. 600-611, 2022.

# CRITIC ve MABAC Yöntemleri İle Türkiye’de Yazılım Sektörünün Performans Analizi

*Araştırma Makalesi/Research Article*

 Şeyma Nur AYDIN<sup>1</sup>,  Ali SEVİNÇ<sup>2</sup>

<sup>1</sup>Bağımsız Araştırmacı, Ankara, Türkiye

<sup>2</sup>KOSGEB, Ankara, Türkiye

[seymanuraydin125@gmail.com](mailto:seymanuraydin125@gmail.com), [alisevinc71@gmail.com](mailto:alisevinc71@gmail.com)

(Geliş/Received:13.06.2024; Kabul/Accepted:25.09.2024)

DOI: 10.17671/gazibtd.1500545

**Özet**— Yazılım sektörü işletmelere, teknolojik gelişmeye ve genç istihdama olan katkılarından dolayı son yıllarda tüm endüstriler için önemli bir noktaya gelmiştir. Özellikle Türkiye gibi gelişmekte olan ülkelerde, küresel rekabet ortamında güçlü olmak için yazılım firmalarının desteğine olan ihtiyaç giderek artmaktadır. Bu çalışmada, 2018-2022 yılları arasındaki finansal veriler kullanılarak yazılım sektörünün Türkiye ortalaması ve İBBS Düzey I ortalaması ile karşılaştırmalı bir performans değerlendirmesi yapılmıştır. Performans ölçütlerinin ağırlıklandırılması için objektif bir metot olan CRITIC yöntemi (Criteria Importance Through Intercriteria Correlation), yıllara göre performans sıralaması için ise MABAC yöntemi (The Multi-Attributive Border Approximation Area Comparison) kullanılmıştır. Elde edilen bulgular, İBBS Düzey I bölgelerinde yazılım sektörü performansının genel olarak 2020 yılında yüksek, 2019 yılında ise daha düşük olduğunu göstermiştir. Türkiye ortalaması sonuçlarına göre de en iyi performans 2020 yılında, en zayıf performans 2018 yılında gerçekleşmiştir.

**Anahtar Kelimeler**— yazılım sektörü, istatistiki bölge birimleri sınıflandırması, performans analizi, çok kriterli karar verme, critic, mabac

## Performance Analysis of the Software Industry in Turkey with CRITIC and MABAC Methods

**Abstract**—The software industry has reached an important point for all industries in recent years due to its contributions to businesses, technological development and young employment. Especially in developing countries such as Turkey, the need for the support of software companies is increasing in order to be strong in the global competitive environment. In this study, a comparative performance evaluation of the software industry with Turkey's average and NUTS Level I average was made using financial data between 2018-2022. The CRITIC method (Criteria Importance Through Intercriteria Correlation), an objective method, was used to weight performance criteria, and the MABAC method (The Multi-Attributive Border Approximation Area Comparison) was used to rank performance by years. The findings showed that the software industry performance in NUTS Level I regions was generally high in 2020 and lower in 2019. According to Turkey average results, the best performance occurred in 2020 and the weakest performance occurred in 2018.

**Keywords**— software industry, the nomenclature of territorial units for statistics, performance analysis, multi-criteria decision making, critic, mabac

## 1. GİRİŞ (INTRODUCTION)

Yazılım sektörü, tüm sektörleri etkilemesi, kuruluşlara rekabet avantajı sağlaması, toplumsal kalkınma imkanı sunması gibi sebeplerden dolayı kritik öneme ve değere sahip bir konuma gelmiştir. Özellikle ülkemiz gibi gelişmekte olan ve genç nüfusun fazla olduğu ülkeler için yazılım sektörünün çok sayıda fırsatı içerisinde barındırdığı söylenebilir [1]. Türkiye sanayisi başta olmak üzere iletişim, ulaştırma, enerji, alt yapı gibi birçok sektörün gelişimine ivme kazandıran lokomotif bir sektör olan yazılım sektörü, bu yönüyle ülke gelişimi için katma değer ve istihdam yaratan doğal bir kaldıraç görevi görmektedir [2]. Bu sebeple bilişim alanında sağlanacak katma değerler ile gelişmiş ülkeler seviyesine ulaşmak ve küresel rekabet koşullarında daha güçlü bir Türkiye'ye sahip olmak için, kalkınma hamlesinde bilişimin stratejik önemi oldukça yüksektir [3].

Bu araştırmada, çok ölçütlü karar verme yöntemleri kullanılarak Türkiye'deki yazılım sektörünün 2018-2022 yılları arasındaki performans değerlendirilmesi yapılmıştır. İncelenen literatürde Türkiye'de bilişim sektörüne yönelik performans çalışmaları mevcuttur [4,5,6]. Ancak doğrudan yazılım sektörü performansını belirlemeye yönelik herhangi bir araştırmaya rastlanmamıştır. Ayrıca bu çalışmada, sektör performansı yıllara göre değerlendirilmiştir ve istatistiki bölge birimleri sınıflandırması dikkate alınmıştır. Türkiye'nin genel ekonomik performansı ile bölgesel kalkınma eğilimlerini anlamak ve yıllara göre değerlendirmeler yapmak için, araştırmada elde edilen bulguların yazıma katkı sağlayacağı düşünülmektedir.

## 2. YAZILIM SEKTÖRÜ (SOFTWARE INDUSTRY)

Yazılım hizmetlerinin sağlanmasını üstlenen ilk firmalar, 1950'li yıllardan kalma programlama firmalarıdır. O yıllarda bu firmaların görevleri, bilgisayar kullanıcıları ve üreticileri için türünün tek örneği olan programlar geliştirmek olmuştur. Paket yazılım tedarikçileri olarak adlandırılan ikinci grup, 1960'lı yıllarda kamu ve özel sektör kuruluşlarındaki bilgisayar kullanıcılarına yönelik bazı programlar geliştirmek için ortaya çıkmıştır. Paket yazılım tedarikçileri, bilgisayar üreticileriyle doğrudan rekabet halinde faaliyet göstererek sermaye malları sektöründe yer alan firmaların özelliklerini geliştirmişlerdir. 1960'ların sonlarında ekonomik ve teknolojik avantajlar ile pazar fırsatlarının birleşmesi sonucu paket yazılım endüstrisi yükselişe geçmiştir. 1960'larda "üçüncü nesil" teknolojinin var olması ile ekonomik fırsatlar ortaya çıkmıştır ve bilgisayar performansı iki kat artmıştır. Bu dönemde yazılım verimliliğinin de neredeyse iki-üç katı arttığı görülmüştür. Yine aynı dönemde büyük firmaların dışında pek çok bilgisayar kullanıcısı teknolojik kapasiteye sahip olsalar dahi geliştirme maliyetlerinin yüksek olması sebebiyle büyük programlar ile bilgisayarların potansiyelinden yararlanmanın mümkün olmadığını fark etmişlerdir. Bu durum yazılım hizmeti sağlayıcılara, maliyeti düşük programlar geliştirme fırsatı yaratmıştır [7].

1970'li yıllarda üçüncü grup olarak adlandırılan kişisel bilgisayar yazılım tedarikçilerinin önemli bir konuma geldiği bilinmektedir. Yazılım endüstrisi, ilk olarak IBM'in 1969'daki ayrıştırma kararının (donanım ve yazılımı ayrıştırma) ardından ve kişisel bilgisayarın yükselişe geçmesi sonucu 1970'lerde önemli ölçüde büyümeye başlamıştır [7]. Bu gelişmelerle birlikte, başlangıçta bilgisayar donanımı üretiminin ana faaliyet alanına hizmet eden ve ikincil bir sektör olarak görülen yazılım geliştirme, zamanla bağımsız bir endüstri haline gelmiştir [8]. 1979 yılına doğru ABD yazılım firmalarının yıllık satışları yaklaşık 2 milyar dolara ulaşmıştır. 1980'lerde yazılım endüstrisinde yılda % 20 ve üzeri büyüme oranları görülmüştür. Böylece ABD firmalarının yıllık gelirleri 1982'de 10 milyar dolara, 1985'te ise 25 milyar dolara ulaşarak 1979 yılındaki rakamın on katına çıkmıştır [7]. 1990'lı yıllarda ise G7 ülkeleri dışındaki bazı ekonomilerde yazılım sektörünün olağanüstü bir şekilde büyümesi dünyada oldukça beklenmedik bir olay olarak görülmüştür. Bunun sebebi, bu ülkelerin genel olarak ileri teknoloji endüstrisinin büyüebileceği bir konumda olmamalarıdır. Bu yıllarda Hindistan, İrlanda ve İsrail gibi ülkeler önemli yazılım ihracatçıları olarak ortaya çıkmıştır. Bununla birlikte Çin ve Brezilya gibi ülkeler büyük bir kısmı iç pazara dayalı olan kapsamlı bir yazılım endüstrisi geliştirmişlerdir ve 2000'li yıllarda ihracata yönelmeye başlamışlardır [9]. Yazılım endüstrisi 2000'li yıllarda ABD ekonomisinde de büyüme kazanmıştır. 1997'den 2012'ye kadar yazılım endüstrisi üretimi 149 milyar dolardan 425 milyar dolara çıkmıştır ve yazılım sektörünün ABD GSYİH içerisindeki doğrudan payı % 1,7'den % 2,6'ya, ulaşmıştır. Yazılım kullanımı aynı zamanda diğer endüstrilerin verimliliğini de artırarak büyüme katkısında bulunmuştur [10].

2000'li yıllar yazılım sektörünün tüm dünyada gelişmesinde ivme kazandığı bir dönem olmuştur. Bu yıllarda özellikle bazı firmaların yazılım sektörüne sağladıkları katkı açısından öne çıktığı görülmüştür. Örneğin, bugün dünyanın en büyük kitapçısı olarak bilinen yazılım şirketi Amazon, neredeyse her şeyi çevrimiçi olarak satabilen, perakende mağazaya gerek duymayan muhteşem bir yazılım motoru olarak faaliyet göstermektedir. Yine yazılım sektöründeki gelişmelerde önemli bir payı olan Netflix, abone sayısına göre günümüzün en büyük video hizmetlerini sunan başka bir yazılım şirkettir. Comcast ve Time Warner gibi firmalar, içeriği fiziksel kablodan kurtarıp akıllı telefonlara ve tabletlere bağlayan TV Everywhere gibi çabalarla kendilerini yazılım şirketlerine dönüştürmüşlerdir. Bugün, başka bir alan olan müzik sektörüne hakim olan firmalarında, Apple'ın iTunes, Spotify ve Pandora'sı gibi yazılım şirketleri olduğu bilinmektedir. Dijital kanallardan elde edilen sektör geliri 2010 yılında toplam 4,6 milyar dolar olarak gerçekleşmiş ve 2004'teki %2 seviyesinden toplam gelirin %29'una yükselmiştir. Günümüzün en hızlı büyüyen eğlence şirketleri yine yazılım sektöründe olan video oyunu yapımcılarıdır. Gelişimi ile en çok dikkat çeken büyük video oyunu şirketi, oyunlarını tamamen çevrimiçi olarak sunan Zynga'dır. Zynga'nın ilk çeyrek geliri 2011 yılında 235 milyon dolara yükselerek bir önceki yılın iki katı gelir

elde edilmiştir. Yazılımın etkisi altına aldığı bir başka sektör de fotoğrafçılıktır. Shutterfly, Snapfish ve Flickr gibi şirketler yazılımdaki üstünlükleri ile Kodak'ın yerine geçmişlerdir. Günümüzün en büyük doğrudan pazarlama platformu da bir yazılım şirketi olan Google'dır. Groupon, Living Social, Foursquare gibi firmalarda pazarlama perakende sektörüne katılarak yemek sektörü için yazılım kullanan firmalar olmuşlardır. Groupon, yalnızca iki yıl faaliyet gösterdikten sonra 2010 yılında 700 milyon doların üzerinde gelir elde etmiştir. Günümüzün en hızlı büyüyen telekom şirketi, Microsoft tarafından 8,5 milyar dolara satın alınan bir yazılım şirketi olan Skype'tır. Yazılımda dikkatleri üzerine çeken bir diğer firma da günümüzün en hızlı büyüyen işe alım şirketi olan LinkedIn olmuştur. [11]. Bahsi geçen şirketlerin uygulamaları, yazılımın son yıllarda tüm sektörleri ele geçiren ve devamlı gelişmekte olan bir sektör olduğunu göstermektedir.

### 3. LİTERATÜR TARAMASI (LITERATURE REVIEW)

İncelenen literatürde, çok kriterli karar verme yöntemlerinin farklı sektörlerin performans analizi için sıklıkla kullanıldığı görülmektedir. Özellikle bankacılık, sigortacılık, lojistik, turizm, bilişim, sağlık, enerji ve ulaştırma gibi sektörlerin performans değerlendirmesinde pek çok araştırmacı bu yöntemlerden yararlanmıştır.

Tayvan'da halihazırda geliştirilen altı yüksek teknoloji endüstrisinin karşılaştırmalı performanslarını değerlendirmek için çok girdili-çoklu çıktı değerlendirme aracı olan Veri Zarflama Analizi (VZA) kullanılmıştır [12]. Türkiye'de hayat dışı sigorta şirketlerinin ve hayat/emeklilik şirketlerinin sektörel bazda 2007-2011 yılları arasındaki finansal performanslarını değerlendirmek için TOPSIS yönteminden yararlanılmıştır [13]. Tayvan'daki varlık yönetimi bankalarının karşılaştırmalı performans değerlendirmesi için AHP ve VZA 'dan (Veri Zarflama Analizi) faydalanılmıştır [14]. Hindistan'daki hem kamu hem de özel sektörde faaliyet gösteren ticari bankaların mali sağlığına ulaşmaları için TOPSIS ve M-TOPSIS yöntemleri ile performans değerlendirmesi yapılmıştır [15]. Türkiye'de faaliyet gösteren mevduat bankalarının, 2009-2013 yılları arasındaki finansal performanslarının değerlendirilmesinde, VIKOR yönteminden yararlanılmıştır [16]. Aktif büyüklüklerine göre küçük, orta ve büyük ölçekli bankaların performansları, ENTROPI ve WASPAS kullanılarak analiz edilmiştir [17]. Lojistik sektöründe faaliyette bulunan ve Türkiye'nin en büyük 500 işletmesi içerisinde yer alan dokuz işletmenin finansal performansını değerlendirmek için Gri İlişkisel Analiz (GİA) yöntemi kullanılmıştır [18]. Borsa İstanbul'a kayıtlı on turizm firmasının, 2015 yılına ait finansal performanslarının analiz edilmesi için Gri İlişkisel Analiz (GİA) yönteminden faydalanılmıştır [19]. "Forbes" dergisinde açıklanan "Global 2000" listesinde, en büyük şirketler arasında yer almış bilişim teknolojisi sektöründeki bilgisayar donanım firmalarının performans değerlendirmesi, ENTROPI ve VIKOR yöntemleri ile yapılmıştır [4]. Sağlık sektöründe faaliyet

gösteren şirketlerin iş sağlığı ve güvenliği performansının ölçülmesi için AHP, PROMETHEE ve GRA yöntemleri kullanılmıştır [20]. Borsa İstanbul A. Ş. Elektrik endeksinde (XELKT) bulunan ve 2016-2017 yıllarında hisse senetleri borsada işlem gören beş şirketin finansal performansları, WASPAS yöntemi ile değerlendirilmiştir [21]. BIST ulaştırma ve depolama sektörü şirketlerinin finansal performanslarının ölçülmesinde, 2015-2019 dönemlerini kapsayan finansal tablo verileri kullanılarak SWARA ve OCRA yöntemlerinden faydalanılmıştır [22]. Doğu Avrupa'daki yedi ülkenin sağlık sektörlerinin performansı BWM, LBWA ve CoCoSo yöntemleri ile değerlendirilmiştir [23]. Sırbistan'daki ticaret şirketlerinin performansını kapsamlı bir şekilde analiz etmek ve geleceğe dönük iyileştirmeler yapmak için bulanık kümeler ile WASPAS yönteminden yararlanılmıştır [24].

### 4. İSTATİSTİKİ BÖLGE BİRİMLERİ SINIFLANDIRMASI

(THE NOMENCLATURE OF TERRITORIAL UNITS FOR STATISTICS)

1970'li yıllarda Avrupa Topluluğu İstatistik Ofisi (Eurostat) tarafından, Avrupa Birliği'nde bölgesel istatistiklerin tek bir mekânsal sınıflandırmaya göre yapılması için İstatistikî Bölge Birimleri Sınıflandırması (*The Nomenclature of Territorial Units for Statistics (NUTS)*) geliştirilmiştir. Bu istatistikî sınıflandırma yöntemi, 1988 yılından itibaren AB mevzuatında yer almıştır. Türkiye'de ise 2001 yılında İstatistikî Bölge Birimleri Sınıflandırması tanımlanmış ve 28 Ağustos 2002 tarihli Bakanlar Kurulu Kararı ile yürürlüğe girerek kullanılmaya başlanmıştır. Bu sınıflandırma sonucu, iller düzey 3 seviyesi olarak kabul edilmiş, komşu illerin gruplandırılması ile de 26 Düzey II ve 12 Düzey I Bölgeleri oluşturulmuştur. Tablo 1'de Düzey I bölgeleri gösterilmiştir [25].

Tablo 1. Türkiye'de İstatistikî Bölge Birimleri Sınıflandırması (İBBS Düzey I)  
(The Nomenclature of Territorial Units for Statistics in Turkey (NUTS I))

Düzey I Kod	Düzey I Bölge Adı
TR1	İstanbul
TR2	Batı Marmara
TR3	Ege
TR4	Doğu Marmara
TR5	Batı Anadolu
TR6	Akdeniz
TR7	Orta Anadolu
TR8	Batı Karadeniz
TR9	Doğu Karadeniz
TRA	Kuzeydoğu Anadolu
TRB	Ortadoğu Anadolu
TRC	Güneydoğu Anadolu

### 5. METODOLOJİ (METHODOLOGY)

Bu çalışmada, 2018-2022 dönemine ait finansal veriler kullanılarak yazılım sektörünün Türkiye ortalaması ve İBBS Düzey I (12 bölge) ortalaması ile karşılaştırmalı bir performans analizi yapılmıştır. Performans ölçütlerinin ağırlıklandırılması için objektif bir metod olan CRITIC

yöntemi, yıllara göre performans sıralaması için ise MABAC yöntemi kullanılmıştır.

### 5.1. CRITIC Yöntemi (CRITIC Method)

Kriter ağırlıklarının belirlenmesi, çok ölçütlü karar problemleri için önemli adımlardan biridir. Kriter ağırlıklarını belirlemeye yönelik yaklaşımlar, subjektif ve objektif yöntemler olarak ikiye ayrılmaktadır. Subjektif yöntemlerde kriter ağırlıkları, karar sürecine dahil olan uzmanların görüşleri doğrultusunda belirlenir. Bu yöntemler, karar vericilerin sezgilerini ve görüşlerini yansıttığından karar verme süreci doğrudan bu görüş ve sezgilerden etkilenir. Objektif yöntemlerde ise karar vericilerin görüş ve sezgileri dikkate alınmaz. Bir başlangıç karar matrisi kullanılarak kriter ağırlıkları hesaplanır [26]. Subjektif ağırlıklandırma metodları, sonuçların güvenilirliği açısından şüpheye yol açabildiğinden, bu tür sorunların üstesinden gelmek için objektif ağırlıklandırma yöntemleri tercih edilir [27].

1995 yılında geliştirilen CRITIC yöntemi, başlangıç karar matrisindeki verilere dayanan bir objektif kriter ağırlıklandırma yaklaşımıdır [28]. Yöntemde, her değerlendirme kriteri için alternatifler arasında ölçülen değerlerin üyelik fonksiyonu şeklinde ifade edilen bir dizi varyasyon vardır. Kullanılan kriterlere göre oluşturulan bileşenlerin her biri standart sapma gibi istatistiksel parametrelere sahiptir. Bu parametreler ilgili standart değerler arasındaki farkın derecesini temsil eder [29]. Kriterlerin kontrastını belirlemek için, sütun başına değişkenlerin standartlaştırılmış kriter değerlerinin standart sapmalarının yanı sıra tüm sütun çiftlerinin korelasyon katsayıları kullanılır [30].

CRITIC yönteminin işlem adımları aşağıda sıralanmıştır [28, 31, 32]:

**1.Adım: Karar Matrisi Oluşturma:** Bu yöntemin ilk aşaması, Eşitlik (1)'de gösterilen başlangıç karar matrisinin oluşturulmasıdır.

$$x = \begin{matrix} A_1 \\ A_2 \\ \vdots \\ A_m \end{matrix} \begin{bmatrix} x_{11} & x_{12} & \dots & x_{1n} \\ x_{21} & x_{22} & \dots & x_{2n} \\ \vdots & \vdots & \dots & \vdots \\ x_{m1} & x_{m2} & \dots & x_{mn} \end{bmatrix} \quad (1)$$

**2.Adım: Matrisin Normalizasyonu:** Yöntemin ikinci adımında, başlangıç matrisi normalize edilir. Normalizasyon işleminde fayda kriterleri için Eşitlik (2), maliyet kriterleri için ise Eşitlik (3)'ten yararlanılır.

$$r_{ij} = \frac{x_{ij} - x_j^{\min}}{x_j^{\max} - x_j^{\min}} \dots \dots \dots j = 1, 2, \dots, n \quad (2)$$

$$r_{ij} = \frac{x_j^{\max} - x_{ij}}{x_j^{\max} - x_j^{\min}} \dots \dots \dots j = 1, 2, \dots, n \quad (3)$$

**3.Adım: İlişki Katsayı Matrisinin ( $P_{jk}$ ) Elde Edilmesi:** Bu aşamada, karar kriterleri arasındaki ilişki seviyesinin belirlenmesi için Eşitlik (4)'ten yararlanılarak korelasyon katsayıları hesaplanır.

$$\rho_{jk} = \frac{\sum_{i=1}^m (r_{ij} - r_j^-) \cdot (r_{ik} - r_k^-)}{\sqrt{\sum_{i=1}^m (r_{ij} - r_j^-)^2 \cdot \sum_{i=1}^m (r_{ik} - r_k^-)^2}} \quad j, k = 1, 2, \dots, n \quad (4)$$

**4.Adım:  $C_j$  Değerlerini Hesaplama:** Yöntemin dördüncü adımında, tüm kriterlerde bulunan toplam bilgi miktarı ( $C_j$ ) Eşitlik (5) yardımı ile hesaplanır. Toplam bilgi miktarı kriterlerde bulunan karşılık yoğunluğu ve çelişkilerden elde edilir. Eşitlik (5) içinde bulunan standart sapma değerleri ( $\sigma$ ) Eşitlik (6) ile hesaplanır.

$$c_j = \sigma_j \cdot \sum_{k=1}^n (1 - p_{jk}) \quad j = 1, 2, \dots, n \quad (5)$$

$$\sigma_j = \sqrt{\frac{\sum_{i=1}^m (r_{ij} - r_j^-)^2}{m-1}} \quad (6)$$

**5.Adım: Kriter Ağırlıklarının ( $W_j$ ) Hesaplanması:** Yöntemin son adımında, kriterlerin ağırlıkları Eşitlik (7)'den yararlanılarak hesaplanır.

$$w_j = \frac{c_j}{\sum_{k=1}^n c_j} \quad (7)$$

### 5.2. MABAC Yöntemi (MABAC Method)

MABAC yönteminin işlem adımları aşağıda sıralanmıştır [33, 34]:

**1.Adım: Karar Matrisi Oluşturma:** Yöntemin ilk adımında başlangıç matrisi oluşturulur ve m alternatif, n kritere göre değerlendirilir. Eşitlik (8)'de başlangıç karar matrisi gösterilmiştir.

$$X = \begin{matrix} & C_1 & C_2 & \dots & C_n \\ A_1 & \begin{bmatrix} x_{11} & x_{12} & \dots & x_{1n} \\ x_{21} & x_{22} & \dots & x_{2n} \\ \dots & \dots & \dots & \dots \\ A_m & \begin{bmatrix} x_{m1} & x_{m2} & \dots & x_{mn} \end{bmatrix} \end{matrix} \end{matrix} \quad (8)$$

**2.Adım: Matrisin Normalizasyonu:** Yöntemin ikinci adımında, başlangıç matrisi normalize edilir. Eşitlik (9)'da normalize karar matrisi gösterilmiştir.

$$N = \begin{matrix} & C_1 & C_2 & \dots & C_n \\ A_1 & \begin{bmatrix} n_{11} & n_{12} & \dots & n_{1n} \\ n_{21} & n_{22} & \dots & n_{2n} \\ \dots & \dots & \dots & \dots \\ A_m & \begin{bmatrix} n_{m1} & n_{m2} & \dots & n_{mn} \end{bmatrix} \end{matrix} \end{matrix} \quad (9)$$



Bu adımda fayda kriterleri için kriterin en büyük değeri, maliyet kriterleri için de kriterin en küçük değeri tercih edilir. Fayda kriterleri için Eşitlik (10), maliyet kriterleri için ise Eşitlik (11)'den yararlanılır.

$$n_{ij} = \frac{x_{ij} - x_i^-}{x_i^+ - x_i^-} \quad (10)$$

$$n_{ij} = \frac{x_{ij} - x_i^+}{x_i^- - x_i^+} \quad (11)$$

**3.Adim: Ağırlıklı Karar Matrisi Oluşturma:** Karar matrisinin normalize edilmesinin ardından, Eşitlik (12)'den yararlanılarak ağırlıklı karar matrisi oluşturulur.

$$v_{ij} = w_i \cdot (n_{ij} + 1) \quad (12)$$

Eşitlik (12) ile oluşturulan ağırlıklı karar matrisi (V) Eşitlik (13)'te gösterilmiştir.

$$V = \begin{bmatrix} v_{11} & v_{12} & \dots & v_{1n} \\ v_{21} & v_{22} & \dots & v_{2n} \\ \dots & \dots & \dots & \dots \\ v_{m1} & v_{m2} & \dots & v_{mn} \end{bmatrix} = \begin{bmatrix} w_1 \cdot (n_{11} + 1) & w_2 \cdot (n_{12} + 1) & \dots & w_n \cdot (n_{1n} + 1) \\ w_1 \cdot (n_{21} + 1) & w_2 \cdot (n_{22} + 1) & \dots & w_n \cdot (n_{2n} + 1) \\ \dots & \dots & \dots & \dots \\ w_1 \cdot (n_{m1} + 1) & w_2 \cdot (n_{m2} + 1) & \dots & w_n \cdot (n_{mn} + 1) \end{bmatrix} \quad (13)$$

**4.Adim: Sınır Yakınlık Alanı Matrisi Elde Etme:** Bu aşamada, Eşitlik (14)'ten faydalanılarak tüm kriterler için sınır yakınlık alanı elde edilir.

$$g_i = \left( \prod_{j=1}^m v_{ij} \right)^{1/m} \quad (14)$$

Kriterler için sınır yakınlık alanı değeri hesaplandıktan sonra Eşitlik (15)'te gösterilen sınır yakınlık alanı matrisi (G) elde edilmiş olur.

$$G = \begin{bmatrix} C_1 & C_2 & \dots & C_n \\ g_1 & g_2 & \dots & g_n \end{bmatrix} \quad (15)$$

**5.Adim: Alternatiflerin Sınır Yakınlık Alanına Olan Mesafesini Hesaplama:** Bu aşamada, karar matrisi elemanlarının sınır yakınlık alanından uzaklıkları belirlenir ve Eşitlik (16)'da gösterilen Q matrisi oluşturulur.

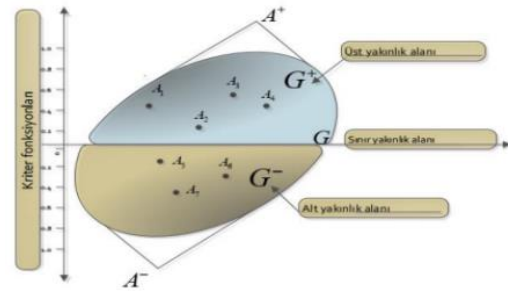
$$Q = \begin{bmatrix} q_{11} & q_{12} & \dots & q_{1n} \\ q_{21} & q_{22} & \dots & q_{2n} \\ \dots & \dots & \dots & \dots \\ q_{m1} & q_{m2} & \dots & q_{mn} \end{bmatrix} \quad (16)$$

Seçeneklerin sınır yakınlık alanına uzaklığı olan  $q_{ij}$ , ağırlıklandırılmış matristeki öğeler (V) ile sınır yakınlık alanı değeri (G) arasındaki farktır ve bunu hesaplamak için Eşitlik (17) kullanılır. Elde edilen Q matrisi Eşitlik (18)'de gösterilmiştir.

$$Q = V - G = \begin{bmatrix} v_{11} & v_{12} & \dots & v_{1n} \\ v_{21} & v_{22} & \dots & v_{2n} \\ \dots & \dots & \dots & \dots \\ v_{m1} & v_{m2} & \dots & v_{mn} \end{bmatrix} - \begin{bmatrix} g_1 & g_2 & \dots & g_n \\ g_1 & g_2 & \dots & g_n \\ \dots & \dots & \dots & \dots \\ g_1 & g_2 & \dots & g_n \end{bmatrix} \quad (17)$$

$$Q = \begin{bmatrix} v_{11} - g_1 & v_{12} - g_2 & \dots & v_{1n} - g_n \\ v_{21} - g_1 & v_{22} - g_2 & \dots & v_{2n} - g_n \\ \dots & \dots & \dots & \dots \\ v_{m1} - g_1 & v_{m2} - g_2 & \dots & v_{mn} - g_n \end{bmatrix} = \begin{bmatrix} q_{11} & q_{12} & \dots & q_{1n} \\ q_{21} & q_{22} & \dots & q_{2n} \\ \dots & \dots & \dots & \dots \\ q_{m1} & q_{m2} & \dots & q_{mn} \end{bmatrix} \quad (18)$$

Karar seçeneği, sınır yakınlık alanı olan  $G^-$ 'ye, üst yakınlık alanı olan  $G^+$ 'ya veya alt yakınlık alanı olan  $G^-$ 'ye ait olabilir.  $G^+$  alanı ideal alternatifi ( $A^+$ ) içeren alan iken,  $G^-$  alanı ideal olmayan alternatifi ( $A^-$ ) göstermektedir. Şekil 1'de sınır yakınlık alanı gösterilmiştir.



Şekil 1. Sınır Yakınlık Alanı

Karar seçeneğinin sınır yakınlık alanına ait olma durumu Eşitlik (19)'da verilmiştir.

$$A_i \in \begin{cases} G^+ & \text{if } q_{ij} > 0 \\ G & \text{if } q_{ij} = 0 \\ G^- & \text{if } q_{ij} < 0 \end{cases} \quad (19)$$

Karar probleminde herhangi bir karar seçeneğinin, en iyi alternatif olarak kabul edilmesi, kriter değerlerinin büyük çoğunluğunun üst yakınlık alanına ait olmasına bağlı bir durumdur.

**6.Adim: Seçeneklerin Sıralanması:** Yöntemin son adımında, Eşitlik (20) kullanılarak karar seçeneklerinin kriter fonksiyonları hesaplanır. En yüksek  $S_i$  değerine sahip olan seçenek, karar problemi için en iyi seçenektir.

$$S_i = \sum_{j=1}^n q_{ij}, j = 1, 2, \dots, n, i = 1, 2, \dots, m \quad (20)$$

## 6. ARAŞTIRMANIN BULGULARI (FINDINGS OF THE RESEARCH)

### 6.1. Performans Değerlendirmede Kullanılan Veri Seti ve Karar Kriterleri (Dataset and Decision Criteria Used in Performance Evaluation)

KOSGEB tarafından hem şirket hem de şahıs işletmelerine ait bilanço değerlerinin Türkiye ortalaması ve İBBS Düzey I ortalamaları hesaplanmaktadır. Bu araştırmada, yazılım sektörü Türkiye bilanço ortalamaları ve İBBS Düzey I ortalamaları veri seti, T.C. Küçük ve Orta Ölçekli İşletmeleri Geliştirme ve Destekleme İdaresi Başkanlığı'ndan (KOSGEB) temin edilmiştir [35].

Tablo 2'de yazılım sektörünün performans değerlendirmesinde kullanılan 13 finansal kriter ile kriterlerin fayda ve maliyet yönleri gösterilmiştir.

Tablo 2. Performans Değerlendirmede Kullanılan Kriterler

(Criteria Used in Performance Evaluation)

Kriter	Kod	Yön
Cari oran	K1	Fayda
FAVÖK (Amortisman Öncesi Faaliyet Karı)	K2	Fayda
Gelirler	K3	Fayda
Giderler	K4	Maliyet
İşletme Sermayesi Gün Sayısı(Gün)	K5	Maliyet
Kar-Zarar	K6	Fayda
Likidite Oranı(Asit-Test)	K7	Fayda
Nakit Oran	K8	Fayda
Net Çalışma Sermayesi	K9	Fayda
Net Satış Hasılatı	K10	Fayda
Öz Kaynakların Karlılığı	K11	Fayda
Satışların Karlılığı	K12	Fayda
Satışların Maliyeti	K13	Maliyet

### 6.2. CRITIC Yöntemi ile Kriterlerin Ağırlıklandırılması (Weighting of Criteria with CRITIC Method)

CRITIC yönteminde kriter ağırlıklandırılması, bir başlangıç karar matrisine dayanmaktadır. Bu çalışmada 12 bölge ve Türkiye ortalaması için farklı başlangıç karar matrisleri kullanıldığından her bir bölgenin kriter ağırlıklandırılması da ayrı ayrı gerçekleştirilmiştir. Bu nedenle çalışmada, örnek olarak İBBS Düzey I TR1 bölgesi için tüm aşamaların yer aldığı tablolar gösterilmiştir. Diğer bölgeler ve Türkiye ortalaması için ise  $C_j$  ve  $W_j$  değerlerinin yer aldığı tablo verilmiştir.

Çalışmada CRITIC yönteminin kriterlerin ağırlıklandırılması için tercih edilme sebebi, objektif verileri ağırlıklandırmada kullanılan bir yöntem olmasıdır. Araştırmada yer alan tüm kriterler sayısal verilerden oluştuğu için CRITIC yönteminin uygun bir metot olduğu düşünülmüştür. Yöntemdeki tüm hesaplamalar için Microsoft Excel programı kullanılmıştır. Tablo 3.'te İBBS

Düzey I TR1 bölgesi için oluşturulan başlangıç karar matrisi gösterilmiştir (bkz. EK1).

Kriterler ağırlıklarının elde edilmesi için ikinci adımda başlangıç matrisinin normalizasyon işlemi gerçekleştirilmiştir. Tablo 4.'te normalize edilmiş karar matrisi gösterilmiştir (bkz. EK1).

Normalizasyon işleminin ardından, kriterler arasındaki ilişkilerin derecesini belirlemek için Eşitlik (4)'ten yararlanarak doğrusal ilişki katsayılarının oluşturduğu ilişki katsayı matrisi ( $P_{jk}$ ) değerleri hesaplanmıştır. Tablo 5'te ilişki katsayı matrisine yönelik değerler verilmiştir (bkz. EK1).

İlişki katsayı matrisi oluşturulduktan sonra, kriterlerde bulunan karşıtlık yoğunluğu ve çelişkileri birleştiren toplam bilgi ( $C_j$ ) değerleri hesaplanmıştır.  $C_j$  değerlerinin hesaplanması için öncelikle  $(1-P_{jk})$  değerleri belirlenmiştir. Tablo 6'da  $(1-P_{jk})$  değerleri gösterilmiştir (bkz. EK1).

$(1-P_{jk})$  değerlerinden sonra  $C_j$  değerlerinin hesaplanması için bir diğer değer olan  $\sigma_j$  değerleri Eşitlik (6)'dan yararlanılarak elde edilmiş ve Tablo 7'de gösterilmiştir (bkz. EK1).

$(1-P_{jk})$  ve  $\sigma_j$  değerlerinin elde edilmesinin ardından Eşitlik (5) ile  $C_j$  değerleri, Eşitlik (7) ile de kriterlerin ağırlığını gösteren  $W_j$  değerleri hesaplanmıştır. Tablo 8'de  $C_j$  ve  $W_j$  değerleri gösterilmiştir (bkz. EK1).

Tablo 8'e göre İBBS Düzey I TR1 Bölgesi için en yüksek ağırlığa 0,142 ile K4 kriteri (giderler), en az ağırlığa ise 0,054 ile K12 kriteri (satışların karlılığı) sahip olmuştur. Tablo 9'da diğer bölgeler ve Türkiye ortalaması (TR) için  $C_j$  ve  $W_j$  değerleri gösterilmiştir (bkz. EK1).

Tablo 9'a göre, TR2 bölgesi için en yüksek ağırlığa K1 (cari oran) kriteri, en düşük ağırlığa ise K12 (satışların karlılığı) kriteri sahip olmuştur. TR3 bölgesi için en önemli kriter, K4 (giderler) olmuştur. TR4 bölgesinde K7 (likidite oranı), TR5 bölgesinde K4 (giderler), TR6'da K5 (işletme sermayesi gün sayısı), TR7'de K4 (giderler), TR8'de K1 (cari oran), TR9'da K4 (giderler), TRA'da K10 (net satış hasılatı), TRB'de K2 (FAVÖK) ve TRC'de ise K1 (cari oran) kriterleri birinci sırada yer almıştır. TR (Türkiye ortalaması) için ise K5 (işletme sermayesi gün sayısı) kriteri en yüksek öneme sahip olmuştur.

### 6.3. MABAC Yöntemi ile Yıllara Göre Performans Sıralaması (Performance Ranking by Years with MABAC Method)

CRITIC yöntemi ile kriter ağırlıklarının hesaplanmasının ardından çalışmanın ikinci aşamasında, MABAC yönteminden yararlanılarak yazılım sektörünün 2018-2022 dönemine yönelik performans değerlendirmesi yapılmıştır. Bu aşamada da İBBS Düzey I TR1 bölgesi için tüm adımların gösterildiği tablolar örnek olarak verilmiştir. Diğer bölgeler ve Türkiye ortalaması için ise yıllara göre performans sıralamalarının yer aldığı tablolar

gösterilmiştir. MABAC yönteminde de CRITIC yönteminde olduğu gibi tüm hesaplamalar için Microsoft Excel programı kullanılmıştır.

İBBS Düzey I TR1 bölgesi için oluşturulan başlangıç karar matrisi, kriter ağırlıklarının hesaplandığı aşamada Tablo 3.'te verilmiştir. MABAC yönteminde karar matrisinin oluşturulmasının ardından, normalizasyon işlemi gerçekleştirilmektedir. Normalizasyon işlemi kolaylaştırmak adına öncelikle tüm kriterler için maksimum ve minimum değerler elde edilmiş ve Tablo 10'da bu değerler gösterilmiştir (bkz. EK1).

Kriterler için maksimum ve minimum değerlerin yer aldığı Tablo 10 değerleri elde edildikten sonra, Eşitlik (10) ve Eşitlik (11) kullanılarak normalizasyon işlemi gerçekleştirilmiştir. Tablo 11'de İBBS Düzey I TR1 bölgesi için normalize karar matrisi verilmiştir (bkz. EK1).

Normalizasyon işleminin ardından, Eşitlik (12)'den yararlanılarak Tablo 12'de gösterilen ağırlıklandırılmış normalize karar matrisi oluşturulmuştur (bkz. EK1).

Bir sonraki adımda, Eşitlik (14)'ten faydalanılarak tüm karar kriterleri için sınır yakınlık alanı belirlenmiştir ve Eşitlik (15)'te gösterilen sınır yakınlık alanı matrisi elde edilmiştir. Tablo 13'te bu matris gösterilmiştir (bkz. EK1).

Sonraki adımda, alternatiflerin sınır yakınlık alanına uzaklığını gösteren  $q_{ij}$ , Eşitlik (17)'den yararlanılarak hesaplanmıştır. Elde edilen  $Q$  matrisi Eşitlik (18)'deki gibi gösterilmiştir. Tablo 14'te sınır yakınlık alanına olan uzaklıkların yer aldığı matris verilmiştir (bkz. EK1).

MABAC yönteminin son adımında, Eşitlik (20)'den yararlanılarak karar seçeneklerinin kriter fonksiyonları  $S_i$  hesaplanmıştır. Bu çalışmada en yüksek  $S_i$  değerine sahip olan yıl, en iyi performansın gerçekleştiği yılı göstermektedir. Tablo 15'te İBBS Düzey I TR1 bölgesinin yıllara göre yazılım sektörü performans sıralaması verilmiştir. Tabloya göre, İBBS Düzey I TR1 bölgesi için en iyi performansa sahip olan yıl, 2020 yılıdır. 2020 yılını sırasıyla 2022, 2021, 2018 ve 2019 yılları izlemiştir.

Tablo 15. İBBS Düzey I TR1 Bölgesi İçin Yıllara Göre Yazılım Sektörü Performans Sıralaması  
(Software Industry Performance Ranking by Years for NUTS Level I TR1 Region)

Yıl	Si	Sıralama
2018	-0,153	4
2019	-0,15993	5
2020	0,248943	1
2021	0,088099	3
2022	0,1999	2

Tablo 16'da İBBS Düzey I TR2 bölgesinin yıllara göre yazılım sektörü performans sıralaması verilmiştir. Tabloya göre, İBBS Düzey I TR2 bölgesinin performansının en iyi olduğu yıl, 2021 yılıdır. 2021 yılını sırasıyla 2020, 2022, 2019 ve 2018 yılları takip etmiştir.

Tablo 16. İBBS Düzey I TR2 Bölgesi İçin Yıllara Göre Yazılım Sektörü Performans Sıralaması  
(Software Industry Performance Ranking by Years for NUTS Level I TR2 Region)

Yıl	Si	Sıralama
2018	-0,14349	5
2019	0,024282	4
2020	0,103076	2
2021	0,204373	1
2022	0,038483	3

Tablo 17'de İBBS Düzey I TR3 bölgesinin yıllara göre yazılım sektörü performans sıralaması verilmiştir. Tabloya göre, 2020 yılı İBBS Düzey I TR3 bölgesi için yazılım sektörü performansının en yüksek olduğu yıldır. 2020 yılını sırasıyla 2021, 2022, 2018 ve 2019 yılları izlemiştir.

Tablo 17. İBBS Düzey I TR3 Bölgesi İçin Yıllara Göre Yazılım Sektörü Performans Sıralaması  
(Software Industry Performance Ranking by Years for NUTS Level I TR3 Region)

Yıllar	Si	Sıralama
2018	-0,09773	4
2019	-0,16998	5
2020	0,215532	1
2021	0,158348	2
2022	0,136076	3

Tablo 18'de İBBS Düzey I TR4 bölgesinin yıllara göre yazılım sektörü performans sıralaması verilmiştir. Buna göre, İBBS Düzey I TR4 bölgesi için yazılım sektörü performansının en yüksek olduğu yıl, 2020 yılıdır. Bu bölge için 2020 yılını sırasıyla, 2022, 2018, 2021 ve 2019 yılları takip etmiştir.

Tablo 18. İBBS Düzey I TR4 Bölgesi İçin Yıllara Göre Yazılım Sektörü Performans Sıralaması  
(Software Industry Performance Ranking by Years for NUTS Level I TR4 Region)

Yıl	Si	Sıralama
2018	-0,02199	3
2019	-0,05169	5
2020	0,237387	1
2021	-0,05036	4
2022	0,120612	2

Tablo 19'da İBBS Düzey I TR5 bölgesinin yıllara göre yazılım sektörü performans sıralaması gösterilmiştir. Tabloya göre, İBBS Düzey I TR5 bölgesi için yazılım sektörü performansının en iyi olduğu yıl, 2020 yılı olarak belirlenmiştir. 2020 yılını, sırasıyla 2022, 2021, 2018 ve 2019 yılları izlemiştir.

Tablo 19. İBBS Düzey I TR5 Bölgesi İçin Yıllara Göre Yazılım Sektörü Performans Sıralaması  
(Software Industry Performance Ranking by Years for NUTS Level I TR5 Region)

Yıl	Si	Sıralama
2018	-0,04518	4
2019	-0,13621	5
2020	0,234538	1
2021	0,054095	3
2022	0,128204	2

Tablo 20'de İBBS Düzey I TR6 bölgesinin yıllara göre yazılım sektörü performans sıralaması gösterilmiştir. Buna göre 2022 yılı, İBBS Düzey I TR6 bölgesi için en iyi performansın gerçekleştiği yıldır. Performans sıralamasında 2022 yılını sırasıyla, 2020, 2019, 2021 ve 2018 yılları takip etmiştir.

Tablo 20. İBBS Düzey I TR6 Bölgesi İçin Yıllara Göre Yazılım Sektörü Performans Sıralaması  
(Software Industry Performance Ranking by Years for NUTS Level I TR6 Region)

Yıl	Si	Sıralama
2018	-0,08635	5
2019	0,059663	3
2020	0,149045	2
2021	-0,07434	4
2022	0,169622	1

Tablo 21'de İBBS Düzey I TR7 bölgesinin yıllara göre yazılım sektörü performans sıralaması gösterilmiştir. Tabloya göre, İBBS Düzey I TR7 bölgesi için en iyi performans 2020 yılında gerçekleşmiştir. 2020 yılını sırasıyla, 2022, 2021, 2018 ve 2019 yılları izlemiştir.

Tablo 21. İBBS Düzey I TR7 Bölgesi İçin Yıllara Göre Yazılım Sektörü Performans Sıralaması  
(Software Industry Performance Ranking by Years for NUTS Level I TR7 Region)

Yıl	Si	Sıralama
2018	-0,1193	4
2019	-0,12564	5
2020	0,246527	1
2021	0,026121	3
2022	0,199618	2

Tablo 22'de İBBS Düzey I TR8 bölgesinin yıllara göre yazılım sektörü performans sıralaması gösterilmiştir. Tabloya göre 2021 yılı, İBBS Düzey I TR8 bölgesi için en yüksek performansın olduğu yıldır. 2021 yılını sırasıyla, 2022, 2020, 2018 ve 2019 yılları takip etmiştir.

Tablo 22. İBBS Düzey I TR8 Bölgesi İçin Yıllara Göre Yazılım Sektörü Performans Sıralaması  
(Software Industry Performance Ranking by Years for NUTS Level I TR8 Region)

Yıl	Si	Sıralama
2018	-0,05618	4
2019	-0,25588	5
2020	0,099691	3
2021	0,280645	1
2022	0,172453	2

Tablo 23'te İBBS Düzey I TR9 bölgesinin yıllara göre yazılım sektörü performans sıralaması gösterilmiştir. Tabloya göre, İBBS Düzey I TR9 bölgesi için en yüksek performansa sahip olan yıl, 2022 yılıdır. 2022 yılını sırasıyla, 2018, 2021, 2020 ve 2019 yılları izlemiştir.

Tablo 23. İBBS Düzey I TR9 Bölgesi İçin Yıllara Göre Yazılım Sektörü Performans Sıralaması  
(Software Industry Performance Ranking by Years for NUTS Level I TR9 Region)

Yıl	Si	Sıralama
2018	0,072318	2
2019	-0,10048	5
2020	0,012242	4
2021	0,017568	3
2022	0,226409	1

Tablo 24'te İBBS Düzey I TRA bölgesinin yıllara göre yazılım sektörü performans sıralaması gösterilmiştir. Tabloya göre 2020 yılı, İBBS Düzey I TRA bölgesi için en yüksek performansın gerçekleştiği yıldır. 2020 yılını sırasıyla, 2019, 2018, 2022 ve 2021 yılları takip etmiştir.

Tablo 24. İBBS Düzey I TRA Bölgesi İçin Yıllara Göre Yazılım Sektörü Performans Sıralaması  
(Software Industry Performance Ranking by Years for NUTS Level I TRA Region)

Yıl	Si	Sıralama
2018	0,022628	3
2019	0,023363	2
2020	0,373968	1
2021	-0,17988	5
2022	-0,02043	4

Tablo 25'te İBBS Düzey I TRB bölgesinin yıllara göre yazılım sektörü performans sıralaması gösterilmiştir.

Tabloya göre, İBBS Düzey I TRB bölgesi için en yüksek performans 2020 yılında gerçekleşmiştir. 2020 yılını sırasıyla, 2022, 2021, 2018 ve 2019 yılları izlemiştir.

Tablo 25. İBBS Düzey I TRB Bölgesi İçin Yıllara Göre Yazılım Sektörü Performans Sıralaması  
(Software Industry Performance Ranking by Years for NUTS Level I TRB Region)

Yıl	Si	Sıralama
2018	0,010144	4
2019	-0,08821	5
2020	0,165931	1
2021	0,068716	3
2022	0,086819	2

Tablo 26'da İBBS Düzey I TRC bölgesinin yıllara göre yazılım sektörü performans sıralaması gösterilmiştir. Tabloya göre 2020 yılı, İBBS Düzey I TRC bölgesi için en yüksek performansın gerçekleştiği yıldır. 2020 yılını sırasıyla, 2018, 2022, 2021 ve 2019 yılları takip etmiştir.

Tablo 26. İBBS Düzey I TRC Bölgesi İçin Yıllara Göre Yazılım Sektörü Performans Sıralaması  
(Software Industry Performance Ranking by Years for NUTS Level I TRC Region)

Yıl	Si	Sıralama
2018	0,075324	2
2019	-0,07115	5
2020	0,152181	1
2021	0,026834	4
2022	0,056152	3

Tablo 27'de Türkiye'nin (TR) yıllara göre yazılım sektörü performans sıralaması verilmiştir. Tabloya göre Türkiye ortalamasında en yüksek performansa sahip olan yıl, 2020 yılıdır. 2020 yılını sırasıyla, 2022, 2021, 2019 ve 2018 yılları izlemiştir.

Tablo 27. Türkiye (TR) Yıllara Göre Yazılım Sektörü Performans Sıralaması  
(Software Industry Performance Ranking by Years for Turkey (TR))

Yıl	Si	Sıralama
2018	-0,14454	5
2019	-0,06315	4
2020	0,218486	1
2021	0,095139	3
2022	0,124286	2

## 7. SONUÇ (RESULT)

Yenilikçi ürünler, hizmetler ve iş modelleri geliştirmek, pazara hızlı bir şekilde giriş yapmak ve rakiplerden farklılaşmak için, işletmelerin yazılıma olan talebi gün geçtikçe daha da artmaktadır. Yapay zeka, makine

öğrenimi ve büyük veri analitiği gibi gelişmiş teknolojilerin entegrasyonu sayesinde, stratejik kararların alınması kolaylaşmaktadır. İşletme maliyetlerinin azaltılması ve verimliliğin artmasında da yazılım sektörü, büyük katkılar sunmaktadır. Esnek çalışma modellerini destekleyen yazılımlar, işletmelerin küresel pazarlara erişimini kolaylaştırmaktadır. Bununla birlikte güvenlik ve risk yönetimi için geliştirilen yazılımlar, olası sorunların ortaya çıkmasına engel olmaktadır. İşletmelere sağladığı bu avantajlardan dolayı son yıllarda yazılım sektörünün tüm endüstriler için vazgeçilmez bir konuma geldiği görülürken, gelecekte de bu önemin artarak devam edeceği tahmin edilmektedir.

Bu çalışmada istatistiki bölge birimleri sınıflandırması temelinde, Türkiye'de 2018-2022 dönemine yönelik yazılım sektörü performansı analiz edilmiştir. Performans değerlendirmesi için 13 finansal ölçüt kullanılmıştır. Araştırmada kullanılan veri seti, T.C. Küçük ve Orta Ölçekli İşletmeleri Geliştirme ve Destekleme İdaresi Başkanlığı'ndan (KOSGEB) temin edilmiştir [35]. Performans kriterlerinin ağırlıklandırılması için objektif bir metot olan CRITIC yöntemi, yıllara göre performans sıralaması için ise MABAC yöntemi kullanılmıştır. CRITIC yönteminden elde edilen bulgulara göre, İBBS Düzey I TR1, TR3, TR5, TR7, TR9 bölgeleri için en yüksek ağırlığa K4 (giderler) kriteri sahip olmuştur. TR2, TR8, TRC bölgeleri için K1 (cari oran) kriteri birinci sırada yer almıştır. TR4 bölgesinde K7 (likidite oranı), TR6'da K5 (işletme sermayesi gün sayısı), TRA'da K10 (net satış hasılatı), TRB'de K2 (FAVÖK) kriterleri en önemli kriter olarak belirlenmiştir. TR (Türkiye ortalaması) için ise K5 (işletme sermayesi gün sayısı), önem ağırlığı en yüksek kriter olmuştur.

Her bir bölge için kriter ağırlıklarının hesaplanmasının ardından, MABAC yönteminden yararlanılarak İBBS Düzey I bölgeleri ile Türkiye ortalaması için yıllara göre performans sıralaması gerçekleştirilmiştir. MABAC yönteminden elde edilen bulgulara göre, TR1, TR3, TR4, TR5, TR7, TRA, TRB, TRC bölgeleri için en yüksek performansın 2020 yılında gerçekleştiği görülmüştür. TR2 ve TR8 bölgeleri için 2021, TR6 ve TR9 bölgeleri için ise 2022 yılı en iyi performansa sahip yıllar olarak belirlenmiştir. Bununla birlikte 2019 yılı TR1, TR3, TR4, TR5, TR7, TR8, TR9, TRB ve TRC bölgeleri için en kötü performansın gerçekleştiği yıl olarak tespit edilmiştir. TR2 ve TR6 bölgeleri için 2018, TRA bölgesi için ise 2021 yılı diğer yıllara göre düşük performans göstermiştir. İBBS Düzey I bölgelerinin yazılım sektörü performans analizinin ardından Türkiye ortalaması için gerekli hesaplamalar yapılmıştır. Türkiye ortalaması veri setine göre en iyi performans 2020 yılında, en zayıf performans ise 2018 yılında gerçekleşmiştir.

Araştırmadan elde bulgular, 2018-2022 yılları arasında Türkiye'nin yazılım sektöründe bölgesel farklılıklar gösterdiğini ortaya koymaktadır. Özellikle COVID 19 döneminin ekonomik ve sosyal olarak büyük izler bıraktığı 2020 yılında, genel olarak yazılım sektörü performansının yükseldiği görülmektedir. Pandemi sürecinde dijital çözümlere duyulan ihtiyaç, yazılım

firmalarının çabalarını artırırken hem dünyada hem de ülkemizde yazılım sektörünü tüm dönemlerden farklı bir aşamaya taşımıştır. Bu dönemde birçok sektörün dijitalleşmeye ve online hizmetlere yönelmesi, yazılım talebini artırarak performansa olumlu katkı sağlamış olabilir.

Yazılım sektörünün Türkiye ve iktisadi bölge sınıflandırması performans değerlendirmesi, çeşitli açılardan önemli bilgiler sunabilir. Türkiye'nin genel ekonomik performansını ve bölgesel kalkınma eğilimlerini anlamak ve yıllara göre değerlendirmeler yapmak için çalışmada elde edilen bulgular aydınlatıcı olabilir. Yazılım sektörünün bölgesel performansı, bölgesel kalkınma politikalarının oluşturulması ve hayata geçirilmesinde önemli bir veri kaynağı sunabilir. Bu veriler bölgeler arası gelişmişlik farklarını tespit ederek, kalkınma politikalarının daha etkili ve hedefe yönelik bir şekilde planlanmasına yardımcı olabilir. Yerel ve ulusal düzeyde yatırım yapmak isteyen şirketler veya yatırımcılar için, yazılım sektörünün yıllara göre performansı ve büyüme potansiyeli yatırım kararlarını etkileyebilir. Bununla birlikte yazılım sektörü, genç ve nitelikli işgücü açısından önemli bir istihdam desteği sunduğu için performans değerlendirmeleri, bölgesel istihdam eğilimlerini ve gereksinimlerini anlamak için de faydalı olabilir. Bu da işgücü politikalarının oluşturulması ve yönetilmesinde yardımcı bir kaynak sağlayabilir.

Yazılım sektörünün Türkiye'nin genel rekabet gücüne katkısını değerlendirmek için gelecekteki araştırmacılar, farklı çok kriterli karar verme yöntemlerinden yararlanarak performans değerlendirmeleri yapabilir. Ayrıca İBBS Düzey I bölgelerinin kendi içerisindeki performans sıralaması için de çok kriterli karar verme yöntemleri kullanılabilir. Böylelikle bölgesel karşılaştırmalar yapmak kolaylaşabilir. Bu tarz araştırmalar, Türkiye'nin bölgesel olarak yazılım sektöründe ne kadar rekabetçi olduğunu anlamak ve rekabet avantajlarını vurgulamak için önemlidir. Özellikle yıllara göre değişim ve gelişimin fark edilmesi için gelecek dönemlerde yeni araştırmalar yapılması, hem sektör hem de literatür açısından bilgilendirici olabilir.

#### KAYNAKLAR (REFERENCES)

- [1] M. Damar, "Yazılım Sektörünün İki Lider Ülkesi Hindistan ve İrlanda, Gelişmekte Olan Ülkeler İçin Öneriler", *Ege Eğitim Teknolojileri Dergisi*, 6(1), 29-52, 2022.
- [2] K. M. Şenel, "Türkiye'de Yerli ve Milli Yazılımlar Üzerine Derinlemesine Görüşme", *International Social Sciences Studies Journal*, 8(105), 4073-4094, 2022.
- [3] İ. Aydın, "Bilişim Sektörü ve Türkiye'nin Sektördeki Potansiyeli", *International Journal of New Trends in Arts, Sports & Science Education*, 1(1), 180-200, 2012.
- [4] C. Gök-Kısa, S. Perçin, "Bütünlük Entropi Ağırlık-Vikor Yöntemi ile Bilişim Teknolojisi Sektöründe Performans Ölçümü", *Ekonomik ve Sosyal Araştırmalar Dergisi*, 14(1), 1-13, 2018.
- [5] S. Akginel, **Çok kriterli karar verme teknikleriyle bilişim sektöründe performans değerlendirmesi**, Yüksek Lisans Tezi, Dokuz Eylül Üniversitesi, Sosyal Bilimler Enstitüsü, 2019.
- [6] Ş. Sakarya, S. İlkdoğan, "BIST Bilişim Sektöründe Faaliyet Gösteren İşletmelerin Finansal Performanslarının Nakit Akış Oranları Kapsamında CRITIC Temelli TOPSIS Yöntemi ile Değerlendirilmesi", *Pamukkale Üniversitesi İşletme Araştırmaları Dergisi*, 9(2), 421-445, 2022.
- [7] M. Campbell-Kelly, "Development and Structure of the International Software Industry, 1950-1990", *Business and Economic History*, 73-110, 1995.
- [8] P. Tyrvainen, J. Warsta, V. Seppanen, "Evolution of Secondary Software Businesses: Understanding Industry Dynamics", **In Open IT-Based Innovation: Moving Towards Cooperative IT Transfer and Knowledge Diffusion: IFIP TC8 WG 8.6 International Working Conference**, Madrid, Spain (pp. 381-401), October 22-24, 2008.
- [9] A. Arora, A. Gambardella, "The Globalization of the Software Industry: Perspectives and Opportunities for Developed and Developing Countries", *Innovation Policy and the Economy*, 5, 1-32, 2005.
- [10] R. J. Shapiro, "The U.S. Software Industry as an Engine for Economic Growth and Employment", *Georgetown McDonough School of Business Research Paper*, (2541673), 2014.
- [11] M. Andreessen, "Why Software is Eating the World". *Wall Street Journal*, 20(2011), C2, 2011.
- [12] C. J. Chen, Q. J. Yeh, "A Comparative Performance Evaluation of Taiwan's High-Tech Industries", *International Journal of Business Performance Management*, 7(1), 16-33, 2005.
- [13] Y. Akyüz, Z. Kaya, "Türkiye'de Hayat Dışı Ve Hayat/Emeklilik Sigorta Sektörünün Finansal Performans Analiz Ve Değerlendirilmesi", *Sosyal Ekonomik Araştırmalar Dergisi*, 13(26), 355-371, 2013.
- [14] S. Liang, "A Comparative Performance Assessment of Wealth Management Banks Using Multicriteria Analysis", *Journal of Applied Finance and Banking*, 3(1), 13.
- [15] T. D. Chaudhuri, I. Ghosh, "A Multi-Criteria Decision Making Model-Based Approach for Evaluation of the Performance of Commercial Banks in India", *IUP Journal of Bank Management*, 13(3), 23-33, 2014.
- [16] S. A. Tezergil, "Vikor Yöntemi İle Türk Bankacılık Sektörünün Performans Analizi", 38(1), 357-373, 2016.
- [17] Ö. Akçakanat, H. Eren, E. Aksoy, V. Ömürbek, "Bankacılık Sektöründe ENTROPI ve WASPAS Yöntemleri ile Performans Değerlendirmesi", *Süleyman Demirel Üniversitesi İktisadi ve İdari Bilimler Fakültesi Dergisi*, 22(2), 285-300, 2017.
- [18] A. Başdeğirmen, M. Z. Tunca, "Lojistik Sektöründe Faaliyet Gösteren İşletmelerin Finansal Performanslarının Gri İlişkisel Analiz ile Değerlendirilmesi", *Süleyman Demirel Üniversitesi İktisadi ve İdari Bilimler Fakültesi Dergisi*, 22(2), 327-340, 2017.
- [19] O. Karkacier, A. E., Yazgan, "Turizm Sektöründe Gri İlişkisel Analiz GİA Yöntemiyle Finansal Performans Değerlemesi", *Selçuk Üniversitesi Sosyal Bilimler Enstitüsü Dergisi*, (37), 154-162, 2017.

- [20] A. Korkusuz, U. İnan, Y. Özdemir, H. Başlıgil, "Occupational Health and Safety Performance Measurement in Healthcare Sector Using Integrated Multi Criteria Decision Making Methods", *Journal of the Faculty of Engineering and Architecture of Gazi University*, 35(1), 81-96, 2020.
- [21] Ç. Orçun, "Enerji Sektöründe WASPAS Yöntemiyle Performans Analizi", *Bolu Abant İzzet Baysal Üniversitesi Sosyal Bilimler Enstitüsü Dergisi*, 19(2), 439-453, 2019.
- [22] B. Elmas, T. Özkan, "Ulaştırma ve Depolama Sektörü İşletmelerinin Finansal Performanslarının SWARA-OCRA Modeli ile Değerlendirilmesi", *İşletme Akademisi Dergisi*, 2(3), 240-253, 2021.
- [23] A. E. Torkayesh, D. Pamucar, F. Ecer, P. Chatterjee, "An Integrated BWM-LBWA-CoCoSo Framework for Evaluation of Healthcare Sectors in Eastern Europe", *Socio-Economic Planning Sciences*, 78, 2021.
- [24] R. Lukic, "Evaluation of the Performance Trading Companies in Serbia based on the FF-WASPAS and WASPAS Methods", *The European Journal of Applied Economics*, 20(2), 41-63, 2023.
- [25] Kalkınma Ajansları, <https://www.ka.gov.tr/>, 30.03.2024.
- [26] M. Zizovic, B. Miljkovic, D. Marinkovic, "Objective Methods for Determining Criteria Weight Coefficients: A Modification of the CRITIC Method", *Decision Making: Applications in Management and Engineering*, 3(2), 149-161, 2020.
- [27] H. Kazan, O. Özdemir, "Financial Performance Assessment of Large Scale Conglomerates via TOPSIS and CRITIC Methods", *International Journal of Management and Sustainability*, 3(4), 203-224, 2014.
- [28] D. Diakoulaki, G. Mavrotas, L. Papayannakis, "Determining Objective Weights in Multiple Criteria Problems: The CRITIC Method", *Computers & Operations Research*, 22(7), 763-770, 1995.
- [29] M.G. Naik, R. Kishore, S. A. M. Dehmourdi, "Modeling A Multi-Criteria Decision Support System for Prequalification Assessment of Construction Contractors Using CRITIC and EDAS Models", *Operational Research in Engineering Sciences: Theory and Applications*, 4(2), 79-101, 2021.
- [30] A. Puska, M. Nedeljkovic, R. Prodanovic, R. Vladislavjevic, R. Suzic, "Market Assessment of Pear Varieties in Serbia Using Fuzzy CRADIS and CRITIC methods", *Agriculture*, 12(2), 139, 2022.
- [31] A. Jahan, F. Mustapha, S. M. Sapuan, M. Y. Ismail, M. Bahraminasab, "A Framework For Weighting of Criteria in Ranking Stage of Material Selection Process", *The International Journal of Advanced Manufacturing Technology*, 58, 411-420, 2012.
- [32] S. Zhong, Y. Chen, Y. Miao, "Using Improved CRITIC Method to Evaluate Thermal Coal Suppliers", *Scientific Reports*, 13(1), 195, 2023.
- [33] D. Pamucar, G. Cirovic, "The Selection of Transport and Handling Resources in Logistics Centers Using Multi-Attributive Border Approximation Area Comparison (MABAC)", *Expert Systems with Applications*, 42(6), 3016-3028, 2015.
- [34] Z. Nunic, "Evaluation and Selection of the PVC Carpentry Manufacturer Using the FUCOM-MABAC Model", *Operational Research in Engineering Sciences: Theory and Applications*, 1(1), 13-28, 2018.
- [35] T.C. Küçük ve Orta Ölçekli İşletmeleri Geliştirme ve Destekleme İdaresi Başkanlığı (KOSGEB).

## EK1 (ADDITION1)

Tablo 3. İBBS Düzey I TR1 Bölgesi İçin Başlangıç Karar Matrisi  
(Initial Decision Matrix for NUTS Level I TR1 Region)

	K1	K2	K3	K4	K5	K6	K7	K8	K9	K10	K11	K12	K13
Yön	Maks	Maks	Maks	Min	Min	Maks	Maks	Maks	Maks	Maks	Maks	Maks	Min
2018	1,59	19,07	65232,76	47801,32	62,81	3801,7	1,22	0,56	42527,41	60824,65	0,3	0,09	352,2
2019	1,56	6,86	64165,39	43837,08	67,21	8159,68	1,23	0,59	46934,24	61620,12	0,32	0,06	208,49
2020	1,86	21,5	60124,25	36931,53	58,49	12563,25	1,4	0,73	118411,22	56723,87	0,27	0,18	118,37
2021	1,58	22,02	107822,29	59248,92	47,67	23270,39	1,2	0,67	65006,53	103205,42	0,35	0,11	222,55
2022	1,54	23,51	182127,41	94621,5	54,29	33934,93	1,15	0,65	114620,24	174923,66	0,43	0,15	79,28

Tablo 4. İBBS Düzey I TR1 Bölgesi İçin Normalize Karar Matrisi  
(Normalized Decision Matrix for NUTS Level I TR1 Region)

	K1	K2	K3	K4	K5	K6	K7	K8	K9	K10	K11	K12	K13
Yön	Maks	Maks	Maks	Min	Min	Maks	Maks	Maks	Maks	Maks	Maks	Maks	Min
2018	0,156	0,733	0,042	0,812	0,225	0,000	0,280	0,000	0,000	0,035	0,188	0,250	0,000
2019	0,063	0,000	0,033	0,880	0,000	0,145	0,320	0,176	0,058	0,041	0,313	0,000	0,527
2020	1,000	0,879	0,000	1,000	0,446	0,291	1,000	1,000	1,000	0,000	0,000	1,000	0,857
2021	0,125	0,911	0,391	0,613	1,000	0,646	0,200	0,647	0,296	0,393	0,500	0,417	0,475
2022	0,000	1,000	1,000	0,000	0,661	1,000	0,000	0,529	0,950	1,000	1,000	0,750	1,000

Tablo 5. İBBS Düzey I TR1 Bölgesi İçin İlişki Katsayı Matrisi ( $P_{jk}$ ) Değerleri  
(Relationship Coefficient Matrix ( $P_{jk}$ ) Values for NUTS Level I TR1 Region)

	K1	K2	K3	K4	K5	K6	K7	K8	K9	K10	K11	K12	K13
K1	1,000	0,240	-0,478	0,567	-0,032	-0,267	0,967	0,705	0,528	-0,480	-0,676	0,673	0,282
K2	0,240	1,000	0,512	-0,462	0,774	0,574	-0,003	0,531	0,607	0,504	0,296	0,760	0,273
K3	-0,478	0,512	1,000	-0,989	0,573	0,937	-0,663	0,146	0,477	1,000	0,964	0,318	0,552
K4	0,567	-0,462	-0,989	1,000	-0,490	-0,877	0,738	-0,005	-0,381	-0,988	-0,975	-0,222	-0,453
K5	-0,032	0,774	0,573	-0,490	1,000	0,743	-0,223	0,566	0,396	0,572	0,458	0,468	0,308
K6	-0,267	0,574	0,937	-0,877	0,743	1,000	-0,462	0,445	0,615	0,939	0,866	0,472	0,703
K7	0,967	-0,003	-0,663	0,738	-0,223	-0,462	1,000	0,567	0,340	-0,663	-0,809	0,471	0,155
K8	0,705	0,531	0,146	-0,005	0,566	0,445	0,567	1,000	0,815	0,148	-0,055	0,845	0,720
K9	0,528	0,607	0,477	-0,381	0,396	0,615	0,340	0,815	1,000	0,476	0,264	0,942	0,882
K10	-0,480	0,504	1,000	-0,988	0,572	0,939	-0,663	0,148	0,476	1,000	0,966	0,314	0,557
K11	-0,676	0,296	0,964	-0,975	0,458	0,866	-0,809	-0,055	0,264	0,966	1,000	0,064	0,436
K12	0,673	0,760	0,318	-0,222	0,468	0,472	0,471	0,845	0,942	0,314	0,064	1,000	0,686
K13	0,282	0,273	0,552	-0,453	0,308	0,703	0,155	0,720	0,882	0,557	0,436	0,686	1,000



Tablo 6. İBBS Düzey I TR1 Bölgesi İçin  $(1-P_{jk})$  Değerleri $((1-P_{jk})$  Values for NUTS Level I TR1 Region)

	K1	K2	K3	K4	K5	K6	K7	K8	K9	K10	K11	K12	K13
K1	0,000	0,760	1,478	0,433	1,032	1,267	0,033	0,295	0,472	1,480	1,676	0,327	0,718
K2	0,760	0,000	0,488	1,462	0,226	0,426	1,003	0,469	0,393	0,496	0,704	0,240	0,727
K3	1,478	0,488	0,000	1,989	0,427	0,063	1,663	0,854	0,523	0,000	0,036	0,682	0,448
K4	0,433	1,462	1,989	0,000	1,490	1,877	0,262	1,005	1,381	1,988	1,975	1,222	1,453
K5	1,032	0,226	0,427	1,490	0,000	0,257	1,223	0,434	0,604	0,428	0,542	0,532	0,692
K6	1,267	0,426	0,063	1,877	0,257	0,000	1,462	0,555	0,385	0,061	0,134	0,528	0,297
K7	0,033	1,003	1,663	0,262	1,223	1,462	0,000	0,433	0,660	1,663	1,809	0,529	0,845
K8	0,295	0,469	0,854	1,005	0,434	0,555	0,433	0,000	0,185	0,852	1,055	0,155	0,280
K9	0,472	0,393	0,523	1,381	0,604	0,385	0,660	0,185	0,000	0,524	0,736	0,058	0,118
K10	1,480	0,496	0,000	1,988	0,428	0,061	1,663	0,852	0,524	0,000	0,034	0,686	0,443
K11	1,676	0,704	0,036	1,975	0,542	0,134	1,809	1,055	0,736	0,034	0,000	0,936	0,564
K12	0,327	0,240	0,682	1,222	0,532	0,528	0,529	0,155	0,058	0,686	0,936	0,000	0,314
K13	0,718	0,727	0,448	1,453	0,692	0,297	0,845	0,280	0,118	0,443	0,564	0,314	0,000

Tablo 7. İBBS Düzey I TR1 Bölgesi İçin  $\sigma_j$  Değerleri $(\sigma_j$  Values for NUTS Level I TR1 Region)

	K1	K2	K3	K4	K5	K6	K7	K8	K9	K10	K11	K12	K13
$\sigma_j$	0,413	0,405	0,426	0,395	0,387	0,405	0,378	0,395	0,483	0,426	0,382	0,397	0,388

Tablo 8. İBBS Düzey I TR1 Bölgesi İçin  $C_j$  ve  $W_j$  Değerleri $(C_j$  and  $W_j$  Values for NUTS Level I TR1 Region)

	K1	K2	K3	K4	K5	K6	K7	K8	K9	K10	K11	K12	K13
$C_j$	4,119	2,997	3,685	6,536	3,052	2,961	4,383	2,593	2,914	3,686	3,893	2,465	2,678
$W_j$	0,090	0,065	0,080	0,142	0,066	0,064	0,095	0,056	0,063	0,080	0,085	0,054	0,058

Tablo 9. İBBS Düzey I Diğer Bölgeler ve Türkiye Geneli İçin  $C_j$  ve  $W_j$  Değerleri $(C_j$  and  $W_j$  Values for NUTS Level I Other Regions and Turkey in General)

		K1	K2	K3	K4	K5	K6	K7	K8	K9	K10	K11	K12	K13
TR2	$C_j$	6,184	3,329	3,390	6,167	6,165	3,351	5,921	3,296	3,646	3,534	3,334	3,187	3,388
	$W_j$	0,113	0,061	0,062	0,112	0,112	0,061	0,108	0,060	0,066	0,064	0,061	0,058	0,062
TR3	$C_j$	3,647	3,422	4,647	5,978	5,063	4,454	3,883	3,335	3,239	4,694	5,430	4,063	3,748
	$W_j$	0,066	0,062	0,084	0,108	0,091	0,080	0,070	0,060	0,058	0,084	0,098	0,073	0,067
TR4	$C_j$	5,839	4,027	4,800	5,884	4,549	4,294	5,951	4,295	4,167	4,708	4,690	3,876	4,562
	$W_j$	0,095	0,065	0,078	0,095	0,074	0,070	0,097	0,070	0,068	0,076	0,076	0,063	0,074
TR5	$C_j$	4,497	4,335	4,389	6,159	5,707	5,015	4,102	3,996	3,552	4,406	4,994	3,449	3,483
	$W_j$	0,077	0,075	0,076	0,106	0,098	0,086	0,071	0,069	0,061	0,076	0,086	0,059	0,060
TR6	$C_j$	5,213	3,954	4,205	5,776	5,940	4,223	5,137	4,074	3,636	4,196	4,319	3,574	4,778
	$W_j$	0,088	0,067	0,071	0,098	0,101	0,072	0,087	0,069	0,062	0,071	0,073	0,061	0,081
TR7	$C_j$	4,100	3,452	4,046	6,249	5,323	3,403	4,260	2,918	3,201	4,046	4,152	3,332	2,860
	$W_j$	0,080	0,067	0,079	0,122	0,104	0,066	0,083	0,057	0,062	0,079	0,081	0,065	0,056
TR8	$C_j$	5,730	2,908	3,585	4,832	2,594	3,017	3,548	2,752	5,287	3,595	3,452	3,180	3,216
	$W_j$	0,120	0,061	0,075	0,101	0,054	0,063	0,074	0,058	0,111	0,075	0,072	0,067	0,067
TR9	$C_j$	4,620	3,110	3,153	7,053	6,120	3,646	3,540	4,204	4,397	3,367	3,008	4,647	4,941
	$W_j$	0,083	0,056	0,056	0,126	0,110	0,065	0,063	0,075	0,079	0,060	0,054	0,083	0,089
TRA	$C_j$	2,196	2,855	4,334	3,162	3,602	4,721	2,271	2,187	3,264	5,537	3,246	2,610	3,955
	$W_j$	0,050	0,065	0,099	0,072	0,082	0,107	0,052	0,050	0,074	0,126	0,074	0,059	0,090
TRB	$C_j$	5,691	6,055	4,789	5,882	4,627	5,731	4,666	4,437	4,687	4,921	4,341	4,697	5,790
	$W_j$	0,086	0,091	0,072	0,089	0,070	0,086	0,070	0,067	0,071	0,074	0,065	0,071	0,087
TRC	$C_j$	5,667	5,503	5,404	5,555	4,354	5,643	4,937	4,728	5,207	5,417	4,990	4,732	5,514
	$W_j$	0,084	0,081	0,080	0,082	0,064	0,083	0,073	0,070	0,077	0,080	0,074	0,070	0,082
TR	$C_j$	4,397	3,887	4,631	5,804	6,265	4,344	4,285	3,226	3,460	4,639	4,626	2,983	2,815
	$W_j$	0,079	0,070	0,084	0,105	0,113	0,078	0,077	0,058	0,062	0,084	0,084	0,054	0,051

Tablo 10. İBBS Düzey I TR1 Bölgesi İçin Kriterlerin Maksimum ve Minimum Değerleri  
(Maximum and Minimum Values of Criteria for NUTS Level I TR1 Region)

	K1	K2	K3	K4	K5	K6	K7	K8	K9	K10	K11	K12	K13
<b>Xi+</b>	1,86	23,51	182127,4	94621,5	67,21	33934,93	1,4	0,73	118411,2	174923,7	0,43	0,18	352,2
<b>Xi-</b>	1,54	6,86	60124,25	36931,53	47,67	3801,7	1,15	0,56	42527,41	56723,87	0,27	0,06	79,28

Tablo 11. İBBS Düzey I TR1 Bölgesi İçin Normalize Karar Matrisi  
(Normalized Decision Matrix for NUTS Level I TR1 Region)

	K1	K2	K3	K4	K5	K6	K7	K8	K9	K10	K11	K12	K13
<b>2018</b>	0,156	0,733	0,042	0,812	0,225	0,000	0,280	0,000	0,000	0,035	0,188	0,250	0,000
<b>2019</b>	0,063	0,000	0,033	0,880	0,000	0,145	0,320	0,176	0,058	0,041	0,313	0,000	0,527
<b>2020</b>	1,000	0,879	0,000	1,000	0,446	0,291	1,000	1,000	1,000	0,000	0,000	1,000	0,857
<b>2021</b>	0,125	0,911	0,391	0,613	1,000	0,646	0,200	0,647	0,296	0,393	0,500	0,417	0,475
<b>2022</b>	0,000	1,000	1,000	0,000	0,661	1,000	0,000	0,529	0,950	1,000	1,000	0,750	1,000

Tablo 12. İBBS Düzey I TR1 Bölgesi İçin Ağırlıklandırılmış Normalize Karar Matrisi  
(Weighted Normalized Decision Matrix for NUTS Level I TR1 Region)

	K1	K2	K3	K4	K5	K6	K7	K8	K9	K10	K11	K12	K13
<b>W<sub>j</sub></b>	<b>0,090</b>	<b>0,065</b>	<b>0,080</b>	<b>0,142</b>	<b>0,066</b>	<b>0,064</b>	<b>0,095</b>	<b>0,056</b>	<b>0,063</b>	<b>0,080</b>	<b>0,085</b>	<b>0,054</b>	<b>0,058</b>
<b>2018</b>	0,104	0,113	0,084	0,258	0,081	0,064	0,122	0,056	0,063	0,083	0,101	0,067	0,058
<b>2019</b>	0,095	0,065	0,083	0,267	0,066	0,074	0,126	0,066	0,067	0,084	0,111	0,054	0,089
<b>2020</b>	0,179	0,123	0,080	0,284	0,096	0,083	0,191	0,113	0,127	0,080	0,085	0,107	0,108
<b>2021</b>	0,101	0,125	0,112	0,229	0,133	0,106	0,114	0,093	0,082	0,112	0,127	0,076	0,086
<b>2022</b>	0,090	0,130	0,160	0,142	0,110	0,129	0,095	0,086	0,124	0,160	0,169	0,094	0,117

Tablo 13. İBBS Düzey I TR1 Bölgesi İçin Sınır Yakınlık Alanı Matrisi  
(Border Approximation Area Matrix for NUTS Level I TR1 Region)


	K1	K2	K3	K4	K5	K6	K7	K8	K9	K10	K11	K12	K13
<b>g<sub>i</sub></b>	0,110	0,108	0,100	0,230	0,095	0,088	0,126	0,081	0,089	0,100	0,115	0,077	0,089

Tablo 14. İBBS Düzey I TR1 Bölgesi İçin Sınır Yakınlık Alanına Olan Uzaklıklar  
(Distances to the Border Approximation Area for NUTS Level I TR1 Region)

	K1	K2	K3	K4	K5	K6	K7	K8	K9	K10	K11	K12	K13
<b>2018</b>	-0,006	0,005	-0,016	0,028	-0,013	-0,024	-0,004	-0,024	-0,025	-0,017	-0,015	-0,010	-0,031
<b>2019</b>	-0,015	-0,043	-0,017	0,038	-0,028	-0,015	0,000	-0,014	-0,022	-0,016	-0,004	-0,024	0,000
<b>2020</b>	0,069	0,015	-0,020	0,055	0,001	-0,005	0,065	0,032	0,038	-0,020	-0,031	0,030	0,019
<b>2021</b>	-0,009	0,017	0,012	0,000	0,038	0,018	-0,012	0,012	-0,006	0,012	0,012	-0,001	-0,003
<b>2022</b>	-0,020	0,022	0,061	-0,087	0,016	0,040	-0,031	0,006	0,035	0,060	0,054	0,017	0,027

# Gürültülü Verilere Dayanıklı kNN Algoritması Temelli Yeni Bir İkili Sınıflandırma Algoritması

*Araştırma Makalesi/Research Article*

 Müge ACAR

Endüstri Mühendisliği Bölümü, Eskişehir Teknik Üniversitesi, Eskişehir, Türkiye

[msoyuz@eskisehir.edu.tr](mailto:msoyuz@eskisehir.edu.tr)

(Geliş/Received:16.08.2024; Kabul/Accepted:15.10.2024)

DOI: 10.17671/gazibtd.1534334

**Özet**— Sınıflandırma, belirlenmiş bazı kriterlere göre kategoriler halinde sistematik olarak verilerin analizinde kullanılan etkili bir tekniktir. Sınıflandırıcının başarısı, sınıflandırıcının kendisine ve verilerin kalitesine bağlıdır. Bununla birlikte, gerçek hayat uygulamalarında, veri kümelerinin yanlış etiketlenmiş örnekler içermesi kaçınılmazdır. Gerçek hayat verileri gürültü olarak bilinen yanlış etiketlenmiş örnekler içerebilir. Bu da yanlış sınıflandırmalara neden olabilir. Bu çalışma, yeni bir kNN (k en yakın komşuluk algoritması) tabanlı sınıflandırma algoritması ile gürültü verilerinin sınıflandırılmasının nicel bir değerlendirmesini ve verileri verimli bir şekilde sınıflandırarak klasik kNN'nin performansını artırmayı amaçlamaktadır. Bu yeni tekniğin, gürültü verileriyle ikili sınıflandırma problemlerinde yüksek standart doğruluk seviyeleri sağlayabileceğini öneriyoruz. Bu çalışma, sınıflandırmadan önce gürültü noktaları tespit edilmesini dikkate alarak ikili sınıflandırma problemlerinde kNN tekniğinin performansını arttırmaktadır. Yeni kNN ve klasik kNN algoritmalarını farklı gürültü seviyelerinde (%10, %20, %30 ve %40) farklı veri setlerinde test doğruluğu açısından ölçerek karşılaştırdık ve özellikle yüksek gürültü seviyelerinde %84.33, %93.63, %81.81, %88.00 ye varan test doğruluk değerleri ile klasik kNN algoritmasına göre yüksek değerler elde edildi. Ayrıca geliştirdiğimiz algoritmayı popüler sınıflandırma algoritmalarıyla karşılaştırdığımızda bazı verilerde daha yüksek doğruluk değerleri elde edilerek %85.19, %97.07, %96.63' e varan doğruluk değerleri gözlemlendi.

**Anahtar Kelimeler**— ikili sınıflandırma, gürültü verisi, veri madenciliği

## A New Binary Classifier Robust on Noisy Domains Based on kNN Algorithm

**Abstract**— Classification is an effective technique commonly used in data analysis by systematically arranging groups or categories according to established criteria. The classifier's success relies on the classifier itself and the quality of the data. However, in real-world applications, it is inevitable for datasets to contain mislabeled instances, which may cause misclassification challenges that classifiers have to handle. This study aims for a quantitative assessment of the classification of noisy data through a new kNN-based classification algorithm and to increase the performance of classical kNN by efficiently classifying the data. We perform various numerical experiments on real-world data sets to prove our new algorithm's performance. We obtain high standards of accuracy levels on various noisy datasets. We propose that this new technique can provide high standard accuracy levels in binary classification problems. We compared the new kNN and the classical kNN algorithms in terms of test accuracies on different datasets under varying noise levels (10%, 20%, 30%, and 40%). Particularly at higher noise levels, we achieve significantly higher test accuracy values compared to the classical kNN algorithm, with results reaching 84.33%, 93.63%, 81.81%, and 88.00%. Additionally, when we compare our new algorithm with popular classification algorithms, we observed higher accuracy rates on some datasets, with values reaching up to 85.19%, 97.07%, and 96.63%.

**Keywords**— binary classification, noisy data, data mining

## 1. INTRODUCTION

In recent years, data science has high impact in many areas and continues to develop. Various decisions such as budget planning, sales strategies even social media analysis are made based on the data analysis. In all kinds of analysis, classification is commonly applied in making such crucial decisions. All classification techniques perform the analysis based on the training data. However, the training data may not always be perfect. In real-life, the data often contain noise. According to Zhu and Wu [5], noise refers to any irrelevant or meaningless data that interferes with the processing, transmission, or interpretation of information. They states that noise can occur in various forms such as random errors, distortions, or unwanted signals that affect the quality and accuracy of data or communication. Bishop explains that [30] in data science and machine learning, noise is often considered as random variations or outliers in data that do not represent the underlying patterns or trends. Goodfellow et. al. [29] mention that noise can degrade the performance of algorithms and models, leading to inaccurate results or predictions. So according to studies states about noise reveals that noisy data can be challenging for classifiers to carry out the qualified separations.

According to the literature, different types of noisy points are classified. According to Catal [26], noisy points can be seen in different data parts. First data labels may be incorrect. This type is defined as class noise. Second type is called attribute noise, that is about wrong attribute values. Third type is seen as the combination of these two types. Catal [26] says that class noise has performance decreasing effects on accuracy of data rather than attribute noise. This fact can be explained as two main reasons: One of them is that for each data sample, there are multiple attributes but there is a single label value. While it is possible to compensate an error on a certain attribute with another attribute, the error on the label will be less tolerable. Other reason may be explained that as each attribute can have a bold or weak effect on learning methods but labels always have strong effects on sample data. Therefore, we considered class noise in our study.

In the literature, many classification techniques have been presented to perform classification of noisy data in different ways. These may be generalized in two different approaches. The first approach involves pre-processing of the data. In this approach, noise points in the data is cleaned by filtering, and then the cleaned data is classified using a classification algorithm. Although this seems to be an effective way, it can be long and difficult to implement because it requires two different stages to implement. Other approach is directly classify the data containing noise. In this case, the algorithm should have a mechanism to classify the data by considering there will be noise points.

There are many popular classification algorithms (e.g., SVM, decision trees, neural networks, kNN, etc.) that have been used in the literature. Among these, k nearest neighbor (kNN) is an effective and powerful lazy learning algorithm that is also easy-to-implement. However, its performance heavily relies on the quality of training data. Since this algorithm is dependent on data quality, it makes strong predictions when quality data is available, but may not achieve the expected success when there exist data

losses or noisy data. Considering the fact that the real-life data cannot also be controlled and may contain noise or missing values, the kNN algorithm led to decision makers to make wrong decisions.

In this study, we present a new method based on kNN algorithm to be used in the classification of data containing class noise. Our method does not need to consider whether the data is noisy or if needed for preprocessing. We argue that the new algorithm, which is not data dependent and easy to implement, will yield more consistent predictions. We demonstrate this claim with numerical studies in which we test our algorithm with the data containing 10%, 20%, 30%, 40% noise level and compare the results obtained with those of the classical kNN algorithm. The results we obtain indicated that the estimation success of the new kNN algorithm remained more stable even if the noise level increased, whereas the estimation success of the classical kNN algorithm decreased with increasing the noise level. We also compared the test accuracy values of our new algorithm and those of popular classification algorithms in the literature. Moreover, it is also shown that the test success of the new method is independent of the parameter.

## 2. LITERATURE REVIEW

In the classification literature, there are several studies that consider classification with noise. ([1]; [2]; [3]; [4]; [19];[20]; [18];[27];[28]). It is seen that different evaluation methods have been used in these studies. Some studies examine noisy data classification with pre-processing. They prefer to process and clean the data beforehand and classify the cleaned data with a classification algorithm. Saez and Corhado [4] consider pre-processing the data first to repair the attribute noise. They proposed not to filter the data but repair it to increase the performance of the classification algorithm. Sluban and Lavrač [6] suggest filtering the noise. They explored the class noise detection by studying different diversity measures on a range of heterogeneous noise detection ensembles. Luengo et al [7] utilize bag noise filtering for negative instance noise cleaning. Garcia et al [8] focus on a new label noise injection method for the evaluation of noise filters. Wang et al [9] used a new method for cleaning noise by using the LNC-SDAE framework. Mansour et al [31] proposed an adaptive synthetic sampling-based noise detection technique for mobile edge computing. Dash et al [32] suggest an outliers detection and elimination framework in classification, they consider winsorizing method to deal with the outliers.

Unlike the studies focusing on pre-processing, some studies in the literature suggested that pre-processing followed by classification may become complex and be difficult to implement. Thus, they adopted the idea of using an extended classification algorithm for noise data and performing classification in one stage. There are several successful algorithms (e.g., SVM, decision trees, neural networks, kNN, etc.) implemented that provide successful test accuracy results assuming the data is clean. Although these methods are successful in data classification, since their estimation capabilities decrease when noise level increases in data, their new versions have been used in the literature and therefore these new versions are directly resistant to noise data.

Marsala and Petturiti [10] considered a decision tree based method to classify the noisy data. They focus on order discrimination criteria to be used in decision tree induction, that is, functions that can measure the discriminatory power of an attribute with respect to class, taking into account the monotony of the class with respect to the attribute. Zhu et al [11] proposed a new classification technique noise data by using neural networks. They demonstrate that their method has an effective classification ability while comparing the methods in literature. Chao et al [12] developed a new SVM (support vector machine) technique (RTS-SVM) for classifying data with high percentage of noise. They conclude that the main advantage of the proposed SVM is to improve the classification performance by avoid the effect of the entangled noise data when compared with the classic SVM. When we evaluate these methods, either it is difficult and complex to implement but the quality of classification has increased, or the method has compromised success despite being easy to implement. In our opinion, this between implementation difficulty and test success stands out as a trade-off that is worth to be evaluated.

kNN is an effective classification algorithm which is very popular, easy to implement and have successful test results. It has been widely used for clustering and classification of data for various of different fields. ([15]). However, in recent years due to the complexity of the data and the demand to produce accurate and high-quality information with new technological developments, the performance of the kNN algorithm is not sufficient. Thus, there are other studies in the literature that have focused on increasing the performance of the kNN algorithm and improve the disadvantages of the kNN. Some of these studies are based on finding k close samples in the attribute space. Pedrajas and Boyer [14] develop two method for boosting k-NN. Their two approaches both modify the view of the data that the classifier receives so that the accurate classification of difficult instances is favored. Liao and Vemuri, [13] proposes a new technique for develop a new kNN boosting algorithm. Their approach employs the k-Nearest Neighbor (kNN) classifier to categorize each new program behavior into either normal or intrusive classes. Liu et al [25] proposed a kNN based technique. They propose a critical time difference fall incident detection system to detect fall incident events. In addition, some studies focused on big data classification with the kNN method. García-Gil et al., [8] consider a study that focus on suitable noise filtering approach in big data domains. They apply two algorithms removing noisy examples composed of Random Forest, Logistic Regression and K-Nearest Neighbors (KNN) classifiers. [24] consider kNN based classifier for big data applications. They focus on grouped the data into some partitions to classify the whose data in accurate, fast, and robust manner. Also the re are some studies concentrated on the pre-processing of data by kNN as a filtering technique. Triguero et al., [16] apply kNN based filtering algorithm. They focus to mitigate the computational complexity, storage requirements and noise tolerance by eliminating redundant, irrelevant and noisy information. Maillo et al [21] consider the Spark framework due to its balance between scalability and accuracy that improves previous kNN proposals in the literature. Li et al [33] propose an algorithm named quantum kNN algorithm in order to speed up the classical algorithm. Considering the studies in the literature there are many techniques that tackle with class noise and conduct

effective classifications with extra many stages but it is still a need to develop easy implemented and high quality classifier which is aimed to classify noise data even with high level of noise.

In this study, we propose a new classification algorithm that is resistant to noisy data without making any prior application to the data. We generate a new noise robust kNN (k nearest neighborhood) based algorithm in order to classify any data that may include noise even at high percentage. We extend the classical kNN algorithm to provide the algorithm to show more noise robust performance. We focus on the algorithm is easy to implement and also make qualified classifications with even high level of noise. We conduct our experiments in the aim of mentioned ideas. We make numerical experiments for real datasets as for the traditional kNN method and proposed kNN method for different k parameters. The experimental results indicate that the proposed method has high performance and robust results on different datasets. Also, it is shown that the proposed algorithm is parameter independent. Moreover, all the comparisons are evaluated on the datasets when the class label of data samples (10%, 20%, 30%, 40%) are changed to be considered as noisy data. The results indicate that the method can be considered as a robust to noise method that its performance is comparable with other methods.

In the following section, the kNN algorithm is briefly described. In Section 4, the proposed algorithm with an illustrative example is given. Section 5 provides numerical results, and Section 6 concludes the study.

### 3. METHOD

#### 3.1. kNN Algorithm

The kNN is a supervised learning algorithm used for classification. Unlike other supervised learning algorithms, it does not have a training stage. The kNN algorithm includes a lazy type of learning in which training and testing stages are pretty much the same thing. It basically searches the closest points to the new point and classifies each point according to nearest neighborhoods. The parameter k represents the number of nearest neighbors of the unknown point. The mechanism of the kNN algorithm has been explained, but like all classification algorithms, it has some disadvantages. Those disadvantages are itemized below to highlight the important aspects of new n-kNN algorithm that we develop.

The classical kNN algorithm may cause several problems:

- As it is known, the kNN algorithm searches the distances between all points in the sample set. Then it labels a new point determined by the maximum number of labels in k nearest neighborhood set. This means that the kNN algorithm heavily relies on sample data which may include noise.
- Because of its sensitiveness to noise, the algorithm performs well on some data sets while not on others according to the noise level.

In our new n-kNN algorithm, we proposed solutions for these disadvantages as explained in the next section.

### 3.2. The Proposed n-kNN Algorithm

The n-kNN algorithm focuses on increasing its reliability, since it may decrease in dealing noise data due to its heavily dependence on data.

Suppose we have a data set with two classes, named A and B. Assume that the sample data set is given as  $sampleList = \{sampleList_1, sampleList_2, sampleList_3, \dots, sampleList_n\}$ . The goal is to design a mechanism that detects if a point in kth nearest neighborhood is noise. The algorithm may decide whether the point is noise or not. The algorithm classifies the data according to this information. The steps of n-kNN algorithm are presented below.

```

BEGIN
SET sampleList = ( 1 ≤ COUNT(sampleList) ≤ n);
INPUT unknownSample;

NUMERIC k = ( 1 ≤ k ≤ n);

SET nearestNeighbourSamples =
(nearestNeighbourSamples count = k);

NUMERIC i = 1, j=1;

// ---- SIGNING NOISE ----
FOR i = 1 to COUNT(sampleList)
  FOR j = 1 to COUNT(sampleList)

    estimatedClassSampleListi = CALL
    kNN(sampleListj, sampleListi);

    IF(estimatedClassSampleListi != sampleListi.class)
    THEN
      sampleListi.noise = TRUE;
    END IF
  END FOR
END FOR

// ---- CLASSIFICATION ----

FOR i = 1 to COUNT(sampleList)
  CALL EuclidDistance(sampleListi, unknownSample);

  IF (sampleListi.noise != TRUE) THEN
    IF ( i ≤ k) THEN
      INCLUDE sampleListi to the
      nearestNeighbourSamples;
    ELSE
      IF (sampleListi is closer unknownSample than
      nearestNeighbourSamples) THEN
        ELEMENATE the farthest neighbour in the
        nearestNeighbourSamples;

        INCLUDE sampleListi to the
        nearestNeighbourSamples;
      END IF
    END IF
  END IF
END FOR

DETERMINE the maximum class label in the
nearestNeighbourSamples;
ASSIGN unknownSample into the maximum class label
category;

```

END;

In the n-kNN algorithm, we perform a determination of noise labeling for each data points in the sample (even if the class of the point is already known). First, we aim to determine the distances between all sample points  $p_i$  and find k nearest neighbors for each. According to k neighbors, the number of points belongs to different classes was considered. If the number of points belonging to which class is higher in that neighborhood, the point to be classified is accepted as a point belonging to that class. According to this determination, we find the class of  $p_i$  and compared the original class with the one that we detected from neighborhoods. If the previously known class and the one calculated are not the same, then that point is labeled as noise. After that, resembling to the classical kNN algorithm structure,

the n-kNN algorithm determines the distances for the new point, then selects k nearest neighbors for the new data point. If there is a data point that is predicted as noise by n-kNN mechanism, then the algorithm do not take that point into consideration and selects the next nearest point instead of the noise point. Then, the algorithm decides the new point's class according to its new neighbours. The algorithm performs all these steps and stops.

In order to explain the performance difference between the n-kNN algorithm and classic k-NN algorithm, we present an illustrative example. In the example, two classes are defined, class A and class B. The class A is denoted with green squares and the class B with red pentagons. Let's assume that there is a blue circle point which is a test point to be classified. It actually belongs to class A and is named as  $p_0$ . For both kNN and n-kNN algorithms, neighborhoods for  $p_0$  are determined for different k parameters as  $k=\{3,5,7\}$ . When the performance of two algorithms is evaluated in the same example for  $k = 3$  both k-NN and n-kNN algorithms considers that  $p_0$  belongs to class A with 3 green squares in  $k=3$  neighborhood. For  $k=5$  both algorithms classify  $p_0$  as class A with 3 green squares and 2 red pentagons. But for  $k=7$ , the algorithms yield different results for  $p_0$ . According to the classical kNN for the  $k=7$  neighborhood, there are 4 red pentagons as class B and 3 green squares as class A. Thus, it is obtained that  $p_0$  belongs to class B. But according to the n-kNN algorithm, there are two noise points and those are marked as noise. Those two points are not considered to belong in the neighborhood and next two nearest points are searched. Consequently, for  $k=7$  neighborhood, the n-kNN algorithm classifies the  $p_0$  point as class A according to 5 green squares and 2 red pentagons. When all these are taken into account, without any noise consideration, there can be some misclassifications for any points. Our n-kNN algorithm takes into account the detection of noise points as it can be seen in Figure 1 and Figure 2.

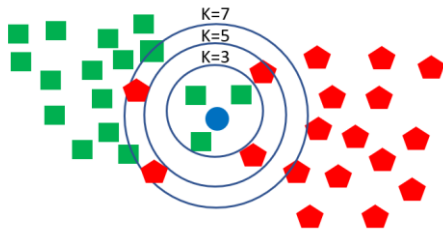


Figure 1. Illustrative example classified by classic kNN algorithm

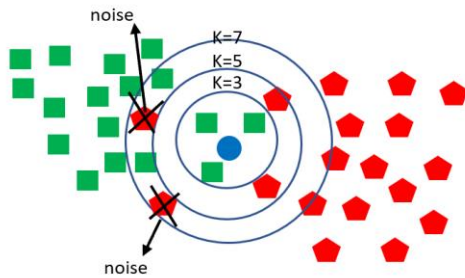


Figure 2. Illustrative example classified by n-kNN algorithm

Dataset	Short Names	# of points	# of features
Fertility	FERTILITY	100	9
Hepatitis	HEPATIT	142	19
Hearth	HEART	303	13
Ionosphere	ION	351	34
Wisconsin	WIS	683	10
Pima	PIMA	768	8
Data Banknote	DATA	1372	4
Spambase	SPAM	4601	57
Puma	PUMA	8192	32
Htru	HTRU	17898	8

Table 1 Description of the datasets.

#### 4. RESEARCH FINDINGS AND DISCUSSION

In this section, the numerical results of the n-kNN algorithm applied to real-world datasets are presented and compared with the results obtained by the popular methods in the literature. 10 binary datasets from the UCI repository have been used in our experiments [22]. The dataset information is given in Table 1. We perform 10-fold cross-validation method to conduct our

experiments for each dataset. We conduct numerical experiments according to three aspects: Firstly, it is aimed to show the

comparison between the accuracies of the proposed n-kNN algorithm and the classical kNN algorithm with different levels of noise (%0, %10, %20, %30, %40). In order to carry out the noise data experiments, we add random class noise with different levels using Weka filters. We follow a procedure described by Mantas and Abell'an (2014b) to add random class noise. We use the default parameter settings in Weka for applying the numerical results just as Mantas and Abell'an (2014b) applied. Also we use same dataset for each comparison. Secondly, In order to show the robustness of our algorithm we conduct experiments with different k parameter values (k= {3,5,7,9}). It is important to show that the accuracy results of the proposed algorithm are also successful for different parameter values, and the algorithm should be able to give consistent results regardless of the parameter value. Finally, we compared the n-kNN algorithm with different methods (support vector machines (SVM), Credit 4.5 (C 4.5), random forest (RF), artificial neural networks (ANN), Decision Tree (DT)) recently used in the literature. We apply those methods SVM, C 4.5, RF, ANN, DT algorithms via Weka tool as Mantas and Abell'an (2014b) applied. We use the default parameters on Weka to apply the analysis.

We use accuracy in our study as the performance measures in our comparisons. It is the mostly common measure for benchmarking in data mining field. It is based on a the consideration that a test sample could be either a false positive (FP), or a false negative (FN), or a true positive (TP), or a true negative (TN). If the classifier predicts the label of the positive and negative test samples correctly, they are named TP and TN, respectively. If the test sample is classified into a positive class, while it is negative, it is called false positive (FP). If it is classified as a negative, but it is positive, it is known as false negative (FN). Moreover, It is given the accuracy calculation according to Eq(1).

$$Accuracy = (TP + TN)/(TP + TN + FN + FP) \quad (1)$$

The experiments with 0%, 10%, 20%, 30% and 40% levels of noise in each dataset are performed to compare the n-kNN algorithm and the classical kNN algorithm for different k parameters. It is given in Table 2. Table 2 shows that better accuracy results are obtained by n-kNN algorithm when comparing classical kNN algorithm with different k parameter values. Although the values of the parameter k changes, the n-

Table 2 Noise data test accuracy results for kNN and n-kNN for different noise levels and k values

DATA SET	NOISE LEVEL %	KNN k=3	n-kNN k=3	kNN k=5	n-kNN k=5	kNN k=7	n-kNN k=7	kNN k=9	n-kNN k=9
ION	%0	<b>84.33</b>	82.34	84.90	84.90	83.48	<b>84.33</b>	83.20	83.77
	%10	83.19	82.91	84.04	<b>85.19</b>	83.48	84.06	83.77	83.48
	%20	76.94	82.91	78.91	84.05	80.63	84.05	80.64	<b>84.06</b>
	%30	70.08	84.33	72.36	84.33	74.91	84.33	76.05	<b>84.34</b>
	%40	60.36	83.18	62.93	84.04	60.64	<b>84.33</b>	64.94	82.90
WIS	%0	<b>97.21</b>	96.92	97.65	<b>97.21</b>	97.07	96.92	97.22	96.77
	%10	93.70	96.92	96.19	96.78	96.19	<b>97.07</b>	96.49	96.92
	%20	86.62	<b>96.92</b>	91.21	96.78	93.99	96.63	95.46	96.33
	%30	77.45	<b>97.36</b>	80.98	96.33	86.66	96.33	88.86	96.48
	%40	60.04	<b>96.63</b>	62.95	96.04	68.94	95.75	73.78	96.04
HEPATITIS	%0	75.43	<b>81.81</b>	79.00	80.38	79.00	81.81	79.67	<b>81.81</b>
	%10	70.62	<b>81.81</b>	74.14	79.67	76.29	<b>81.81</b>	74.95	<b>81.81</b>
	%20	65.67	<b>82.52</b>	70.62	81.81	72.10	81.81	75.62	81.81
	%30	60.48	81.10	62.05	<b>81.81</b>	67.71	<b>81.81</b>	71.19	<b>81.81</b>
	%40	55.57	<b>80.38</b>	57.57	<b>81.81</b>	60.43	<b>81.81</b>	63.95	<b>81.81</b>
HEART	%0	62.32	<b>64.71</b>	62.02	64.34	59.39	65.35	64.34	66.32
	%10	62.02	62.70	62.71	<b>63.70</b>	59.39	65.68	63.69	65.34
	%20	55.81	63.73	60.12	<b>64.70</b>	58.74	63.74	60.76	63.4
	%30	57.75	<b>63.39</b>	58.08	63.38	58.4	63.39	58.76	63.38
	%40	56.12	61.71	55.82	<b>63.09</b>	56.15	61.42	56.8	62.41
FERTILITY	%0	87.00	87.00	87.00	88.00	88.00	88.00	88.00	88.00
	%10	84.00	<b>88.00</b>	85.00	88.00	86.00	88.00	87.00	88.00
	%20	75.00	86.00	77.00	<b>88.00</b>	79.00	<b>88.00</b>	79.00	<b>88.00</b>
	%30	70.00	<b>89.00</b>	72.00	88.00	74.00	88.00	77.00	88.00
	%40	52.00	86.00	56.00	<b>88.00</b>	56.00	<b>88.00</b>	67.00	<b>88.00</b>
SPAM	%0	<b>80.61</b>	78.79	80.13	78.31	79.87	77.55	79.85	76.98
	%10	77.70	78.27	78.40	78.11	78.44	77.03	<b>78.59</b>	76.22
	%20	72.74	<b>77.77</b>	73.70	77.37	75.35	76.66	75.61	75.92
	%30	67.77	<b>77.72</b>	68.94	77.66	70.46	76.4	71.77	75.85



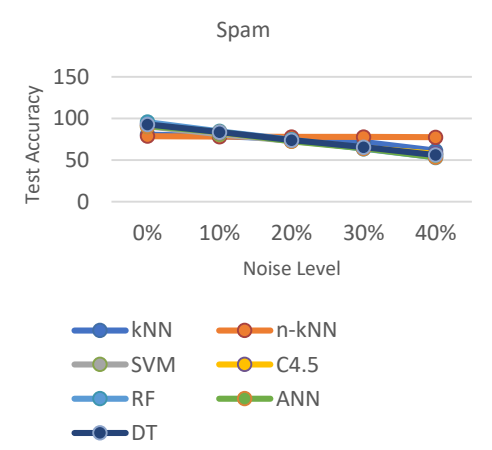
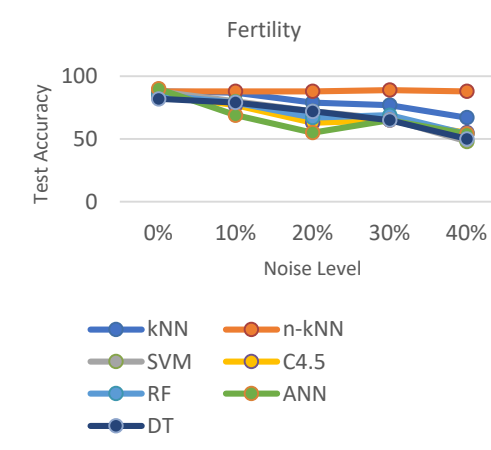
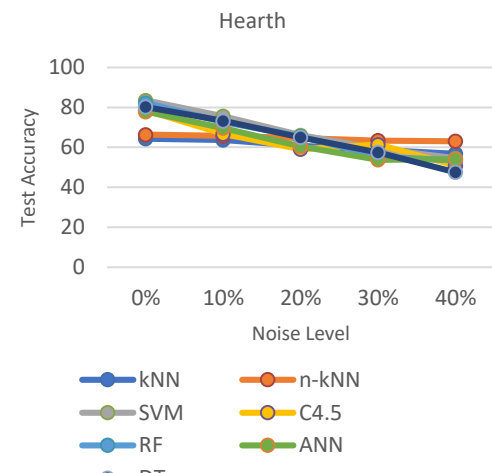
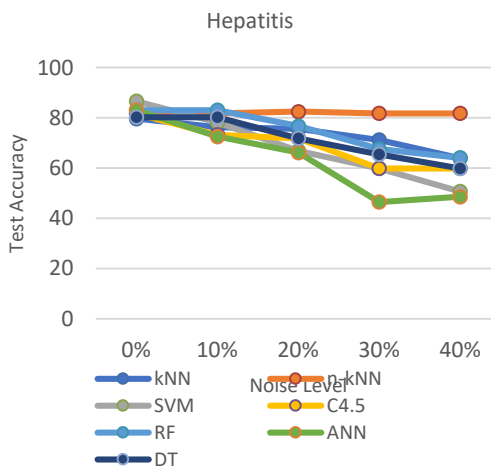
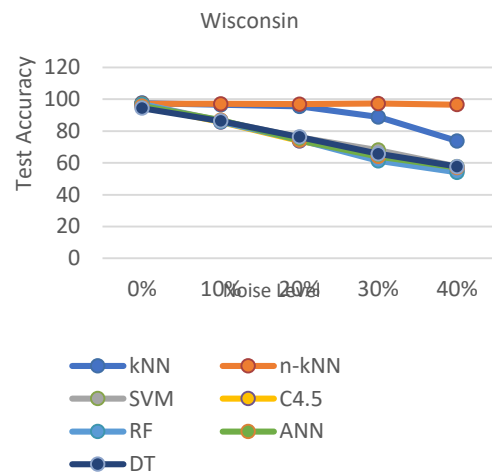
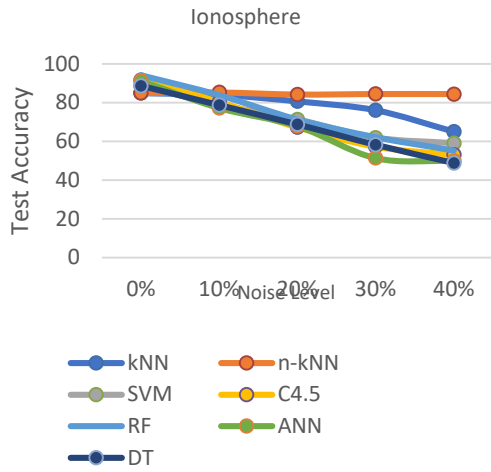
	<b>%40</b>	58.62	<b>77.37</b>	59.47	76.55	61.25	75.66	60.77	75.33
<b>PUMA</b>	<b>%0</b>	57.85	61.97	59.22	63.45	60.28	64.39	61.29	<b>64.50</b>
	<b>%10</b>	56.24	61.69	57.42	63.61	58.35	64.12	59.02	<b>64.33</b>
	<b>%20</b>	54.38	61.94	55.71	63.68	56.62	64.22	57.68	<b>64.48</b>
	<b>%30</b>	54.04	62.11	55.05	62.99	55.47	63.90	55.6	<b>64.32</b>
	<b>%40</b>	51.68	60.90	51.23	62.61	51.70	63.48	52.31	<b>63.87</b>
<b>PIMA</b>	<b>%0</b>	64.99	64.99	64.99	64.99	64.99	64.99	64.99	64.99
	<b>%10</b>	61.05	61.35	61.05	<b>64.99</b>	58.93	<b>64.99</b>	61.5	<b>64.99</b>
	<b>%20</b>	60.88	64.08	59.82	<b>64.99</b>	64.53	<b>64.99</b>	64.53	<b>64.99</b>
	<b>%30</b>	62.26	64.99	56.66	<b>64.99</b>	58.31	<b>64.99</b>	56.92	<b>64.99</b>
	<b>%40</b>	52.87	61.05	52.24	<b>64.99</b>	51.74	<b>64.99</b>	58.00	<b>64.99</b>
<b>DATA BANKNOTE AUTH</b>	<b>%0</b>	<b>96.86</b>	96.43	96.43	95.77	96.64	94.75	95.55	94.09
	<b>%10</b>	93.87	<b>95.99</b>	95.19	95.19	95.48	94.31	94.97	94.46
	<b>%20</b>	86.51	<b>95.77</b>	90.16	94.82	91.91	94.46	92.93	94.09
	<b>%30</b>	77.18	<b>95.41</b>	82.28	94.02	83.67	94.17	84.76	93.80
	<b>%40</b>	64.07	<b>95.55</b>	66.19	94.97	68.59	94.31	72.16	93.51
<b>HTRU</b>	<b>%0</b>	<b>97.18</b>	97.04	97.17	96.96	97.14	96.9	97.11	96.82
	<b>%10</b>	94.03	<b>97.01</b>	95.85	96.88	96.63	96.84	96.86	96.82
	<b>%20</b>	86.71	<b>97.01</b>	91.06	96.89	93.36	96.83	94.68	96.74
	<b>%30</b>	76.06	<b>97.00</b>	80.99	96.85	84.46	96.79	86.65	96.73
	<b>%40</b>	63.26	<b>96.95</b>	66.49	96.86	69.37	96.78	71.36	96.71

Table 3 Test accuracy results for benchmarking

DATA SET	NOISE LEVEL%	BEST classical kNN	BEST n-KNN	SVM	C4.5	RF	ANN	DT
<b>ION</b>	<b>%0</b>	84.90	84.90	88.60	91.45	<b>94.02</b>	91.17	88.60
	<b>%10</b>	84.04	<b>85.19</b>	78.92	80.34	83.76	77.21	78.63
	<b>%20</b>	80.64	<b>84.06</b>	71.23	67.24	71.23	67.81	68.66
	<b>%30</b>	76.05	<b>84.34</b>	61.82	57.27	62.11	51.28	58.12
	<b>%40</b>	64.94	<b>84.33</b>	58.97	52.99	55.27	49.86	48.72
<b>WIS</b>	<b>%0</b>	<b>97.65</b>	97.21	96.93	96.05	96.93	95.75	94.44
	<b>%10</b>	96.49	<b>97.07</b>	86.53	85.65	85.94	86.82	86.38
	<b>%20</b>	95.46	<b>96.92</b>	75.99	73.79	74.52	74.23	76.28

	<b>%30</b>	88.86	<b>97.36</b>	68.08	64.28	61.35	63.98	65.74
	<b>%40</b>	73.78	<b>96.63</b>	57.69	55.20	54.03	57.10	57.69
<b>HEPATITIS</b>	<b>%0</b>	79.67	81.81	<b>86.62</b>	81.69	83.10	83.10	80.28
	<b>%10</b>	76.29	81.81	78.17	73.24	<b>83.10</b>	72.54	80.28
	<b>%20</b>	75.62	<b>82.52</b>	66.90	71.83	76.76	66.20	71.83
	<b>%30</b>	71.19	<b>81.81</b>	59.86	59.86	67.61	46.48	65.49
	<b>%40</b>	63.95	<b>81.81</b>	50.7	59.86	64.09	48.59	59.86
<b>HEART</b>	<b>%0</b>	64.34	66.32	<b>83.5</b>	78.55	82.18	77.89	80.2
	<b>%10</b>	63.69	65.68	<b>75.58</b>	66.67	72.61	69.64	73.27
	<b>%20</b>	60.76	64.70	<b>66.01</b>	59.08	65.35	60.40	65.02
	<b>%30</b>	58.76	<b>63.39</b>	59.41	61.39	55.45	53.80	57.43
	<b>%40</b>	56.80	<b>63.09</b>	54.46	50.83	53.47	54.13	47.53
<b>FERTILITY</b>	<b>%0</b>	88.00	88.00	88.00	85.00	85.00	<b>90.00</b>	82.00
	<b>%10</b>	87.00	<b>88.00</b>	80.00	77.00	79.00	69.00	79.00
	<b>%20</b>	79.00	<b>88.00</b>	72.00	63.00	67.00	55.00	72.00
	<b>%30</b>	77.00	<b>89.00</b>	66.00	65.00	69.00	65.00	65.00
	<b>%40</b>	67.00	<b>88.00</b>	48.00	55.00	54.00	54.00	50.00
<b>SPAM</b>	<b>%0</b>	80.61	78.79	90.44	92.98	<b>95.65</b>	90.98	92.76
	<b>%10</b>	78.59	78.27	80.90	83.13	<b>84.48</b>	83.55	83.48
	<b>%20</b>	75.61	<b>77.77</b>	72.70	73.79	74.61	72.53	73.88
	<b>%30</b>	71.77	<b>77.72</b>	64.68	64.68	63.55	64.01	65.44
	<b>%40</b>	61.25	<b>77.37</b>	57.01	57.57	53.05	53.29	56.03
<b>PUMA</b>	<b>%0</b>	61.29	64.50	65.21	86.21	<b>88.26</b>	84.45	87.50
	<b>%10</b>	59.02	64.33	62.18	79.08	<b>80.44</b>	73.17	80.30
	<b>%20</b>	57.68	64.48	58.13	<b>72.30</b>	71.70	62.67	72.29
	<b>%30</b>	55.60	64.32	52.65	<b>64.33</b>	61.77	52.43	64.22
	<b>%40</b>	52.31	<b>63.87</b>	49.84	56.8	53.69	49.51	55.19
<b>PIMA</b>	<b>%0</b>	64.99	64.99	<b>77.34</b>	73.83	75.52	75.13	75.52
	<b>%10</b>	61.50	64.99	70.31	<b>68.49</b>	67.06	<b>68.49</b>	67.45

	<b>%20</b>	64.53	<b>64.99</b>	64.71	62.76	62.11	61.46	60.42
	<b>%30</b>	62.26	<b>64.99</b>	54.82	52.6	52.73	53.78	55.08
	<b>%40</b>	58.00	<b>64.99</b>	52.73	52.99	43.49	49.22	51.04
<b>DATA BANKNOTE AUTH</b>	<b>%0</b>	96.86	96.43	98.03	98.54	99.34	<b>99.93</b>	98.47
	<b>%10</b>	95.48	<b>95.99</b>	88.78	88.12	88.85	89.87	88.19
	<b>%20</b>	92.93	<b>95.77</b>	78.94	76.53	75.58	79.37	77.7
	<b>%30</b>	84.76	<b>95.41</b>	68.95	66.69	62.83	68.22	67.93
	<b>%40</b>	72.16	<b>95.55</b>	59.84	57.14	53.57	57.51	59.69
<b>HTRU</b>	<b>%0</b>	97.18	97.04	97.56	97.84	<b>98.02</b>	97.97	97.8
	<b>%10</b>	96.86	<b>97.01</b>	86.98	88.18	88.2	88.28	88.17
	<b>%20</b>	94.68	<b>97.01</b>	76.99	78.51	78.28	78.6	78.61
	<b>%30</b>	86.65	<b>97.00</b>	67.77	69.16	67.89	69.25	69.08
	<b>%40</b>	71.36	<b>96.95</b>	58.75	59.58	56.5	59.54	59.31



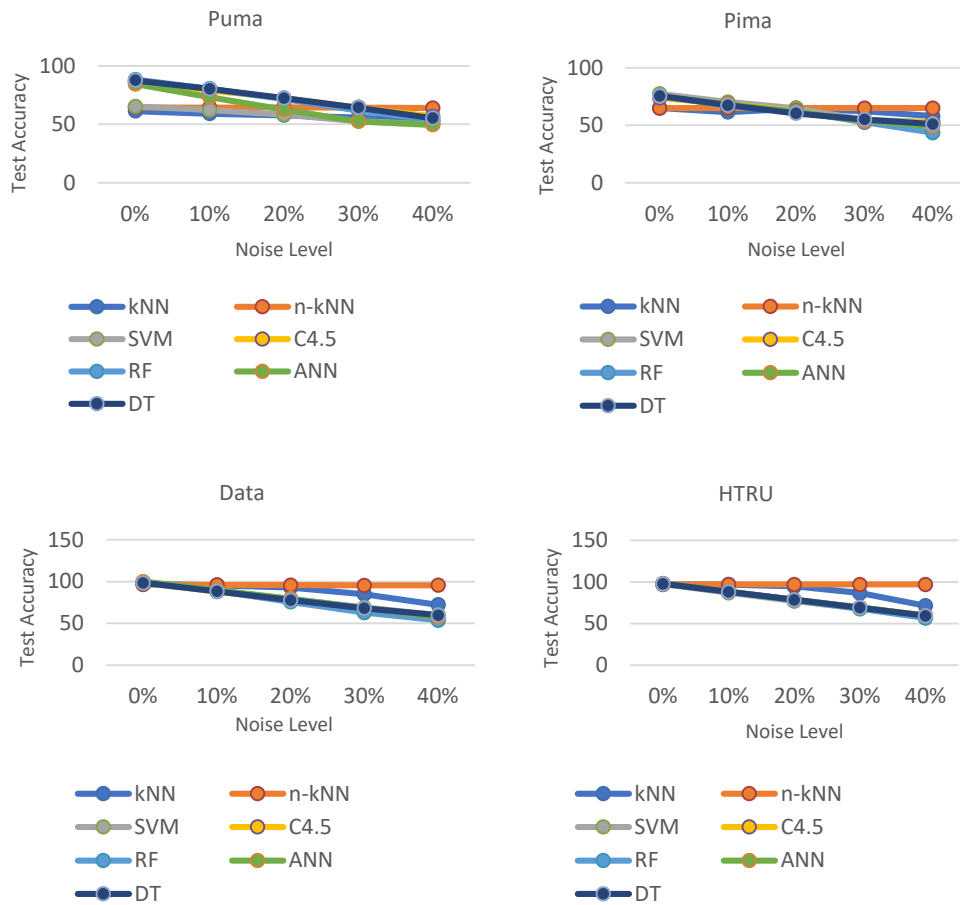


Figure 3. Noise level relationships between the n-kNN and other algorithms on different datasets

kNN algorithm gives better results than the kNN algorithm. Also, as can be seen in the Table 2 that the accuracy values decreases rapidly in classical kNN as the noise level increases. In the n-kNN algorithm, we can see a smaller decrease than classical kNN. Moreover, it can be seen that better accuracy results are obtained with n-KNN algorithm for many datasets. We can make some inferences according to the results obtained from Table 2. Firstly, the proposed kNN algorithm is more resistant to noise level than classical kNN because as noise level increases accuracy results

decreases fastly for kNN algorithm but according to the results of n-kNN algorithm results decreases slowly than classical kNN. Secondly, better accuracy results are obtained by n-kNN comparing with classical kNN. So it can be said that our proposed algorithm has a better performance than classical kNN. Thirdly, for different k values, n-kNN has still the better results than kNN and also better results can change according to the value of k parameter's value. This means that our algorithm is independent from the parameter values.

We also conducted further experiments to benchmark the performance of the n-kNN algorithm with some of the best-known algorithms in the literature (Table 3). We apply most popular classification methods from the literature as SVM, C 4.5, RF, ANN, DT algorithms via Weka tool and take best k values in order to make the comparison. It can be seen that for most of the

noise levels and for various datasets, the n-kNN provided best accuracy results than other methods. Also, the results indicated that when the noise level increased, the test accuracy values did not change much in the n-kNN algorithm while decreased evidently in the other algorithms tested. This is also can be seen from Figure 3.

We visualize our experiments with the graphs it can be seen in Figure 3. We compare our proposed algorithm and other algorithms for different noise levels and test accuracies. It can be obtained that n-kNN algorithm does not changes much as accuracy results when noise level increases. But other algorithms decreases dramatically when the noise level increases. According to this results, we can say that the proposed algorithm is more robust to noise data than the compared algorithms and it is more reliable for binary classification with noisy data.

### 5. CONCLUSION

This paper proposes an effective binary classification algorithm named n-kNN algorithm that is based on the well-known kNN algorithm. We suggest that our algorithm is resistant to noisy data and has reliable performance with high levels of noise data. To approve the performance we conduct numerical experiments by comparing the classical kNN and our new algorithm. We achieve significantly higher test accuracy values compared to the classical

kNN algorithm, with results reaching 84.33%, 93.63%, 81.81%, and 88.00% on real-world datasets. By these numerical results it is also confirmed that in noisy environments, the n-kNN algorithm can be a reliable classifier as the algorithm is not heavily affected by noise in the data. Furthermore, according to the experiments performed on real-world datasets, comparing proposed algorithm and several other methods in the literature based on the popular algorithms, most of the best accuracy results observed by our n-kNN algorithm reaching up to 85.19%, 97.07%, and 96.63%. Furthermore, it can be seen that in different values of the parameter k have better results and this means that the n-kNN algorithm is not dependent on the parameter k. Thus, it can be concluded that the n-kNN is a reliable algorithm because its accuracy results do not change much, stay stable in other words, with increasing noise levels. In this study, we apply our algorithm to binary classification data. However, in the future, we plan to modify and apply it to the data containing more than two classes. We aim to study the multi-class version of the n-kNN algorithm that can be applied more types of datasets. As a future work, we also will explore the data coming from the following popular areas and make some application on different fields cases.

## 6. Author contribution statements

In the scope of this study, the Author 1 in the formation of the idea, the design and the literature review, assessment of obtained results, supplying the materials used and examining the results; the spelling and checking the article in terms of content were contributed.

## 7. Ethics committee approval and conflict of interest statement

There is no conflict of interest with any person / institution in the article prepared. There is no need for an ethics committee approval in the prepared article.



## REFERENCES

- [1] Bootkrajang J. "A generalised label noise model for classification in the presence of annotation errors." *Neurocomputing*, 192, 61-71, 2016.
- [2] García LP, De Carvalho AC, Lorena AC. "Effect of label noise in the complexity of classification problems." *Neurocomputing*, 160, 108-119, 2015.
- [3] Sáez JA, Galar M, Luengo J, Herrera, F. "Tackling the problem of classification with noisy data using multiple classifier systems: analysis of the performance and robustness." *Information Sciences*, 247, 1-20, 2013.
- [4] Sáez, JA, Corchado, E. "ANCES: A novel method to repair attribute noise in classification problems." *Pattern Recognition*, 121, 108198, 2022
- [5] Zhu X, Wu X. "Class noise vs. attribute noise: A quantitative study." *Artificial Intelligence Review*, 22(3), 177-210, 2004.
- [6] Sluban B, Lavrač, N. "Relating ensemble diversity and performance: A study in class noise detection." *Neurocomputing*, 160, 120-131, 2015.
- [7] Luengo J, Sánchez-Tarragó D, Prati RC, Herrera F. "Multiple instance classification: Bag noise filtering for negative instance noise cleaning." *Information Sciences*, 579, 388-400, 2021.
- [8] García-Gil D, Luengo J, García S, Herrera F. "Enabling smart data: noise filtering in big data classification." *Information Sciences*, 479, 135-152, 2019.
- [9] Wang ZY, Luo XY, Liang J. "A Label Noise Robust Stacked Auto-Encoder Algorithm for Inaccurate Supervised Classification Problems." *Mathematical Problems in Engineering*, 2019.
- [10] Marsala C, Petturiti D. "Rank discrimination measures for enforcing monotonicity in decision tree induction." *Information Sciences*, 291, 143-171, 2015.
- [11] Zhu J, Liao S, Lei Z, Li S Z "Multi-label convolutional neural network based pedestrian attribute classification." *Image and Vision Computing*, 58, 224-229, 2017.
- [12] Chao L, Zhipeng J, Yuanjie Z. "A novel reconstructed training-set SVM with roulette cooperative coevolution for financial time series classification." *Expert Systems with Applications*, 123, 283-298, 2019.
- [13] Liao Y, Vemuri VR. "Use of k-nearest neighbor classifier for intrusion detection." *Computers & Security*, 21(5), 439-448, 2002.
- [14] García-Pedrajas N, Ortiz-Boyer D. "Boosting k-nearest neighbor classifier by means of input space projection." *Expert Systems with Applications*, 36(7), 10570-10582, 2009.
- [15] Wang ZY, Luo XY, Liang J. "A Label Noise Robust Stacked Auto-Encoder Algorithm for Inaccurate Supervised Classification Problems." *Mathematical Problems in Engineering*, 2019.
- [16] Triguero I, García-Gil D, Mailló J, Luengo J, García S, Herrera F. "Transforming big data into smart data: An insight on the use of the k-nearest neighbors algorithm to obtain quality data." *Wiley Interdisciplinary Reviews: Data Mining and Knowledge Discovery*, 9(2), e1289, 2019.
- [17] Mantas C J, Abellan J. "Analysis and extension of decision trees based on imprecise probabilities: Application on noisy data." *Expert Systems with Applications*, 41(5), 2514-2525, 2014a.
- [18] Alam MM, Gazuruddin M, Ahmed N, Motaleb A, Rana M, Shishir RR, Rahman RM., "Classification of deep-SAT images under label noise. Applied" *Artificial Intelligence*, 35(14), 1196-1218, 2021.
- [19] Mantas CJ, Abellan J. "Credal-C4.5 decision tree based on imprecise probabilities to classify noisy data." *Expert Systems with Applications*, 41(10), 4625-4637, 2014b.
- [20] Mantas, C. J., Abellan, J., & Castellano, J. G. "Analysis of Credal-C4.5 for classification in noisy domains." *Expert Systems with Applications*, 61, 314-326, 2016.
- [21] Mailló J, García S, Luengo J, Herrera, F, Triguero, I. "Fast and scalable approaches to accelerate the fuzzy k-Nearest neighbors classifier for big data." *IEEE Transactions on Fuzzy Systems*, 28(5), 874-886, 2019.
- [22] Dua D, Graff C. "UCI Machine Learning Repository" [<http://archive.ics.uci.edu/ml>]. Irvine, CA: University of California, School of Information and Computer Science. 2019.

- [23] Hall M, Frank E, Holmes G, Pfahringer B, Reutemann P, Witten IH, "The WEKA data mining software: an update". *ACM SIGKDD explorations newsletter*, 11(1), 10-18, 2009.
- [24] Shokrzade A, Ramezani M, Tab FA, Mohammad, MA. "A novel extreme learning machine based kNN classification method for dealing with big data." *Expert Systems with Applications*, 115293, 2021.
- [25] Liu CL, Lee CH, Lin PM. "A fall detection system using k-nearest neighbor classifier." *Expert systems with Applications*, 37(10), 7174-7181, 2010.
- [26] Catal C., "Software fault prediction: A literature review and current trends." *Expert Systems with Applications*, 38(4), 4626-4636, 2011
- [27] Yıldırım S, Yıldız T. "Türkçe için karşılaştırmalı metin sınıflandırma analizi" *Pamukkale Üniversitesi Mühendislik Bilimleri Dergisi*, 24(5), 879-886, 2018
- [28] Sağlam A, Baykan NA. "Continuous time threshold selection for binary classification on polarized data" *Pamukkale Üniversitesi Mühendislik Bilimleri Dergisi* 25(5), 596-602, 2019
- [29] Goodfellow, I., Bengio, Y., & Courville, A. *Deep learning*. MIT Press.,2016
- [30] Bishop, C. M. *Pattern recognition and machine learning*. Springer, 2006
- [31] Mansour, R. F., Abdel-Khalek, S., Hilali-Jaghdam, I., Nebhen, J., Cho, W., & Joshi, G. P. An intelligent outlier detection with machine learning empowered big data analytics for mobile edge computing. *Cluster Computing*, 1-13. 2023.
- [32] Dash, C. S. K., Behera, A. K., Dehuri, S., & Ghosh, A. An outliers detection and elimination framework in classification task of data mining. *Decision Analytics Journal*, 6, 100164. 2023
- [33] Li, J., Zhang, J., Zhang, J., & Zhang, S., Quantum KNN classification with K Value selection and neighbor selection. *IEEE Transactions on Computer-Aided Design of Integrated Circuits and Systems*. 2023

# İnşaat Şirketi Müşterilerinin Gelecekteki Konut Satın Alma Davranışlarının Metin Madenciliği ve Makine Öğrenmesi ile Tahmin Modellerinin Oluşturulması

*Araştırma Makalesi/Research Article*

 Haydar EKELİK<sup>1\*</sup>,  Şenol EMİR<sup>2</sup>

<sup>1</sup>İstanbul Üniversitesi (Ekonometri Bölümü, İstanbul Üniversitesi, İstanbul, Türkiye)

<sup>2</sup>İstanbul Üniversitesi (Ekonometri Bölümü, İstanbul Üniversitesi, İstanbul, Türkiye)

[haydar.ekelik@istanbul.edu.tr](mailto:haydar.ekelik@istanbul.edu.tr), [senol.emir@istanbul.edu.tr](mailto:senol.emir@istanbul.edu.tr)

(Geliş/Received:14.05.2024; Kabul/Accepted:19.10.2024)

DOI: 10.17671/gazibtd.1484123

**Özet**— Bu çalışmada, inşaat sektöründe faaliyet gösteren bir işletmenin müşterileriyle yüz yüze veya telefonla yapılan görüşmelerinin kayıtlarına çeşitli metin madenciliği ve makine öğrenmesi teknikleri uygulanmıştır. Temel amaç, bu metin tabanlı doküman kümesinden (korpus), yeni görüşme yapılan herhangi bir müşterinin ileride şirketten konut satın almayacağını doğru bir şekilde tahmin edebilecek bir model geliştirmektir. Bu amaçla metinsel verilere bir takım veri ön işleme aşamaları uygulandıktan sonra anahtar kelimeler ve vektör uzay modeli oluşturmuş ve metin tabanlı veri analize uygun formata dönüştürülmüştür. CART(Classification And Regression Tree), RF(Random Forest) ve XGBoost(eXtreme Gradient Boosting) makine öğrenmesi yöntemleri uygulanarak farklı tahmin modelleri oluşturulmuş ve daha sonra bu modeller farklı sınıflandırma ölçütlerine göre karşılaştırılmıştır. Sınıflandırma problemlerinde sınıflardaki gözlem sayıları arasında dengesizlikler olması durumunda yaygın sınıflandırma ölçütlerine göre modellerin karşılaştırılması yanlış sonuçlar verebilmektedir. Bu nedenle literatürde bu gibi durumlar için genel karşılaştırma ölçütlerine ek olarak yeni ölçütler geliştirilmiştir. Çalışmadaki uygulamada da sınıflar arası dengesizlik olduğundan bu ölçütlerden birisi olan PR (Precision- Recall) eğrileri kullanılmıştır. Analiz sonucunda, PR eğrileri dikkate alındığında, görüşme yapılan yeni müşterilerin ileride konut almayacağını en iyi tahmin eden yöntemin Random Forest olduğu görülmüştür.

**Anahtar Kelimeler**— metin madenciliği, pazarlama, random forest, CART, XGBoost

## Predictive Modeling of Future Home Purchase Behavior in Construction Company Customers Using Text Mining and Machine Learning

**Abstract**—In this study, various text mining and machine learning techniques were applied to the recordings of face-to-face or telephone interviews with customers of a company operating in the construction industry. The main objective is to develop a model from this set of text-based documents (corpus) that can accurately predict whether a new customer interviewed will purchase a house from the company in the future. For this purpose, a number of data preprocessing steps were applied to the textual data, then keywords and vector space model were created and the text-based data were converted into a format suitable for further analysis. Different prediction models were created by applying CART(Classification And Regression Tree), RF(Random Forest) and XGBoost(eXtreme Gradient Boosting) methods and then these models were compared according to different classification metrics. In classification problems, imbalances between classes make it difficult to compare models. For this reason, in literature new metrics have been developed in addition to the classical performance metrics. Since there is an imbalance between classes in the application in this study, PR (precision-recall) curves, one of the developed criteria, were used. As a result of the analysis, when the PR curves are taken into account, it is seen that Random Forest shows the best performance for predicting whether interviewed new customers will buy a house in the future.

**Keywords**—text mining, marketing, random forest, CART, XGBoost



## 1. GİRİŞ (INTRODUCTION)

Günümüzde bilgi ve internet teknolojilerindeki gelişmeler ve bu teknolojilerin yaygınlaşmasıyla analizlerde kullanılacak veri kaynaklarının her geçen gün çeşitlendiğini gözlemlemekteyiz. İnternet kullanım oranlarında artış ile gerek internet sitelerinin gerekse sosyal medyanın kullanımının hızla arttığına tanık olmaktadır. Bütün bu gelişmeler bir yandan ulaşılabilen verinin hacimce büyümesini sağlarken bir yandan da verinin çeşitliliğini artırarak araştırmacılara yapısal verinin (structured data) yanında yapısal olmayan veri (unstructured data) ve yarı yapısal veri (semi-structured data) türleri ile analiz yapma gereksinimi doğurmaktadır.

Yapısal olmayan veri, en kolay toplanabilecek veri türüdür [1]. Yapısal olmayan veri, analiz edilmesi için standart bir formatı olmayan veri olarak tanımlanır. Genellikle metinsel veriler, görseller, ses ve video dosyaları bu veri türüne örnek gösterilebilir [2]. Son yıllarda çeşitli sosyal ağlarda, web'de ve diğer bilgi işlem uygulamaları tabanlı uygulamalarda oluşturulan büyük miktarda metin verisi nedeniyle çeşitli metin verilerini etkin bir şekilde işleyebilecek yöntem ve algoritmalara ilgi giderek artmaktadır [1]. Yapısal olmayan verinin saklanması, sorgulanması ve analiz edilerek değer üretilmesi aşamalarında kullanılan teknikler geleneksel yöntemlerden farklıdır. Metin verisinden anlamlı bilgiler çıkartılarak değer üretilmesi aşamasında kullanılacak yöntemlerin başlıcalarından biri metin madenciliği (text mining) yöntemleridir.

Metin madenciliği, metin verisi üzerine klasik veri madenciliği yöntemlerinin uygulanması gibi basit tanımlardan “dünya hakkındaki gerçekleri ve yeni eğilimleri keşfetmek için büyük çevrimiçi metin verilerinin kullanılması” şeklinde sofistike tanımlara kadar geniş bir yaklaşım ve yöntem alanını kapsamaktadır [3]. Metin madenciliği genel olarak, veri madenciliği, dil bilimi, hesaplamalı istatistik ve bilgisayar bilimi ile etkileşimde olan disiplinler arası bir uygulama alanıdır. Metin madenciliği metin sınıflandırma (text classification), metin kümeleme (text clustering), belge özetleme (document summarization), ontoloji ve taksonomi yaratma (ontology and taxonomy), gizil korpus analizi (latent corpus analysis) gibi kendine özgü yöntemlerin yanı sıra ilişkili olduğu disiplinlerde kullanılan pek çok yöntem ve teknikten de yararlanmaktadır [4]. Metin madenciliği başka bir deyişle metin verisinden bilgi keşfi ilk kez Feldman ve Dagan (1995) tarafından metnin makine destekli analizi olarak tanımlanmıştır [5]. Metin madenciliği doğal dil işleme teknikleri ile birlikte bilgi elde etme, bilgi

çıkarma süreçlerini kullanarak veri madenciliği, makine öğrenmesi, istatistiksel yöntemler ve algoritmalara bağlayan bir süreç olarak da görülebilir. Bu açıdan metin madenciliği, veri tabanından bilgi keşfi olarak da tanımlanan veri madenciliği ile benzer bir prosedürdür. Veri madenciliğinde veriye odaklanılırken metin madenciliğinde analizin temelini metinsel verileri içeren belgelerin kümesi (korpus) oluşturmaktadır [6].

Veri madenciliği ile metin madenciliği arasında kavramsal olarak kapsayıcı bir ilişki vardır. Literatürde metin madenciliğinin veri madenciliğinin bir alt dalı olarak da kabul edilebileceği yönünde yaygın bir görüş birliği vardır. Metin madenciliği de veri madenciliği gibi, ilgi çekici örüntülerin tanımlanması ve araştırılması yoluyla veri kaynaklarından faydalı bilgiler çıkarmayı amaçlamaktadır. Ancak veri madenciliğinden farklı olarak metin madenciliğindeki veri kaynakları yapılandırılmamış metin verileridir [7]. Veri madenciliğinde örüntülerin keşfi için yapılandırılmış veri tabanlarından kullanılırken metin madenciliğinde bilginin kaynağı çoğunlukla yapılandırılmamış doğal dil metinleridir [8]. Bu durumda iki kavram arasındaki temel ayrımın metin madenciliğindeki bilgi keşfi sürecinin doğal dil kalıplarının kullanılarak gerçekleştirilmesi olduğunu söylemek yanlış olmayacaktır. Metin madenciliğinin alt konuları ise metin sınıflandırma, metin kümeleme ve metni benzer metinlerle ilişkilendirme olarak genellenebilir [1].

Bu çalışmada inşaat sektöründe faaliyet gösteren ve konut satışı yapan bir firmanın müşterileriyle yaptığı görüşmelerin derlendiği metin tabanlı dosyalar metin madenciliği ve makine öğrenmesi yöntemleri kullanılarak analiz edilmiştir. Bu sayede bir müşteri ile görüşme yapıldıktan sonra ileride ilgili müşterinin konut satın alıp almayacağını tahmin eden bir model oluşturulmaya çalışılmıştır. Çalışmada metin madenciliği konusunda genel bir bakış açısı sunulmuş, müşteri ilişkileri yönetimi ve pazarlama konularında işletmelerin metin madenciliği ve makine öğrenmesi yöntemlerinin birlikte kullanılmalarından nasıl faydalanabilecekleri incelenmiştir.

Çalışmanın ikinci bölümünde literatür taraması, üçüncü bölümünde veri kümesi ve uygulanan yöntemler hakkında detaylı bilgi, dördüncü bölümünde analiz sonuçları yer almaktadır. Son bölümde ise sonuçların değerlendirilmiştir.

## 2. LİTERATÜR TARAMASI (LITERATURE REVIEW)

Metin madenciliği teknikleri ile makine öğrenmesi yöntemlerinin birlikte kullanıldığı birçok çalışma

bulunmaktadır. Farklı anlarda yapılan bu çalışmalardan birkaçı bu kısımda sunulmuştur. Ayrıca spesifik olarak inşaat yönetimi veya pazarlaması konularında metin madenciliğinin makine öğrenmesi ile birlikte kullanıldığı çalışmalar da bu literatür taramasının sonunda yer almaktadır.

Hosseini vd. (2023) tarafından yapılan çalışmada polis tarafından tutulan metin tabanlı kaza tutanakları incelenerek bir kazanın ters yönde sürüş sonucu oluşup oluşmadığını belirlemek için makine öğrenmesi yöntemlerinden faydalanılmıştır. Doğal dil işleme işlemleri için Bidirectional Encoder Representations from Transformers (BERT) modelleri kullanılmıştır. Daha sonra, Naïve Bayes, Support Vector Machine, Decision Tree, Random Forest, and Single Layer Perceptron makine öğrenmesi yöntemleri kullanılarak kazanın ters yönde sürüşten kaynaklanan bir kaza olup olmadığı sınıflandırılmıştır. Her bir sınıflandırma algoritmasının performansını değerlendirmek için çapraz doğrulama ve farklı performans ölçütleri kullanılmıştır. Oluşturulan modelin bu tür kazaların belirlenmesinde başarılı bir şekilde kullanılabileceği ve kaza tutanaklarından kaza sebebinin belirlenme süresini çok kısaltabileceği sonucuna varılmıştır [9].

Soleimani vd. (2021) yaptıkları çalışmada trenler ve arabalar arasındaki kaza sayısını azaltmak amacıyla gereksiz geçişlerin kapatılmasına ve böylece çarpışma risklerinin azaltılmasına yardımcı olacak bir Karayolu-Demiryolu Hemzemin Geçidi Konsolidasyon Modeli oluşturmuşlardır. Önceki çalışmalarından en iyi performansı gösteren XGboost makine öğrenimi algoritmasına ek olarak metin madenciliği teknikleri ve Jeo-uzamsal (Geo-spatial) analiz kullanılmıştır. Sonuçlar, önerilen model için %88'lik bir genel doğruluk göstermiştir. Çalışmada modelde yer alan değişkenlerin göreceli önemleri de rapor edilmiştir ve böylece modelin davranışının derinlemesine anlaşılması sağlanmıştır. Farklı eşik değerlere göre yapılan yorumlar sonucunda incelenen bölgede mevcut karayolu-demiryolu hemzemin geçitlerinin %15'inin kapatılmasının gerektiği sonucuna varılmıştır [10].

Nilashi vd. (2021) yaptıkları çalışmada müşterilerin sosyal medyada restoranların hizmet kaliteleri hakkındaki görüşlerini inceleyerek müşteri segmentasyonunu sağlayan ve müşteri tercihlerini tahmin eden yeni bir yöntem geliştirmişlerdir. Bu amaçla çalışmada metin madenciliği yöntemleri (Latent Dirichlet Allocation), kümeleme analizi (Self Organizing Map) aşamalarından sonra CART (Classification and Regression Trees) yöntemi ile vejetaryen dostu restoranlarda müşterinin hizmet

kalitesi hakkındaki görüşlerinden memnuniyet düzeyleri tahmin edilmiştir. CART yöntemi ile elde edilen modellerin restoranların kalite faktörlerine dayanarak müşterilerin tercihlerini iyi bir şekilde yansıtabildiğini göstermiştir. Toplanan büyük miktardaki çevrimiçi müşteri yorumlarının makine öğrenmesi yöntemleri ile birlikte kullanılmasının restoranlardaki hizmet kalitesinin iyileştirilebilmesi ve müşteri segmentasyonunun sağlanması için etkili bir yöntem olduğu belirtilmiştir [11].

Petropoulos ve Siakoulis (2021) Merkez Bankalarının yaptıkları açıklamaları inceleyerek açıklamalardaki en önemli sinyalleri filtrelemek ve gelecekteki finansal piyasa davranışını tahmin etmek için bir duyarlılık endeksi oluşturmuşlardır. Bu amaçla makine öğrenimi teknikleriyle (XGboost) birlikte doğal dil işleme tekniklerini kullanmışlardır. Daha önceden tanımlanmış veya açıklamaların yer aldığı korpustan elde edilen sözlüklerin XGBoost ile birlikte kullanılması sonucu elde edilen analizlerin sonucunda piyasalardaki çalkantıları başarılı bir şekilde öngerebilmeyi sağlayan bir duyarlılık endeksi oluşturmanın mümkün olduğu belirtilmiştir [12].

Chatterjee vd. (2021) yaptıkları çalışmalarda e-ticaret sektöründe sağlık ürünleri sunan şirketlerin müşteri memnuniyetini modellemeyi ve sektördeki alt gruplar arasındaki farklılıklarını analiz etmeyi hedeflemişlerdir. Bu amaçla, 2008 ve 2018 yılları arasında bir inceleme sitesinde yayınlanan 29 sağlık/sağlık ürünleri alt kategorisinden 619 e-ticaret firması hakkında 186.057 yorum üzerinde metin madenciliği, makine öğrenimi (Doğrusal Regresyon, XGboost, Random Forest, ve CART) ve ekonometri teknikleri kullanılmıştır. Böylece müşteri memnuniyetini belirleyen hizmetlerin temelde neler olduğu ve hangi duyguların ön plana çıktığı belirlenerek sağlık hizmetleri sunan e-ticaret platformlarının tasarımında ve sunumundan nelere dikkat edilmesi ortaya konmuştur [13].

Lin vd. (2022) çalışmalarında hisse fiyatlarının metin tabanlı veriler kullanılarak tahmin edilmesinde verilerin farklı finansal kaynaklardan alınmasının, farklı metin tabanlı öznitelik temsillerinin (TF-IDF, word embeddings gibi) kullanılmasının ya da farklı makine öğrenmesi yöntemlerinin çalıştırılmasının model performanslarını etkileyip etkilemediklerini analiz etmişlerdir. Bu amaçla Reuters, CNBC, The Motley Fool gibi farklı kaynaklardan alınan finansal veriler üzerinde TF-IDF, Word2vec, ELMo, BERT gibi farklı öznitelik temsillerini (feature representation) ve SVM, CNN, LSTM gibi farklı makine öğrenmesi yöntemlerini test etmişlerdir. Vardıkları sonuç en iyi tahmin

performanslarının CNN+Word2vec ve CNN+BERT kombinasyonları ile elde edildiğidir [14].

Allenbrand (2024) yaptığı çalışmada hastalıkların tedavisinde kişiselleştirilmiş ve optimize edilmiş bir tedavi programı oluşturabilmek için metin madenciliği ve makine öğrenmesinden faydalanmıştır. Hastaların aldıkları tedavinin faydaları ve yan etkileri hakkında beyan ettikleri görüşler yoluyla memnuniyeti tahmin edecek modeller oluşturulmuştur. Bu amaçla Naive Bayes, Support Vector Machines, Random Forests ve CART yöntemleri uygulanmıştır. Model performansını ölçmek amacıyla Matthews korelasyon katsayısı ve F1 ölçütü kullanılmıştır. En iyi sonuçları Random Forest yöntemi vermiştir. Önleme adımında verideki gürültüyü azaltmak ve böylece tahmin modellerinin performansını arttırmak amacıyla kümeleme analizi de çalışmada yer almıştır [15].

Anagün vd. (2022) yaptıkları çalışmada finans kuruluşları için müşteri şikayetlerini otomatik olarak sınıflandıracak derin öğrenme tabanlı müşteri şikayet yönetim sistemi geliştirmişlerdir. Bu amaçla metin tabanlı veriler üzerinde modelin performansını arttırmak için yeni bir ön işleme tekniği tanımlanmıştır. Bu yöntemle %96 oranında bir başarı elde edilmiştir [16].

Işık vd. (2020) yaptıkları çalışmada spam ve spam olmayan epostaları sınıflandırmak için iki farklı öznelik seçimi (feature selection) yöntemi uygulanmıştır. Bu amaçla eposta verileri analize uygun olacak şekilde yapılandırılmış ve üzerinde öznelik seçimi yöntemi uygulanmıştır. Tahmin için üç farklı öğrenme yöntemi kullanılmıştır. Elde edilen sonuçlar Türkçe gibi eklemeli bir dil içeren metinsel veriler üzerinde belirtilen öznelik seçimi ve derin öğrenme yöntemlerinin birlikte kullanımının başarı oranını arttırdığı görülmüştür [17].

Yapılan çalışmalar incelendiğinde çalışmanın konusuna bağlı olarak (kazalar, hastalıklar, krizler, ağ saldırısı, sahtekarlık, alışveriş yapma, spam eposta vb) genellikle sınıflar arasında dengesizlikler bulunduğu görülmüştür. Makine öğrenmesi yöntemleri genellikle sınıf dağılımlarının benzer olduklarını varsayar. Yöntemlerin bazıları bu duruma karşı dirençliken bazıları daha az dirençlidir.

Aşağıda inşaat yönetimi ve pazarlaması alanında yapılan metin analitiği çalışmalarının gözden geçirildiği kapsamlı çalışmalar yer almaktadır. Baek vd. (2021) derlemelerinde inşaat sektörüne odaklanan metin madenciliği çalışmalarının mevcut durumunu ve gelecekteki olası trendleri, karşılaşılan zorlukları, yeni araştırma alanlarını ve kullanılan farklı veri kaynaklarını

gözden geçirmek için kapsamlı bir inceleme yapılmıştır. Bu amaçla inşaat sektörü ile ilgili 103 akademik makale incelenmiştir. Alandaki analizlerde kullanılan metin veri kaynakları, farklı metin madenciliği yaklaşımları ve makine öğrenimi yöntemleri gözden geçirilmiştir. Yazarların vardıkları sonuç, bu alanda metin analitiği yöntem ve tekniklerinin anlamlı bilgi üretmek için yeterince gelişmiş olduğu ve metin madenciliği ve makine öğreniminin birlikte kullanımının ilerisi için daha yeni fırsatlar yaratabileceği yönündedir. Ayrıca, yakın gelecekte doğal dil işleme yöntemlerinin gelişmesi, kullanılan araçların yaygınlaşması ve dijital dönüşümün daha da hızlanması ile beraber emek yoğun metin tabanlı görevlerin yerini otomatik metin analizi araçlarının alacağı düşünülmektedir [18].

Yan vd. (2022) inşaat sektöründe metin madenciliği uygulamalarına ilişkin 2000-2021 yılları arasında yayınlanan 127 akademik dergi makalesi üzerinde hacimli bir araştırma gerçekleştirmiştir. Yayınlarda oluşan trendler, yayınların en çok yapıldığı ülkeler ve bölgeler, öne çıkan araştırmacılar ve anahtar kelimeler, hangi anahtar kelimelerin daha çok birlikte kullanıldığı gibi ayrıntıları belirlemek amacıyla VOSviewer yazılımından faydalanılmıştır. Çalışmada metin madenciliğinin öncelikli uygulama alanları detaylı olarak incelenmiştir. Ayrıca karşılaşılan temel zorluklar ve gelecekteki yönelimler (alan bilgisinin veya ontolojinin modellere entegrasyonu, yenilikçi teknolojilerin uygulanması ve sosyal medyadaki metinsel veri üzerinde duygu analizi yapılması) çalışmada yer almaktadır [19].

Shamshiri vd. 2024 tarafından inşaat sektöründe metin madenciliği ve doğal dil işleme yöntemlerinin derinlemesine incelenmesi için 205 akademik makale ele alınmıştır. Bu makaleler öncelikle uygulama alanlarına göre sınıflandırılarak incelenmiştir. Derleme çalışmasında metin madenciliği ve doğal dil işleme yöntemlerinin inşaat yönetiminin kısıt yönetimi, kapsam yönetimi, entegrasyon yönetimi, kaynak yönetimi, iletişim yönetimi, malzeme yönetimi, tedarik yönetimi, devreye alma ve başlatma ve proje kontrolü gibi alt alanlarındaki potansiyelleri hakkında yorumlar içermektedir. Ayrıca, çalışmada mevcut çalışmalardaki boşlukları tespit ederek gelecekteki inşaat sektöründeki uygulamaların daha az insana bağımlı ve daha az hataya eğilimli hale getirilmesi için öneriler bulunmaktadır [20].

Yapılan literatür araştırmasında metin madenciliğinin farklı amaçlarla birçok uygulama kullanıldığı görülmektedir. Fakat bildiğimiz kadarıyla bu

çalışmadaki gibi bir inşaat şirketinin potansiyel müşterileri ile yaptığı görüşmeler sonucu elde edilen metinsel veriler üzerinde yeni bir müşterinin ev alma niyetini farklı makine öğrenmesi yöntemleri ile tahmin eden bir çalışma bulunmamaktadır.

### 3. MATERYAL VE METODLAR (MATERIAL AND METHODS)

#### 3.1 Veri Kümesi

Veriler, dijital mecralarda ve çeşitli gazetelerin internet sayfalarında reklam gösterimi yapan bir inşaat firmasından elde edilmiştir. Firma yaptığı reklamlar sayesinde internet kullanıcılarının kendi sitesine gelecek konut projeleri hakkında bilgi edinmelerini sağlamıştır. Daha detaylı bilgi almak isteyen kullanıcılar ise iletişim bilgilerini sitedeki bir form aracılığıyla firmaya göndermişlerdir. Kullanıcıların iletişim bilgilerini alan firma kullanıcılarla telefon görüşmesi veya yüz yüze görüşmeler yapıp onların görüşlerini yazılı olarak elektronik ortamda kayıt altına almıştır. Analiz için elde edilen metin verisi bu şekilde toplanmıştır. Yapılan görüşmelerin sonucunda kullanıcıların konut satın alıp almadığı bilgisi de veri setinde yer almaktadır.

Oluşturulan bu veri seti üzerinde farklı makine öğrenmesi yöntemleri kullanılarak yeni bir kullanıcı ile yapılan görüşmenin sonucunun olumlu (konut alma) veya olumsuz (konut almama) olmak üzere iki seçenekten hangisi olacağını tahmin edecek modeller oluşturulmuştur. İkili sınıflandırma problemi şeklinde ele alınan uygulama yüksek performans gösterecek şekilde çözülmeye çalışılmıştır.

Yapılandırılmamış metin tabanlı veri kümesi R [21] dilinde metin madenciliği için en yaygın olarak kullanılan paket olan Text Mining (TM) [4] kullanılarak vektör uzayı modeli ile temsil edilen yapılandırılmış hale getirilmiştir. Bu paketin içerisinde tüm metin ön işleme aşamaları için fonksiyonlar tanımlanmıştır.

Web sayfasına 1492 kullanıcı iletişim bilgilerini bırakmıştır. İletişim bilgileri doğru olan ve yüz yüze veya telefonla karşılıklı görüşme yapılan kişi sayısı ise 579 dur. Metin tabanlı veri seti bu kişilerle yapılan görüşmeler sonucu oluşturulmuştur. Dolayısıyla korpus her bir görüşmeye ait 579 metin dosyasından oluşmaktadır.

Her bir dosya üzerinde tüm karakterlerin küçük harfe dönüştürülmesi, noktalama işaretlerinin kaldırılması, cümlelerin kelimelere ayrılması (tokenization), bir anlam ifade etmeyen durma kelimelerinin (stop words) kaldırılması, kelimelerin köklerinin bulunması

(stemming), aynı anlama gelen kelimelerin tek bir kelime altında birleştirilmesi gibi ön işleme adımları yerine getirilmiştir. Veriler metin ön işleme aşamaları yapıldıktan sonra vektör uzay modeline dönüştürülmüştür. Vektör uzay modeliyle yapısal olmayan bir biçimden yapısal biçime dönüştürülen veriler analize hazır hale getirilmiştir. Veri ön işleme aşaması birden fazla adımdan oluşması ve karşılaşılan çeşitli zorluklar nedeniyle uygulamanın en çok zaman harcanan aşaması olmuştur.

Örnek veri kümesi Tablo 4’de verilmiştir. Çıktı değerinin “1” olması görüşme sonucu satış yapıldığı, “0” satış işleminin yapılmadığını göstermektedir.

Tablo 1: Veri kümesi özet tablo

Görüşme No	lokasyon	daire	metrekare	...	Sınıf Değişkeni
1	1	1	1	...	1
2	0	1	0	...	0
3	0	1	0	...	0
4	0	0	1	...	1
5	0	1	1	...	0
⋮	⋮	⋮	⋮	⋮	⋮
579	0	0	0	...	0

Tablo 1’de gösterilen veri kümesi sınıf dağılımında “0” ile temsil edilen kategoriye ait 534, “1” ile temsil edilen kategoriye ait 45 gözlem bulunmaktadır. Bunun anlamı görüşmelerin %8’inin satın almayla sonuçlanırken, geri kalan %92 sinde herhangi bir satış işleminin olmadığını göstermektedir. Bu konut satışı gibi bir alanda beklenen bir sonuçtur. Fakat makine öğrenmesi açısından sınıfların dengesiz olması model performansını etkileyen faktörlerden birisidir.

Tablo 2: Anahtar kelime frekansları

Anahtar Kelimeler	Frekanslar
ofis	392
proje	332
daire	297
bilgilendirme	243
davet	208
fiyat	199
avantaj	159
lansman	126
satış	123
1+1	122

Tablo 2’de korpusta en çok yer alan ilk 10 kelime ve frekans sayıları verilmiştir.

Tablo 3: Veri kümesi sınıf dağılımı

Y-Sınıf değişkeni (Bağımlı değişken)		
0-sınıfı gözlem sayısı	1-sınıfı gözlem sayısı	Toplam Gözlem Sayısı
534 (92%)	45 (8%)	579 (100%)

Veri kümesi rastgele %80'i eğitim ve %20'i test verisi olarak ikiye ayrılmıştır. Eğitim kümesinde 465, test kümesinde ise 114 gözlem bulunmaktadır. Rastgele seçim yapıldığından eğitim ve test kümelerindeki sınıf dağılımları benzer olmuştur.

Sınıflandırma tabloları oluşturulurken kesim noktası satın alma gerçekleştirenlerin sınıfının yaygınlık değeri olan %8 olarak alınmıştır. Sınıflandırma analizi sonucu tahmin edilen değeri 0.08 değerinin üzerinde olan gözlemler "1" ile temsil edilen satın alma gerçekleştirmiş olanlar sınıfına, 0.08 değerine eşit ve altında olan gözlemler ise "0" ile temsil edilen satın alma gerçekleştirmeyenlerin olduğu sınıfa atanmıştır.

### 3.2 Metin Madenciliği

Metin sınıflandırma probleminin amacı, metin verisini daha önceden belli olan kategorilere otomatik olarak atanmasıdır. Amaç, makine öğrenmesi algoritmalarını kullanarak, metinleri kategorilere (sınıflara) otomatik olarak atayacak sınıflandırıcıları oluşturmaktır. Veri madenciliğinde bu durum denetimli öğrenme olarak da bilinir. Dokümanların (metinlerin) otomatik olarak sınıflandırılabilmesi için vektörel olarak ifade edilmesi gerekir. Bunun için de "vektör uzay modeli" oluşturulmalıdır [1].

Metin sınıflandırma eğitim ve test aşaması olarak iki aşamadan oluşmaktadır. Eğitim aşamasına geçmeden önce belgeler vektör uzayı modeline dönüştürülür ve sınıflandırma algoritmaları kullanılarak model oluşturulur. Test aşamasında ise eğitim kümesinde olmayan veriler oluşturulan model yardımıyla sınıflandırılır ve model performansı değerlendirilir.

Metin verisinden bilgi çıkarımına giden süreçte genellikle takip edilen adımlar Şekil 1'de görülmektedir.



Şekil 1. Metin Sınıflandırma Süreci

Metin ön işleme, çoğu metin analizinde gerekli olan ve metin temsilini kolaylaştırmak için girdi belgelerini daha tutarlı hale getirmenin amaçlandığı aşamadır. Geleneksel metin ön işleme yöntemleri arasında

durdurucu kelimelerden (stop-words) arındırma ve gövdeleme (stemming) yani kelime köklerini bulma işlemleri yapılır. Durdurucu kelimeleri kaldırma aşamasında sözcüklerin daha genel ve anlamsız olarak değerlendirildiği bir sözcük durdurma listesi kullanılarak (acaba, ama, ancak, için, madem vb.) kelimeler kaldırılır. Gövdelemede, kelimeler köklerine indirgemektedir. Örneğin, "izlemek", "izliyor", ve "izledi", ek almış kelimeleri "izlemek" olarak temsil edilir [1]. Ayrıca bu aşamada hatalı yazımları düzeltme, noktalama işaretleri ve gereksiz kelimeleri (edatlar, bağlaçlar) çıkarma işlemleri de yapılır.  $d_i$  tekil bir doküman olmak üzere,  $D = \{d_1, d_2, \dots, d_n\}$  ile ifade edilen dokümanları bir arada bulunduran D veri seti "korpus" olarak adlandırılmaktadır. Korpus, üzerinde metin analizleri yapılacak veri setidir. Korpus nesnesi, metin madenciliğinde analizlerde kullanılan doküman terim matrisine (vektör uzay modeline) çevrilmeye hazır, tüm metinlerin bazı özellikleri ile birlikte tutulduğu matris benzeri bir yapıdır [22]. Ön işleme çalışmaları tamamlandığında çalışma matrisinin hazırlanmasına geçilir. Oluşan sözlüğe göre dokümanlar sayısal olarak ifade edilir. Vektör uzay modeli metinlerin sayısal hale (yapılandırılmış veri) getirilmesini sağlayan bir yöntemdir.

Bir metin kelime dizisi olarak ifade edilebilir. Bir eğitim setinin tüm kelimelerine sözlük veya özellik seti denir. Böylelikle, bir belge ikili vektörle gösterilebilir, eğer belge sözlükteki kelimeyi içeriyorsa 1 değerini, içermiyorsa 0 değerini alır. Bu, bir belgeyi  $R^{|v|}$  alanına yerleştirmek olarak düşünülebilir. Burada,  $|v|$  sözlükte yer alan kelime sayısını gösterir. Oluşturulan sözlüğe göre dokümanlar sayısal olarak ifade edilir. Kelimelerin bu şekilde gösterilmesi vektör uzayı modeli olarak adlandırılır [23].

$$x_i(\text{Kelime}) = \begin{cases} 1, & i. \text{ kelime doküman içinde yer alıyorsa} \\ 0, & i. \text{ kelime doküman içinde yer almıyorsa} \end{cases}$$

Tablo 4: Vektör Uzay Modeli –Doküman Terim Matrisi

	Kelime 1	Kelime 2	...	Kelime n
Doküman 1	1	0	...	1
Doküman 2	0	1	...	0
Doküman 3	0	1	...	0
...	...	...	...	...
Doküman n	0	0		1

### 3.3 CART (Classification and Regression Tree) Yöntemi

CART (Sınıflandırma ve Regresyon Ağaçları) her bir ara düğümün (internal node) bir değişkeni (öznitelik, feature) ve her yaprak düğümün (leaf node) bir sınıf etiketi taşıdığı akış şeması benzeri bir yapıya sahiptir. Ağaçta bulunan en üstteki düğüm, kök düğümdür (root node) ve bu düğüm de veri kümesindeki bir değişkeni temsil etmektedir [24]. Karar ağaçlarındaki kök düğüm ve ara düğümler entropi (entropy), gini endeksi (gini index), kazanç oranı (gain ratio), bilgi kazancı (information gain), twoing gibi bölünme kriteri (değişken seçim) algoritmaları ile oluşturulmaktadır. Bu kriterlerden hangisinin kullanılacağına göre karar ağacı algoritması farklılık göstermektedir.

CART algoritması karar ağaçlarının genelinde olduğu gibi yukarıdan aşağıya iteratif olarak böl ve yönet tarzında oluşan açgözlü (greedy) (geri izlemesiz) bir yaklaşım benimser. CART algoritmasına bölünme yöntemi olarak genellikle gini endeksi kullanılır. Karar ağaçları için çoğu algoritmada, eğitim kümesi iteratif olarak daha küçük alt kümelerle bölünür [24]. Her bölünmede ağaç üzerinde aşağıya doğru iki yeni dal oluşur. Ağaç oluşturma sürecinin her aşamasında, en iyi bölünme, ileriye bakmak ve gelecekteki bazı adımlarda daha iyi bir ağaca yol açacak bir bölünme seçmek yerine, o aşamada yapılır [25].

### 3.4 Random Forest Yöntemi

Random Forest (Rastgele Orman), karar ağacı tabanlı sınıflandırıcılar topluluğudur, ormandaki her ağaç bağımsız olarak örneklenen bir rasgele vektörün değerlerine bağlı olarak oluşur. Bu yöntemi Breiman makalesinde şöyle açıklamaktadır.  $k$ 'inci ağaç için, geçmiş rastgele vektörlerden  $Q_1, \dots, Q_{k-1}$  bağımsız ancak aynı dağılımla sahip  $Q_k$  rasgele vektörü oluşturulur; ve bir ağaç,  $x$ 'in girdi vektörü olduğu eğitim veri seti ve  $Q_k$  kullanılarak  $h(x, Q_k)$  sınıflandırıcısı oluşturulur. Rastgele bölünme seçiminde  $Q$ ,  $1$  ile  $k$  arasında bağımsız rastgele tam sayıdan oluşur.  $Q$ 'nun yapısı ve boyutu, ağaç yapımındaki kullanımına bağlıdır. Rastgele orman yönteminde, torbalama ve rastgele değişken (özellik) seçimi birlikte kullanılır. Her yeni eğitim seti, orijinal eğitim setinden iadeli olarak (bootstrap yöntemiyle) çekilir. Ardından rastgele değişken (özellik) seçimi kullanılarak yeni eğitim setinde bir ağaç yetiştirilir. Yetiştirilen ağaçlarda budama yapılmaz [26]

Bu algoritmada her sınıflandırıcı eğitim kümesine eşit boyutta iadeli örnekleme yolu ile elde edilen veri kümeleri ile eğitilir. İadeli örnekleme kullanıldığı için

eğitim kümesindeki gözlemler elde edilen veri kümelerinde birden fazlada görülebilir ya da hiç yer almayabilir. Yeni bir örneği sınıflandırmak için her sınıflandırıcı kendi sınıf tahminini oluşturur ve torbalanmış sınıflandırıcı (bagged classifier) en fazla tahmin edilen sınıfı sonuç olarak döndürür (oylama yöntemi - voting method) [27].

### 3.5 XGBoost (Extreme Gradient Boosting) Yöntemi

Gradyan artırma karar ağacı (Gradient Boosting Decision Tree) sınıflandırma doğruluğu ve etkinliği nedeniyle yaygın olarak kullanılan bir makine öğrenme algoritmasıdır. GBDT sırayla eğitilmiş karar ağaçlarının bir topluluk modelidir [28]. GBDT zayıf modeller olarak genellikle regresyon ağaçlarını kullanır [29]. Her yinelemede, negatif gradyanları (hatalar olarak da adlandırılır) kullanarak karar ağaçları oluşturur [30].

XGBoost gradyan artırma algoritmasının ek parametre ayarları ile gradyan artırmanın genelleştirilmiş hali olarak düşünülebilir. Tahmin performansı yüksek olmakla birlikte çok çekirdekli ve paralel (dağıtılmış) makine uygulamasına da sahiptir [31]. XGBoost algoritması yenilik olarak; seyrek veriler (sparse data) için yeni bir ağaç öğrenme algoritması sunar, ağırlıklı nicel çizim (weighted quantile sketch) ile ağaç öğreniminde örnek ağırlıklarını belirler ve paralel bilgi işleme ile öğrenmeyi daha hızlı hale getirir [32].

XGBoost, türevlenebilir kayıp fonksiyonuna ek olarak aşırı öğrenmeyi önlemek için düzenleme (regularization) terimi kullanır.  $\{(x_i, y_i)\}_{i=1}^n$  veri kümesi ve  $l(y_i, \hat{y}_i)$  diferansiyallenebilir kayıp fonksiyon (loss function) olmak üzere;

$$L(\varphi) = \sum_i l(y_i, \hat{y}_i) + \sum_k \Omega(f_k) \quad (1)$$

$$\Omega(f_k) = \gamma T + \frac{1}{2} \lambda w^2$$

olacak şekilde Denklem 1 elde edilir.  $\Omega(f_k)$ ,  $f_k$  modelinin (karar ağacı) karmaşıklığı olarak adlandırılır ve bu terim aşırı öğrenmeyi engellemektedir. Bu terimin sıfır olması gradyan artırma algoritmasına karşılık gelmektedir [32].  $\Omega(f_k)$  fonksiyonunda  $T$ ,  $f_k$  ağacındaki yaprak sayısı  $w$  yaprak ağırlığıdır ve yapraklardaki tahmin edilen değerler yardımıyla hesaplanmaktadır.  $\gamma T$  her ek ağaç yaprağı için sabit bir ceza puanı oluşturur ve  $\lambda w^2$  aşırı ağırlıkları cezalandırır.  $\gamma$  ve  $\lambda$  uygulayıcı tarafından belirlenen parametrelerdir [31].

### 3.6 Model Değerlendirme

Eğitilen modelin performansı bağımsız test verileri üzerindeki tahmin başarısına bağlıdır [33]. Model değerlendirme, gözlemlerin sınıf etiketini tahmin etmede ne kadar iyi olduğunu değerlendirmek için ölçümler sunmaktadır [24]. Sınıflandırma modellerini değerlendirme de genellikle hata matrisi (confusion matrix) kullanılır. İki sınıflı modellerin hata matrisi 2 satır ve 2 sütundan oluşmaktadır. Bir sınıflandırma modelinin iyi bir doğruluğa sahip olması için gözlemlerin çoğunun asal köşegen üzerinde yer alması ve asal köşegen dışında kalan yerlerinde sifira yakın olması istenir [24]. İki sınıflı bir sınıflandırma problemi için oluşturulan hata matrisi Tablo 5’de gösterilmektedir. Gözlemler pozitif ve negatif olarak etiketlenmiştir.  $P$  veri kümesinde yer alan pozitif gözlem sayısını,  $N$  ise negatif gözlem sayısını temsil etmekte olup  $P'$  ve  $N'$  ise model tarafından tahmin edilen pozitif ve negatif gözlem sayısını temsil etmektedir [24].

Tablo 5: Sınıflandırma tablosu- Hata Matrisi

		Gerçek Sınıf (Actual Class)		
		Hayır(0)	Evet(1)	Toplam
Tahmin Edilen Sınıf (Predicted Class)	Hayır(0)	$DN$	$YN$	$N'$
	Evet (1)	$YP$	$DP$	$P'$
	Toplam	$N$	$P$	$N+P$

**Doğru Pozitif (DP):** Sınıflandırıcı tarafından doğru şekilde tahmin edilmiş pozitif gözlemleri ifade etmektedir.  $DP$  doğru pozitif gözlemlerin sayısına karşılık gelmektedir.

**Doğru Negatif [34]:** Sınıflandırıcı tarafından doğru şekilde tahmin edilmiş negatif gözlemleri ifade etmektedir.  $DN$  doğru negatif gözlemlerin sayısına karşılık gelmektedir.

**Yanlış Pozitif (YP):** Gerçekte negatif olan ancak sınıflandırıcı tarafından pozitif olarak tahmin edilen gözlemleri ifade etmektedir.  $YP$  yanlış pozitif gözlemlerin sayısına karşılık gelmektedir.

**Yanlış Negatif (YN):** Gerçekte pozitif olan ancak sınıflandırıcı tarafından negatif olarak tahmin edilen gözlemleri ifade etmektedir.  $YN$  yanlış negatif gözlemlerin sayısına karşılık gelmektedir.

Sınıflandırma modellerini birbiriyle karşılaştırabilmek için hata matrisinden birçok farklı ölçüt oluşturulmuştur. Tablo 6’da en yaygın kullanılan performans ölçütlerinden bazıları ve bunların hesaplanma yolları verilmiştir.

Tablo 6: Değerlendirme Ölçütleri

Değerlendirme Ölçütleri	Formüller
Doğruluk (Accuracy)	$\frac{DP + DN}{P + N}$
Yanlış Sınıflandırma Oranı (Misclassification rate)	$\frac{YN + YP}{P + N}$
Duyarlılık (Sensitivity) Hatırlama (Recall) Doğru Pozitif Oran (True Positive Rate)	$\frac{DP}{P}$
Özgüllük (Specificity) Doğru Negatif Oran (True Negative Rate)	$\frac{DN}{N}$
Yanlış Negatif Oran (False Negative Rate)	$\frac{YN}{P}$
Yanlış Pozitif Oran (False Positive Rate)	$\frac{YP}{N}$
Keskinlik (Precision)	$\frac{DP}{P'}$
Yaygınlık (Prevalence)	$\frac{P}{P + N}$

Sınıf dengesizliğinin olduğu problemlerde sınıflandırma modeli, çoğunluk sınıfının gözlemlerini doğru bir şekilde sınıflandırır. Ancak bu durumda azınlık sınıf gözlemleri de yanlış sınıflandırılabilir. Bu nedenle duyarlılık (sensivity) ve özgüllük (specificity) ölçülerini kullanmak daha tutarlı olmaktadır. Duyarlılık, doğru pozitif oran olarak da adlandırılır ve gerçekte pozitif olan gözlemlerin ne kadarının model tarafından doğru bir şekilde sınıflandırıldığı ölçüsünü vermektedir. Özgüllük ise doğru negatif oran olarak adlandırılır ve gerçekte negatif olan gözlemlerin model tarafından ne kadarının doğru bir şekilde sınıflandırıldığı ölçüsünü vermektedir [24]. Keskinlik (precision) ölçüsü ise model tarafından pozitif olarak etiketlenen gözlemlerin ne kadarının doğru bir şekilde sınıflandırıldığı oranını vermektedir [35].

Tablo 6’da hesaplanan tüm doğruluk ölçümleri eşik değer seçimine bağlı olarak hesaplanmaktadır. Eşik değer seçimi sınıflandırma hatalarında değişikliğe neden olabilmektedir. Bu durum dengesiz veri kümeleri için doğru bir yaklaşım değildir [36]. Dengesiz veri kümelerinde yüksek veya düşük düzeyde gözlemlenen yaygınlık (prevelence) bulunmaktadır. Yaygınlık, veri kümesinde her kategorinin ne kadar sıklıkta olduğunu veren ölçüdür [36]. Gözlenen (gerçek) ve tahmin

yaygınlığı olarak ikiye ayrılmaktadır. Birçok uygulamada, gözlenen ve tahmin edilen yaygınlığın benzerlik göstermesi önemlidir. Dolayısıyla hem tahmin edilen hem de gözlemlenen yaygınlık eşik değeri için bir kriter olarak kullanılabilir [37]. Eşik değeri seçimi ile birlikte sınıf tahminleri oluşur. Eşik değeri seçimine bağlı olmayan yöntemlerden olan ROC (Receiver Operating Characteristics) ve kesinlik-doğru pozitif oran (Precision-Recall; PR) eğrileri yaklaşımı model performansını değerlendirmek için kullanılan daha objektif değerlendirme ölçüleridir [35].

İki sınıflı bir problemde, ROC eğrisi, modelin pozitif gözlemleri doğru bir şekilde sınıflandırdığı oran (duyarlılık) ile negatif gözlemleri yanlışlıkla pozitif olarak sınıflandırdığı oran arasındaki dengeyi görselleştirmemizi sağlar. ROC eğrisinin altındaki alan (Area Under Curve – AUC) ise modelin doğruluğunun bir ölçüsüdür [24]. ROC eğrisinin altındaki alan (Area Under Curve -  $AUC_{ROC}$ ), genel model performansı için bir ölçü vermektedir. İyi modellerin AUC değeri 1'e yakinken, zayıf modellerin AUC değeri 0.5'e yakındır. Bu değerlere bakılarak model performansları değerlendirilir [36].

İkili sınıflama modellerinde sınıf dağılımında dengesizlik olması durumunda ROC eğrilerine alternatif olarak PR eğrileri önerilmiştir [35, 38]. ROC alanı ve PR alanı arasındaki önemli fark eğrilerin görselliğidir. PR eğrilerinde, ROC uzayında belirgin olmayan algoritmalar arasındaki farklılıklar ortaya çıkabilmektedir. PR alanında dikey ekseninde kesinlik (precision) ve yatay ekseninde duyarlılık (doğru pozitif oran, recall-sensitivity) bulunmaktadır [39]. ROC eğrisinde olduğu gibi PR eğrisinde de eğri altında kalan alan (Area Under Curve-  $AUC_{PR}$ ) kullanılabilir ancak  $AUC_{PR}$  pozitif sınıfın yaygınlığına göre değişmektedir ve beklenen değeri veri kümesindeki pozitif sınıf oranına yakındır.  $AUC_{PR}$  değeri için alt sınır pozitif sınıfın yaygınlık değeridir.  $AUC_{PR}$  değeri yaygınlık değerinden ne kadar büyükse sınıflandırıcının o oranda iyi olduğu söylenir. [35].  $AUC_{ROC}$  değerinde olduğu gibi iyi modellerin  $AUC_{PR}$  değerinin de 1'e yakın olması beklenir.

#### 4. UYGULAMA (APPLICATION)

##### 4.1 CART Sonuçlar

CART algoritmasında bölünme kriteri olarak gini endeksi kullanılmıştır. CART analizi rpart [40] paketi yardımıyla yapılmıştır. Terminal düğümlerde 2 veya fazla gözlem olduğu sürece bölünme işlemi devam ettirilmiştir. Bulunan bu değer denemeler sonucu test kümesindeki AUC değerini maksimum yapan değerdir

ve bir parametre (minimum gözlem sayısı) olarak CART yönteminde kullanılmıştır. CART yöntemi kullanılarak elde edilen sınıflandırma tablosu eğitim ve test performansları sırasıyla Tablo 7 ve Tablo 8'de verilmiştir.

Tablo 7: CART (Eğitim kümesi performansı)

Predicted (Tahmin)	Actual (Gerçek)	
	0	1
0	363	5
1	70	27

Tablo 8: CART (Test kümesi performansı)

Predicted (Tahmin)	Actual (Gerçek)	
	0	1
0	88	8
1	13	5

Konut satışında önemli olan gerçekten satış yapılabilecek müşterilerin belirlenmesidir. Bu bağlamda CART yöntemi test kümesinde satış yapılan 13 müşteriden 5 tanesini doğru tahmin edebilmiştir. Bu değerler kesim noktasına bağlı olarak değişmektedir.

Tablo 9: CART (Farklı ölçütlere göre eğitim ve test kümesi performansları)

	Duyarlılık	Özgüllük	Kesinlik	Hata	Doğruluk
Eğitim	0.84	0.84	0.28	0.16	0.84
Test	0.38	0.87	0.27	0.18	0.82

Farklı sınıflandırma metriklerinin sonuçları eğitim ve test kümeleri için Tablo 9'da verilmiştir. CART yöntemi eğitim kümesindeki duyarlılığı %84 iken test kümesinde bu oran %38 olmuştur. Test doğruluk değeri ise 0.92 olarak bulunmuştur. Duyarlılık dışında diğer ölçütler eğitim ve test kümelerinde birbirine yakın değerler almıştır. Bu tablodaki hesaplanan tüm ölçütler, kesim noktasının alışveriş gerçekleştirenler sınıfının yaygınlık değeri olan %8 alınmasıyla elde edilmiştir. Kesim noktasının değişmesi durumunda hesaplanan değerler de değişiklik gösterecektir.

Tablo 10: CART (ROC-AUC ve PR-AUC değerleri)

	ROC-AUC	PR-AUC
Eğitim	0.91	0.71



<b>Test</b>	0.63	0.24
-------------	------	------

Tablo 10'da ise kesim noktasından bağımsız olarak hesaplanan ROC-AUC ve PR-AUC değerleri yer almaktadır. ROC-AUC ve PR-AUC değerleri PRROC [41, 42] paketi kullanılarak hesaplanmıştır. Bu değerlere bakıldığında eğitim ve test kümeleri arasında farklılıklar bulunmaktadır. Fakat bu tür dengesiz veri tiplerindeki performansı ölçmek için kullanılan PR-AUC değeri alışveriş gerçekleştirenler sınıfının yaygınlık değeri olan %8 değerinden daha büyük bulunmuştur. Bu durum da modelin sınıfları dengesiz dağılan bu veri kümesi üzerinde iyi bir performans gösterdiğinin bir ölçüsü olarak kabul edilir

#### 4.2 Random Forest Sonuçlar

Random Forest yöntemi için önceden tanımlı parametrelerle model çalıştırılmış ve 25. ağaçta (ntree=25) minimum OOB (Out of Bag) hatasına (0.071) ulaşılmıştır. Ntree parametresi RF algoritmasında yer alan ağaç sayısını göstermektedir. Sonrasında 25 ağaçlı model için değişken sayısı (73) kadar döngü ile modeller oluşturulmuş ve rastgele değişken sayısının 8 olması durumunda model minimum OOB hatasına sahip olduğu görülmüştür. Rastgele orman analizi için R programında randomForest kütüphanesi kullanılmıştır [43].

Bu parametrelere göre Random Forest yönteminden elde edilen eğitim ve test performansları sırasıyla Tablo 11 ve Tablo 12'de hata matrisi biçiminde verilmiştir.

Tablo 11: Random Forest (Eğitim kümesi performansı)

<b>Predicted (Tahmin)</b>	<b>Actual (Gerçek)</b>	
	<b>0</b>	<b>1</b>
<b>0</b>	410	9
<b>1</b>	23	23

Tablo 12: Random Forest (Test kümesi performansı)

<b>Predicted (Tahmin)</b>	<b>Actual (Gerçek)</b>	
	<b>0</b>	<b>1</b>
<b>0</b>	92	10
<b>1</b>	9	3

Random Forest yöntemi test kümesinde satış yapılan 13 müşteriden sadece 3 müşteriyi doğru olarak tahmin edebilmiştir. Bu değerler kesim noktasına bağlı olarak değişmektedir.

Tablo 13: Random Forest (Farklı ölçütlere göre eğitim ve test kümesi performansları)

	<b>Duyarlılık</b>	<b>Özgüllük</b>	<b>Kesinlik</b>	<b>Hata</b>	<b>Doğruluk</b>
<b>Eğitim</b>	0.72	0.95	0.50	0.07	0.93
<b>Test</b>	0.23	0.91	0.25	0.17	0.83

Farklı sınıflandırma metriklerinin sonuçları eğitim ve test kümeleri için Tablo 13'de verilmiştir. Random Forest yönteminin eğitim kümesindeki duyarlılığı %72 iken test kümesinde bu oran %23 olmuştur. CART yönteminden daha düşük bir duyarlılık değeri elde edilmiştir. Bu tablodaki hesaplanan tüm metrikler, kesim noktasının alışveriş gerçekleştirenler sınıfının yaygınlık değeri olan %8 alınmasıyla elde edilmiştir. Kesim noktasının değişmesi durumunda hesaplanan değerler de değişiklik gösterecektir.

Tablo 14: Random Forest (ROC-AUC ve PR-AUC değerleri)

	<b>ROC-AUC</b>	<b>PR-AUC</b>
<b>Eğitim</b>	0.90	0.72
<b>Test</b>	0.66	0.35

Tablo 14'de ise kesim noktasından bağımsız olarak hesaplanan ROC-AUC ve PR-AUC değerleri yer almaktadır. PR-AUC değeri alışveriş gerçekleştirenler sınıfının yaygınlık değeri olan %8 değerinden daha büyük bulunmuştur. Bu durum da modelin sınıfları dengesiz dağılan bu veri kümesi üzerinde yine de iyi bir performans gösterdiğinin bir ölçüsü olarak kabul edilir. PR-AUC değeri CART yönteminden daha iyi sonuç vermiştir

#### 4.3 XGBoost Sonuçlar

XGBoost algoritması için öncelikle modelde yer alan ağaç sayısı (iterasyon) belirlenmiştir. Ağaç sayısı belirlenirken minimum log-kayıp (log-loss) değeri dikkate alınmıştır. Bu değer eğitim ve test kümesi için ayrı ayrı belirlenmiştir. Eğitim ve test kümelerindeki log-kayıp oranının iterasyon (yineleme sayısı, XGBoost algoritmasında ağaç sayısına da karşılık gelmektedir) sayısına göre test kümesi için log-kayıp oranı minimum değerini (0.32) 208.iterayonda almıştır. Öğrenme oranı (learning rate) ise 0.01 olarak alınmıştır. Öğrenme oranının küçük olması aşırı öğrenmeyi engellemektedir. Model önce 1000 iterasyon olacak şekilde çalıştırılmış ve test kümesi için minimum log-kayıp değerini 208. iterasyonda almıştır. XGBoost analizi için R dilinde hazırlanmış xgboost [44] kütüphanesi kullanılmıştır.

Yukarıda belirtilen parametrelere göre XGBoost algoritmasının çalıştırılmasıyla elde edilen eğitim ve test performansları sırasıyla Tablo 15 ve Tablo 16'de sunulmuştur.

Tablo 15: XGBoost (Eğitim kümesi performansı)

Predicted (Tahmin)	Actual (Gerçek)	
	0	1
0	222	1
1	211	31

Tablo 16: XGBoost (Test kümesi performansı)

Predicted (Tahmin)	Actual (Gerçek)	
	0	1
0	51	4
1	50	9

XGBoost yöntemi test kümesinde satış yapılan 13 müşteriden 9 müşteriyi doğru olarak tahmin edebilmiştir. Bu da CART ve Random Forest yöntemine göre duyarlılık yönünden daha iyi sonuçlar verdiğini göstermektedir.

Tablo 17: XGBoost (Farklı ölçütlere göre eğitim ve test kümesi performansları)

	Duyarlılık	Özgüllük	Kesinlik	Hata	Doğruluk
Eğitim	0.96	0.50	0.12	0.47	0.53
Test	0.69	0.50	0.15	0.47	0.53

Farklı sınıflandırma metriklerinin sonuçları eğitim ve test kümeleri için Tablo 17'de verilmiştir. RF yöntemi eğitim kümesindeki duyarlılığı %96 iken test kümesinde bu oran %69 olmuştur. CART ve RF yönteminden daha yüksek bir duyarlılık değeri elde edilmiştir. Bu tablodaki hesaplanan tüm metrikler, kesim noktasının alışveriş gerçekleştirenler sınıfının yaygınlık değeri olan %8 alınmasıyla elde edilmiştir. Kesim noktasının değişmesi durumunda hesaplanan değerler de değişiklik gösterecektir.

Tablo 18: XGBoost (ROC-AUC ve PR-AUC değerleri)

	ROC-AUC	PR-AUC
Eğitim	0.92	0.68
Test	0.67	0.31

Tablo 18'de ise kesim noktasından bağımsız olarak hesaplanan ROC-AUC ve PR-AUC değerleri yer almaktadır. PR-AUC değeri alışveriş gerçekleştirenler sınıfının yaygınlık değeri olan %8 değerinden daha büyük bulunmuştur. Bu durum da modelin sınıfları dengesiz dağılan bu veri kümesi üzerinde yine de iyi bir performans gösterdiğinin bir ölçüsü olarak kabul edilir. Test veri kümesi için PR-AUC değeri CART yönteminden yüksek ancak RF algoritmasından düşük bulunmuştur.

#### 4.4 Genel Değerlendirme

Genel bir değerlendirme için eğitim veri kümesi sonuçları Tablo 18'de sunulmuştur.

Tablo 19: Yöntemlerin farklı ölçütlere göre eğitim kümesi performansları

Algoritma	Duyarlılık	Özgüllük	Kesinlik	Hata	Doğruluk
CART	0.84	0.84	0.28	0.16	0.84
RF	0.72	0.95	0.50	0.07	0.93
XGBoost	0.96	0.50	0.12	0.47	0.53

Tablo 19'a bakıldığında en yüksek doğruluk değeri Random Forest algoritmasında olurken duyarlılığı en yüksek algoritma ise XGBoost olmuştur. Bu tarz sınıf dengesizliğinin olduğu veri kümelerinde duyarlılık ve kesinlik değerlerine bakmak sonuçların daha anlaşılır yorumlanmasını sağlar. Bu tablodaki hesaplanan tüm ölçütler, kesim noktasının alışveriş gerçekleştirenler sınıfının yaygınlık değeri olan %8 alınmasıyla elde edilmiştir.

Kesim noktasının değişmesi durumunda eğitim veri kümesi için hesaplanan değerler de değişiklik gösterecektir. Bu sebeple kesim noktasından bağımsız olarak model performansını değerlendiren metriklerin dikkate alınması sonuçların daha objektif değerlendirilmesini sağlar. Eğitim veri kümesi için kesim noktasından bağımsız olarak hesaplanan ROC-AUC ve PR-AUC değerleri Tablo 19'da sunulmuştur..

Tablo 20: Yöntemlerin eğitim kümesindeki ROC-AUC ve PR-AUC değerleri

Algoritma	ROC-AUC	PR-AUC
CART	0.91	0.71
RF	0.90	0.72
XGBoost	0.92	0.68

Tablo 20'ye bakıldığında ROC-AUC ve PR-AUC değerlerinin birbirlerine yakın olmakla birlikte en yüksek ROC-AUC değeri XGBoost algoritmasında en yüksek PR-AUC değeri ise RF algoritmasında bulunmuştur.

Genel bir değerlendirme için test veri kümesi sonuçları Tablo 21'de sunulmuştur

Tablo 21: Yöntemlerin farklı ölçütlere göre test kümesi performansları

Algoritma	Duyarlılık	Özgüllük	Kesinlik	Hata	Doğruluk
<b>CART</b>	0.38	0.87	0.27	0.18	0.92
<b>RF</b>	0.23	0.91	0.25	0.17	0.83
<b>XGBoost</b>	0.69	0.50	0.15	0.47	0.53

Tablo 21'de test veri kümesi sonuçlarına bakıldığında en yüksek doğruluk değeri CART algoritmasında olurken duyarlılığı en yüksek algoritma ise XGBoost olmuştur. Bu tarz sınıf dengesizliğinin olduğu veri kümelerinde duyarlılık ve kesinlik değerlerine bakmak sonuçların daha anlaşılır yorumlanmasını sağlar. Bu tablodaki hesaplanan tüm metrikler, kesim noktasının alışveriş gerçekleştirenler sınıfının yaygınlık değeri olan %8 alınmasıyla elde edilmiştir. Kesim noktasının değişmesi durumunda hesaplanan değerler de değişiklik gösterecektir. Bu sebeple kesim noktasından bağımsız olarak model performansını değerlendiren metriklerin dikkate alınması sonuçların daha objektif değerlendirilmesini sağlar.

Test veri kümesi için kesim noktasından bağımsız olarak hesaplanan ROC-AUC ve PR-AUC değerleri Tablo 22'de sunulmuştur.

Tablo 22: Yöntemlerin test kümesindeki ROC-AUC ve PR-AUC değerleri

Algoritma	ROC-AUC	PR-AUC
<b>CART</b>	0.63	0.24
<b>RF</b>	0.66	0.35
<b>XGBoost</b>	0.67	0.31

Tablo 22'ye bakıldığında, ROC-AUC değerleri birbirine yakın bulurken, PR-AUC değerlerinde farklılar ortaya çıkmıştır. En yüksek PR-AUC değeri RF algoritmasında, ikinci en yüksek PR-AUC değeri de XGBoost algoritmasında bulunmuştur. Sınıf dengesizliğinin olduğu bu veri kümesinde PR-AUC değerine göre sonuçları değerlendirmek daha doğrudur. Bu sebeple veri kümesi için en iyi öğrenme Random Forest algoritmasında gerçekleşmiştir.

## 5. SONUÇLAR VE TARTIŞMA (RESULTS AND DISCUSSION)

Metin madenciliği, yapılandırılmamış veri türleri olan belgeler, sosyal medya gönderileri ve e-postalar gibi çok farklı formatta kaydedilmiş veri dosyalarından bilgi çıkarımı için kullanımı gittikçe yaygınlaşan bir yöntemdir. Metin madenciliği yöntemleri ile birlikte makine öğrenmesi yöntemleri kullanılarak büyük hacimli metinsel veriler içindeki kalıpların, eğilimlerin ve ilişkilerin ortaya çıkarılabilmesi karar verme süreçlerine yardımcı olarak işletmelerin veriye dayalı kararlar almasını kolaylaştırmaktadır. En önde gelen uygulama alanlarından birisi de müşterilerin belirli bir ürün veya hizmet hakkındaki yorumlarını, görüşlerini ve sosyal medya gönderilerini analiz ederek işletmelerin müşteri duygularını, tercihlerini ve geri bildirimlerinin incelenmesidir. Büyük boyuttaki metinsel verilere göz atarak hemen anlayamayabilecek gizli kalıpları ve ilişkileri tanımlayarak bilgi keşfini kolaylaştırması metin madenciliğinin çok farklı alanlara uygulanabilmesini kolaylaştırmıştır. Ayrıca çok uzun metin dosyalarını analiz etmeyi, özetlemeyi sağlayan, belgede geçen temel konuları ortaya çıkaran metin madenciliği yöntemleri sayesinde zaman ve kaynak tasarrufu sağlanabilmektedir.

Metin madenciliği uygulamalarında yaygın olarak takip şu aşamalar bulunmaktadır. Problemin tanımlanması, metin tabanlı veri kullanılarak ulaşılmak istenen amacın belirlenmesidir. Örneğin müşteri görüşleri kullanılarak duygu analizinin yapılması, metin içindeki konuların belirlenmesi veya korpus içindeki belgelerin kümelenmesi vb.), veri toplama ve veri ön işleme (analizde kullanılacak verilerin web, eposta, sosyal medya vb. gibi farklı ortamlardan ulaşılarak analizde kullanılacak korpus oluşturulduktan sonra verideki gürültüyü azaltabilmek için noktalama işaretlerinin ve özel karakterlerin kaldırılması, durma kelimelerinin kaldırılması, kelimelerinin eklerinin kaldırılarak köklerine ulaşılması vb.), öznitelik çıkarımı (metinsel verilerin tahminlerde kullanılacak makine öğrenmesi yöntemlerinin kullanımına hazır hale getirilebilmesi için nümerik hale getirilmesi gerekir. Bunun için Bag-of-Words, TF-IDF, Word Embedding yöntemlerinden birisi kullanılır), model seçimi ve eğitimi (Tahmin için kullanılacak makine öğrenmesi veya derin öğrenme algoritmalarının seçilmesidir. Bu seçimi veri büyüklüğü, verinin boyutu, oluşturulacak modelden beklenen açıklanabilirlik seviyesi gibi kriterler belirleyicidir. Analize hazır hale getirilmiş veri seti eğitim ve test kümesi şeklinde ayrılarak, eğitim kümesinde performansı en yüksek hale getirecek şekilde model parametreleri optimize edilir.), model değerlendirme

(Oluşturulan modellerin performanslarını belirlemek için örneğin sınıflandırma problemlerinde doğruluk, kesinlik, ROC-AUC gibi birçok farklı ölçüt kullanılır. Model performans ölçümünü daha güvenilir hale getirmek için çapraz geçirme de kullanılabilir), modelin kullanımı (oluşturulan en iyi model daha önceden eğitimde kullanılmayan, yeni veriler üzerinde test edilir. Zaman içinde modelin performansını korumak veya geliştirmek için model parametrelerinde gerekli düzenlemeler yapılabilir), yorumlama ve görselleştirme (elde edilen sonuçlardan karar verme sürecinde nasıl faydalanabileceği yorumlanır, modelin davranışı hakkında daha detaylı bilgi alabilmek ve sonuçları görselleştirmek için hata matrisi, ROC eğrileri veya kelime bulutları gibi teknikler kullanılabilir)

Bu çalışmada konut üretimi ve satışı yapan bir inşaat şirketinin müşterileri ile yaptığı görüşmelerin kayıtlarının yer aldığı metin dosyaları incelenerek gelecekte bir müşterinin yapılan görüşme sonucunda konut alma kararı alıp almayacağını başarılı bir şekilde tahmin edecek bir model oluşturulmuştur. Veri kümesinde müşteriler ile telefonla veya yüz yüze yapılan görüşmelerin kayıtları ve daha sonrasında bu müşterilerin konut alıp almadıkları bilgisi yer almaktadır. Dolayısıyla problem makine öğrenmesi açısından bir ikili sınıflandırma problemi olarak ele alınmıştır. Beklenildiği gibi görüşme sonucu konut alma kararı veren müşteri sayısının, konut alma düşüncesinden vazgeçen veya erteleyen müşteriler göre çok daha düşük olmasından dolayı, sınıfların genelde dengeli olduğu ikili sınıflandırma problemlerinin aksine bu problem dengesiz bir ikili sınıflandırma problemidir. Bu nedenle model performansının biraz düşük olabileceği çalışmanın başında öngörülmüştür. Verileri eksik olan, analize uygun olmayan görüşme kayıtları çıkarıldığında 579 müşteri görüşmesi kaydından oluşan bir veri kümesine ulaşılmıştır. Bu veri kümesi yukarıda detayları verilen ön işleme adımlarından sonra analize uygun hale getirilmiştir. Tüm görüşmeleri özetleyen vektör uzay modelini gösteren matris elde edilmiştir. Beklenildiği gibi bu seyrek (sparse) ve yüksek boyutlu (high dimensional) bir matristir. Bu iki problem makine öğrenmesi algoritmalarının performanslarını negatif yönde etkileyen faktörlerdir. Model oluşturma aşamasında Random Forest, Classification and Regression Tree (CART), ve XGBoost algoritmaları kullanılarak modellerin performansları karşılaştırılmıştır. Karşılaştırma için sınıflandırma problemlerinde yaygın olarak kullanılan kesinlik, doğruluk, özgüllük, ROC-AUC gibi ölçütlere ek olarak dengesiz sınıflandırma problemlerinin performansını analiz etmede daha anlamlı sonuçlar verdiği bilinen PR-AUC ölçütü de yer almıştır.

Duyarlılık ölçütüne göre analizde kullanılan CART, RF ve XGBoost yöntemleri test kümesi üzerinde sırasıyla 0.38, 0.23 ve 0.69 değerlerini vermişlerdir. Bu yönden XGBoost yönteminin gerçekte konut alan müşterilerin daha yüksek bir doğrulukla konut alacaklar sınıfında tahmin ettiği görülmektedir. Her iki sınıfın göz önüne alındığı tahminlerin genel olarak doğruluğuna bakıldığında ise elde edilen doğruluk değerleri sırasıyla 0.92, 0.83 ve 0.53 olarak elde edilmiştir. Bu yönden bakıldığında CART yönteminin önce çıktığı görülmektedir. Duyarlılık ve doğruluk gibi ölçütlerin elde edildiği hata matrislerindeki değerler çalışmada da ifade edildiği gibi tercih edilen eşik değerine göre değişmektedir. Çalışmada eşik değeri olarak tüm müşterilerin içinde görüşmeler sonucu konut alma kararı vermiş müşterilerin yüzdesi olan %8(0.08) değeri esas alınarak analizler yapılmıştır. Dolayısıyla farklı eşik değerleri tercih edilirse bu ölçütlerin değerleri de değişecektir. Eşik değerinden bağımsız olarak bir değerlendirme yapılabilmesini sağlayan PR-AUC değerlerine bakıldığında CART, RF, XGBoost yöntemleri sırasıyla 0.24,0.35 ve 0.31 bulunmuştur. Bu değerlerin hepsi seçilen eşik değeri olan 0.08 den büyük olduğu için dengesiz şekilde dağılan bu veri kümesi üzerinde yeterli performans gösterdikleri söylenebilir. PR-AUC değerine göre en iyi performansı Random Forest (RF) yöntemi göstermiştir.

Çalışma ileride daha fazla sayıda müşteri görüşmesinden elde edilen bir korpus üzerinde derin öğrenme yöntemleri (CNN veya RNN gibi) ve farklı öznitelik temsilleri ile (Word2Vect, BERT) gibi kullanılarak geliştirilebilir. Ayrıca verideki dengesizliğin sebep olduğu sorunları aşmak için literatürde yaygın olarak kullanılan SMOTE(Synthetic Minority Over-Sampling Technique) gibi yöntemlere yer verilerek sonuçlar incelenebilir.

#### KAYNAKLAR (REFERENCES)

- [1] C. C. Aggarwal and C. Zhai, Eds. Mining Text Data (An Introduction to Text Mining. Springer, 2012.
- [2] L. Duan and Y. Xiong, "Big data analytics and business analytics," *Journal of Management Analytics*, vol. 2, no. 1, pp. 1-21, 2015/01/02 2015, doi: 10.1080/23270012.2015.1020891.
- [3] M. A. Hearst, "Untangling text data mining," presented at the Proceedings of the 37th annual meeting of the Association for Computational Linguistics on Computational Linguistics, College Park, Maryland, 1999. [Online]. Available: <https://doi.org/10.3115/1034678.1034679>.
- [4] I. Feinerer, K. Hornik, and D. Meyer, "Text Mining Infrastructure in R," *Journal of Statistical Software*, vol. 25, no. 5, pp. 1 - 54, 03/31 2008, doi: 10.18637/jss.v025.i05.

- [5] R. Feldman and I. Dagan, "Knowledge Discovery in Textual Databases (KDT)," 06/28 1995.
- [6] A. Hotho, A. Nürnberger, and G. Paass, "A Brief Survey of Text Mining," LDV Forum - GLDV Journal for Computational Linguistics and Language Technology, vol. 20, pp. 19-62, 07/01 2005, doi: 10.21248/jlcl.20.2005.68.
- [7] R. Feldman, Ronen, Sanger, and James, The text mining handbook: Advanced approaches in analyzing unstructured data. 2007.
- [8] D. Delen and M. Crossland, "Seeding the survey and analysis of research literature with text mining," Expert Systems with Applications, vol. 34, pp. 1707-1720, 04/01 2008, doi: 10.1016/j.eswa.2007.01.035.
- [9] P. Hosseini, S. Khoshsirar, M. Jalayer, S. Das, and H. Zhou, "Application of text mining techniques to identify actual wrong-way driving (WWD) crashes in police reports," International Journal of Transportation Science and Technology, vol. 12, no. 4, pp. 1038-1051, 2023/12/01/ 2023, doi: <https://doi.org/10.1016/j.ijst.2022.12.002>.
- [10] S. Soleimani, M. Leitner, and J. Codjoe, "Applying machine learning, text mining, and spatial analysis techniques to develop a highway-railroad grade crossing consolidation model," Accident Analysis & Prevention, vol. 152, p. 105985, 2021/03/01/ 2021, doi: <https://doi.org/10.1016/j.aap.2021.105985>.
- [11] M. Nilashi et al., "Big social data and customer decision making in vegetarian restaurants: A combined machine learning method," Journal of Retailing and Consumer Services, vol. 62, no. 102630, 2021.
- [12] A. Petropoulos and V. Siakoulis, "Can central bank speeches predict financial market turbulence? Evidence from an adaptive NLP sentiment index analysis using XGBoost machine learning technique," Central Bank Review, vol. 21, no. 4, pp. 141-153, 2021/12/01/ 2021, doi: <https://doi.org/10.1016/j.cbrev.2021.12.002>.
- [13] S. Chatterjee, D. Goyal, A. Prakash, and J. Sharma, "Exploring healthcare/health-product ecommerce satisfaction: A text mining and machine learning application," Journal of Business Research, vol. 131, pp. 815-825, 2021/07/01/ 2021, doi: <https://doi.org/10.1016/j.jbusres.2020.10.043>.
- [14] W.-C. Lin, C.-F. Tsai, and H. Chen, "Factors affecting text mining based stock prediction: Text feature representations, machine learning models, and news platforms," Applied Soft Computing, vol. 130, p. 109673, 10/01 2022, doi: 10.1016/j.asoc.2022.109673.
- [15] C. Allenbrand, "Supervised and unsupervised learning models for pharmaceutical drug rating and classification using consumer generated reviews," Healthcare Analytics, vol. 5, p. 100288, 2024/06/01/ 2024, doi: <https://doi.org/10.1016/j.health.2023.100288>.
- [16] Y. Anagun, N. S. Bolel, S. Isik, and S. E. Ozkan, "DEEP LEARNING-BASED CUSTOMER COMPLAINT MANAGEMENT," Journal of Organizational Computing and Electronic Commerce, vol. 32, no. 3-4, pp. 217-231, 2022/10/02 2022, doi: 10.1080/10919392.2023.2210049.
- [17] S. Isik, Z. Kurt, Y. Anagun, and K. Ozkan, "Spam E-mail Classification Recurrent Neural Networks for Spam E-mail Classification on an Agglutinative Language," International Journal of Intelligent Systems and Applications in Engineering, vol. 8, no. 4, pp. 221-227, 12/30 2020, doi: 10.18201/ijisae.2020466316.
- [18] S. Baek, W. Jung, and S. H. Han, "A critical review of text-based research in construction: Data source, analysis method, and implications," Automation in Construction, vol. 132, p. 103915, 12/01 2021, doi: 10.1016/j.autcon.2021.103915.
- [19] H. Yan, M. Ma, Y. Wu, H. Fan, and C. Dong, "Overview and analysis of the text mining applications in the construction industry," Heliyon, vol. 8, no. 12, p. e12088, 2022/12/01/ 2022, doi: <https://doi.org/10.1016/j.heliyon.2022.e12088>.
- [20] A. Shamshiri, K. Ryu, and J. Y. Park, "Text mining and natural language processing in construction," Automation in Construction, vol. 158, p. 105200, 02/01 2024, doi: 10.1016/j.autcon.2023.105200.
- [21] R: A Language and Environment for Statistical Computing. (2021). R Foundation for Statistical Computing, Vienna, Austria. [Online]. Available: <https://www.R-project.org/>
- [22] E. Ikonomakis, S. Kotsiantis, and V. Tampakas, "Text Classification Using Machine Learning Techniques," WSEAS transactions on computers, vol. 4, pp. 966-974, 08/01 2005.
- [23] E. Leopold and J. Kindermann, "Text Categorization with Support Vector Machines. How to Represent Texts in Input Space?," Machine Learning, vol. 46, no. 1, pp. 423-444, 2002/01/01 2002, doi: 10.1023/A:1012491419635.
- [24] J. Han, M. Kamber, and J. Pei, Data mining : concepts and techniques, 3 ed. Morgan Kaufmann, 2012.
- [25] G. James, D. Witten, T. Hastie, and R. Tibshirani, An introduction to statistical learning : with applications in R. New York : Springer, 2013.
- [26] L. Breiman, "Random Forests," Machine Learning, vol. 45, no. 1, pp. 5-32, 2001.
- [27] L. Rokach, Pattern Classification Using Ensemble Methods. Singapore: World Scientific Publishing, 2010.
- [28] J. H. Friedman, "Greedy function approximation: a gradient boosting machine," Annals of statistics, pp. 1189-1232, 2001.
- [29] J. Son, I. Jung, K. Park, and B. Han, "Tracking-by-Segmentation with Online Gradient Boosting Decision Tree," in 2015 IEEE International Conference on Computer Vision (ICCV), 7-13 Dec. 2015 2015, pp. 3056-3064, doi: 10.1109/ICCV.2015.350.
- [30] G. Ke et al., "LightGBM: a highly efficient gradient boosting decision tree," presented at the Proceedings of the 31st International Conference on Neural Information Processing Systems, Long Beach, California, USA, 2017.
- [31] R. Mitchell and E. Frank, "Accelerating the XGBoost algorithm using GPU computing," PeerJ Comput. Sci., vol. 3, p. e127, 2017.
- [32] T. Chen and C. Guestrin, "XGBoost: A Scalable Tree Boosting System," presented at the Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, San Francisco, California, USA, 2016.

- [33] T. Hastie, R. Tibshirani, and J. Friedman, *The Elements of statistical learning : data mining, inference, and prediction*. New York: Springer (in English), 2018.
- [34] D. W. Hosmer, S. Lemeshow, and R. X. Sturdivant, *Applied Logistic Regression*, 3 ed. John Wiley & Sons, 2013.
- [35] A. Fernández, S. García, M. Galar, R. C. Prati, B. Krawczyk, and F. Herrera, *Learning from Imbalanced Data Sets*. Springer Nature Switzerland AG, 2018.
- [36] E. A. Freeman and G. G. Moisen, "A comparison of the performance of threshold criteria for binary classification in terms of predicted prevalence and kappa," *Ecological Modelling*, vol. 217, no. 1, pp. 48-58, 2008, doi: <https://doi.org/10.1016/j.ecolmodel.2008.05.015>.
- [37] J. S. Cramer, *Logit Models from Economics and Other Fields*. Cambridge University Press, 2003.
- [38] T. Saito and M. Rehmsmeier, "The Precision-Recall Plot Is More Informative than the ROC Plot When Evaluating Binary Classifiers on Imbalanced Datasets," *PLoS ONE* vol. 10, no. 3, p. e0118432, 2015.
- [39] J. Davis and M. Goadrich, "The relationship between Precision-Recall and ROC curves," presented at the Proceedings of the 23rd international conference on Machine learning, Pittsburgh, Pennsylvania, USA, 2006. [Online]. Available: <https://doi.org/10.1145/1143844.1143874>.
- [40] T. Therneau and B. Atkinson, "rpart: Recursive Partitioning and Regression Trees," 2019. [Online]. Available: <https://CRAN.R-project.org/package=rpart>.
- [41] J. Grau, I. Grosse, and J. Keilwagen, "PRROC: computing and visualizing precision-recall and receiver operating characteristic curves in R," *Bioinformatics*, vol. 31, no. 15, pp. 2595-2597, 2015.
- [42] J. Keilwagen, I. Grosse, and J. Grau, "Area under Precision-Recall Curves for Weighted and Unweighted Data," vol. 9, no. 3, 2014.
- [43] A. Liaw and M. Wiener, "Classification and Regression by randomForest," *R News*, vol. 2, no. 3, pp. 18-22, 2002. [Online]. Available: <https://CRAN.R-project.org/doc/Rnews/>.
- [44] T. Chen et al., "xgboost: Extreme Gradient Boosting," 2021. [Online]. Available: <https://CRAN.R-project.org/package=xgboost>.

# Enhancing Skin Cancer Diagnosis through the Integration of Deep Learning and Machine Learning Approaches

*Araştırma Makalesi/Research Article*

 Yahya DOĞAN<sup>1</sup>,  Cüneyt ÖZDEMİR<sup>1</sup>

<sup>1</sup>Bilgisayar Mühendisliği, Mühendislik Fakültesi, Bilgisayar Mühendisliği, Siirt Üniversitesi, Siirt, Türkiye

[yahyadogan@siirt.edu.tr](mailto:yahyadogan@siirt.edu.tr), [cozdemir@siirt.edu.tr](mailto:cozdemir@siirt.edu.tr)

(Geliş/Received:14.05.2024; Kabul/Accepted:21.10.2024)

DOI: 10.17671/gazibtd.1484037

**Abstract**— Skin cancer is a disease characterized by the uncontrolled proliferation of skin cells, typically manifesting as lesions or abnormal growths. Early diagnosis is critical for improving treatment outcomes. This study proposes an innovative approach to skin cancer diagnosis by integrating modern deep learning models with traditional machine learning algorithms. A three-phase methodology was developed. In the first phase, meaningful features were extracted from skin lesion images using various transfer learning models, including Xception, VGG16, ResNet152V2, InceptionV3, InceptionResNetV2, MobileNetV2, EfficientNetB2, and DenseNet201. In the second phase, dimensionality reduction was performed using Principal Component Analysis (PCA). In the final phase, the reduced feature sets were classified using K-Nearest Neighbors (KNN) and Random Forest (RF) algorithms. Experimental results demonstrated that the highest accuracy of 91.28% was achieved through the combination of DenseNet201 for feature extraction, PCA for dimensionality reduction, and Random Forest for classification. These findings highlight the effectiveness of integrating transfer learning models, dimensionality reduction techniques, and machine learning algorithms in enhancing the accuracy of skin cancer diagnosis.

**Keywords**— Skin cancer diagnosis, feature selection, dimensionality reduction, transfer learning, machine learning

## Cilt Kanseri Teşhisi için Hibrit Derin Öğrenme ve Makine Öğrenmesi Yaklaşımı

**Özet**— Cilt kanseri, cilt hücrelerinin kontrolsüz çoğalması sonucu ortaya çıkan ve genellikle lezyonlar veya yeni büyümeler olarak kendini gösteren bir hastalıktır. Erken teşhis, tedavi sonuçlarını iyileştirmek için kritik bir rol oynamaktadır. Bu çalışmada, modern derin öğrenme modelleri ile geleneksel makine öğrenimi algoritmalarını birleştirilerek cilt kanseri teşhisinde yenilikçi bir yöntem sunulmaktadır. Üç aşamalı bir metodoloji geliştirilmiştir. İlk aşamada, cilt lezyonlarının görüntülerinden anlamlı özellikler çıkarılmış ve bu amaçla Xception, VGG16, ResNet152V2, InceptionV3, InceptionResNetV2, MobileNetV2, EfficientNetB2 ve DenseNet201 gibi çeşitli transfer öğrenme modelleri değerlendirilmiştir. İkinci aşamada, Temel Bileşen Analizi (PCA) ile özellik boyutlarının azaltılması sağlanmış ve üçüncü aşamada ise bu indirgenmiş özellikler, K-En Yakın Komşular (KNN) ve Rastgele Orman (RF) algoritmaları ile sınıflandırılmıştır. Yapılan deneyler sonucunda en yüksek doğruluk %91.28 ile DenseNet201 modelinden elde edilen özelliklerin PCA ile boyutlarının azaltılarak RF algoritması ile sınıflandırılmasıyla elde edilmiştir. Bu bulgular, transfer öğrenme modelleri ile yapılan özellik çıkarma işlemlerinin, PCA ile boyut azaltmanın ve makine öğrenmesi algoritmalarının cilt kanseri teşhisinde yüksek performans sağladığını göstermektedir.

**Anahtar Kelimeler**— cilt kanseri teşhisi, öznetelik seçimi, boyut azaltma, transfer öğrenme, makine öğrenmesi

## 1. INTRODUCTION

The number of cancer patients continues to increase due to factors such as rising radiation exposure, environmental changes, viruses, and shifts in dietary habits [1]. Skin cancer emerges as a neoplastic growth resulting from the uncontrolled proliferation of cells within the skin, which is the body's largest organ. Acting as a protective barrier, the skin covers the body's outer layers and contains structures like hair follicles and sweat glands [2]. Each year, thousands of lives are lost to skin cancer [3],[4]. Early diagnosis can enhance therapeutic effectiveness and improve survival outcomes. Patients diagnosed at an early stage are more likely to respond to less invasive treatments and achieve favorable outcomes [5]. When detected early, skin cancer is highly treatable, with five-year survival rates reaching up to 99% [6].

Skin cancer commonly develops in sun-exposed areas, particularly on the face, neck, hands, and arms. The most prevalent types are basal, squamous, and melanoma, each requiring distinct treatment approaches [7],[8]. Traditionally, skin cancer diagnosis relies on physical examination and visual assessment of lesions. However, these methods have limitations, including diagnostic variability among physicians and the inability to detect deeper lesions [9],[10]. Furthermore, limited access to specialist dermatologists, especially in developing countries and rural regions, can hinder early diagnosis [11].

Automated diagnostic systems hold great promise for the early detection of skin cancer. Artificial intelligence-based methods, such as deep learning models, can identify and classify skin lesions, operating independently of dermatologists and potentially improving diagnostic accuracy [12]. However, challenges remain regarding the integration and reliability of these systems in clinical practice. Their performance compared to human dermatologists has not yet been fully validated, and large datasets are required for effective training and long-term sustainability [13]. Several studies have demonstrated the success of deep learning models in detecting and classifying skin lesions, accelerating the diagnostic process, and enhancing accuracy [14]. These models are increasingly being used as alternatives to traditional methods for skin cancer diagnosis [15]. By analyzing skin lesions, they can identify various cancer types, contributing to early detection efforts [16]. Furthermore, their ability to function independently of dermatologists increases access to diagnosis, particularly for patients in remote areas.

Research demonstrates that employing multiple deep learning models can significantly enhance accuracy in skin cancer classification, as ensemble approaches combine the strengths of different architectures to improve predictive performance [17]. In particular, transfer learning has emerged as a powerful technique, enabling models pre-trained on large datasets, such as ImageNet, to be fine-tuned for skin cancer diagnosis. This method not only accelerates the training process but also mitigates the issue of limited domain-specific data by leveraging knowledge

from other visual recognition tasks [18]. However, the deployment of deep learning models in this field faces several challenges. The need for large, high-quality labeled datasets is a primary constraint, as collecting and annotating dermatological images requires significant time, financial resources, and expertise. Furthermore, variability in image quality and data sources—including differences in resolution, lighting, and skin tones—can affect model performance and generalizability across populations [13]. The scarcity of annotated datasets also introduces the risk of overfitting, where the model performs well on the training data but struggles with unseen data in real-world scenarios.

The literature features numerous studies utilizing machine learning and deep learning techniques for skin cancer classification [19–21]. Garg et al. [5] proposed a methodology employing image processing techniques to detect melanoma skin cancer, integrating the ABCD rule, which evaluates asymmetry, border irregularity, color, and diameter. After performing illumination correction and lesion segmentation, the model achieved an accuracy of 91.6%. Pham et al. [22] utilized deep CNNs with data augmentation to classify skin lesions, compiling a dataset from multiple sources including the ISBI Challenge, the ISIC Archive, and the PH2 dataset. They evaluated the results using classifiers such as support vector machines, random forests, and artificial neural networks, and achieved an accuracy of 89%. Remya et al. [23] developed a framework combining deep learning and multi-modality data analysis to classify skin lesions. This approach demonstrated significant potential for early detection and accurate diagnosis, showing improvements in key parameters like sensitivity, specificity, and accuracy compared to traditional methods. The model's performance on real-world datasets suggests strong clinical application potential. Fahad et al. [24] proposed a shallow CNN architecture that effectively classifies skin lesions with reduced computational resources, achieving a test accuracy of 98.81% on the heavily imbalanced ISIC 2020 dermoscopic dataset. They used preprocessing and data augmentation techniques to remove image noise and balance the dataset, and compared the model's performance against different transfer learning models, showing superior speed and accuracy. Sivakumar et al. [19] introduced a ResNet50-based model that enhances accuracy through a comprehensive process including data acquisition, preprocessing, segmentation, feature extraction, and classification. The model achieved a peak accuracy of 94% and an F1-score of 93.9% for malignant melanoma detection on the ISIC dataset, demonstrating significant improvement over traditional methods. Furthermore, an application developed by the researchers facilitates faster and more accurate diagnosis of malignant melanoma. Manimurugan et al. [21] presented a novel deep learning architecture combining context-aware convolutional and recurrent neural networks (CA-CNN-RNN) for analyzing skin cancer images. Using H&E stained images from the Cancer Genome Atlas database, the model encoded local features into high-dimensional representations, achieving superior performance across



various classification architectures such as DarkNet-53, VGG-19, ResNet50, and Inception. The CA-CNN-RNN model demonstrated superior metrics, achieving 97.14% accuracy, 96.49% precision, 98.21% recall, and 96.50% F1 score.

This study presents a three-stage methodology for classifying different types of skin cancer, focusing on the synergy between deep learning and traditional machine learning techniques. In the first stage, several transfer learning models—including Xception, VGG16, ResNet152V2, InceptionV3, InceptionResNetV2, MobileNetV2, EfficientNetB2, and DenseNet201—were assessed for their ability to extract meaningful features from skin lesion images. These models, pre-trained on large datasets, were leveraged to transfer learned features, thereby addressing the challenge of limited domain-specific data and accelerating the training process. In the second stage, the features extracted from the best-performing transfer learning model were processed through PCA for dimensionality reduction. PCA helps retain essential information while reducing the feature space, which mitigates the risk of overfitting and decreases the computational burden during classification. Finally, in the third stage, the reduced feature set was classified using two machine learning algorithms: KNN and RF. KNN, known for its simplicity and effectiveness in classification tasks, offers robust performance when the dimensionality is reduced. RF, on the other hand, provides high flexibility and resilience against overfitting through its ensemble-based structure.

### 3. MATERIAL AND METHODS

This section provides a detailed description of the datasets, proposed methodology, training procedures, and evaluation metrics employed in this study. The goal is to ensure transparency and reproducibility of the research process. First, the dataset used for model training and testing is introduced, including its structure and class distribution. Following this, the proposed hybrid approach, which integrates transfer learning models with traditional machine learning algorithms, is explained in detail. The training details, such as hyperparameters, batch sizes, learning rates, and training epochs, are also discussed. Finally, the performance metrics used to evaluate the effectiveness of the models—such as accuracy, precision, recall, and F1-score—are outlined, providing a comprehensive view of the evaluation criteria used to assess the classification performance.

#### 3.1. Dataset

The dataset utilized in this study was obtained from the Kaggle online platform, provided by the International Skin Imaging Collaboration (ISIC), a global initiative dedicated to advancing melanoma diagnosis. The images used in this research are available at Kaggle: <https://www.kaggle.com/fanconic/skin-cancer-malignant-vs-benign>. The ISIC Archive is recognized as the most

extensive public repository of dermoscopic images of skin lesions, with all images undergoing rigorous quality control to ensure consistency and reliability. The dataset comprises 1440 images of benign skin lesions and 1197 images of malignant lesions, all standardized to a resolution of 224x224 pixels to ensure compatibility with deep learning models. For this study, the dataset was divided into separate training and test sets. The training set contains 1440 benign and 1197 malignant lesion images, yielding a total of 2637 images. The test set consists of 360 benign and 300 malignant images, amounting to 660 images. Additionally, 10% of the training data was randomly selected and reserved as a validation set to monitor the model's performance during training and prevent overfitting. Figure 1 presents the distribution of benign and malignant lesions in the training set, while Figure 2 illustrates the class proportions in the dataset. This carefully curated dataset ensures that the models are trained and evaluated on high-quality, balanced data, contributing to reliable and meaningful classification outcomes.



Figure 1. Distribution of benign and malignant lesion images in the training set. The bar chart shows that the training set consists of 1440 benign images and 1197 malignant images, highlighting the slight class imbalance present in the dataset.



Figure 2. Proportional distribution of benign and malignant lesion images in the training set. The pie chart illustrates that 54.7% of the images represent benign lesions, while 45.3% correspond to malignant lesions.

Figure 3 provides representative sample images from the dataset, illustrating both benign and malignant skin lesions.

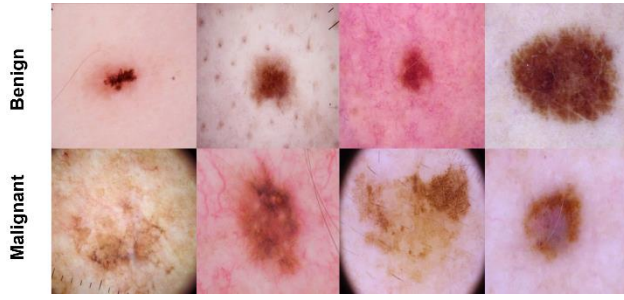


Figure 3: Randomly Selected Benign and Malignant Images from the Dataset

### 3.1. Proposed Approach

This section presents a detailed explanation of the proposed methodology, which integrates transfer learning models with traditional machine learning algorithms to classify skin cancer. The approach combines the strengths of both techniques, leveraging deep learning for feature extraction and machine learning for effective classification. The proposed framework utilizes a combination of feature extraction through deep learning models, dimensionality reduction techniques, and machine learning classifiers to enhance classification performance. The following subsections provide a comprehensive overview of each component and the workflow involved, highlighting the steps taken to ensure optimal model performance and generalization.

Initially, the DenseNet201 model was employed to extract meaningful features from the skin lesion images. This deep-learning architecture was selected due to its proven ability to capture complex patterns and fine-grained details, which are essential for differentiating between benign and malignant skin lesions. In this context, the DenseNet201 model performs a series of convolutional operations on the input image, denoted as  $x$ , to generate a comprehensive feature representation. The output of the DenseNet201 model consists of a feature vector that encapsulates the relevant information extracted from each image, providing a lower-dimensional representation suitable for further processing. Each element within the feature vector corresponds to a distinct characteristic or pattern identified in the image during the feature extraction process. This process is represented mathematically as the function  $F(x)$ , where  $x$  refers to the input image.

$$F(x) = [f_1, f_2, \dots, f_n] \quad (1)$$

Where  $F(x)$  denotes the extracted feature vector, with  $[f_1, f_2, \dots, f_n]$  representing the individual features. Each  $f_i$  (where  $i=0, 1, 2, \dots, n$ ) is a numerical value that captures a specific attribute of the input image. This feature vector serves as a crucial input for subsequent stages, such as dimensionality reduction and classification, facilitating accurate and efficient diagnosis of skin cancer.

Secondly, the PCA method was applied to reduce the dimensionality of the obtained feature vectors. PCA performs dimensionality reduction by performing principal

component analysis on the feature vectors. Accordingly, the covariance matrix  $C$  of the feature vector is computed, and transformation is carried out based on the eigenvalues and eigenvectors of this matrix. These operations are provided mathematically in Equations 2, 3, and 4.

$$C = \frac{1}{n}(X - \bar{x})^T(X - \bar{x}) \quad (2)$$

Where  $C$  represents the covariance matrix of the feature vector,  $F(x)$  denotes the data matrix formed by feature vectors, and  $\bar{x}$  represents the mean of the feature vectors.

$$Cv_i = \lambda_i v_i \quad (3)$$

Where  $v_i$  represents the eigenvectors, forming the principal components of the PCA transformation,  $\lambda_i$  denotes the eigenvalues determining the significance of the eigenvectors, and  $Z$  represents the matrix of newly obtained feature vectors after dimensionality reduction.

$$Z = X_{v_{selected}} \quad (4)$$

After dimensionality reduction, feature vectors are classified using the RF classifier. RF operates by constructing multiple decision trees during training and outputting the mode of the classes for classification. For each new example, every decision tree within the forest independently provides a classification. The final class is then determined by taking the majority vote across all trees, meaning the class predicted by the most trees becomes the overall prediction. Mathematically, this process is expressed as:

$$y = mode(y_{trees}) \quad (5)$$

Where  $y$  represents the predicted class of the new example,  $y_{trees}$  denotes the set of class predictions from all decision trees in the forest, and the mode operator indicates the most frequently occurring value among these predictions.

### 3.1. Metrics

The performance of the proposed skin cancer classification model was assessed using several key metrics to evaluate its accuracy, reliability, and ability to distinguish between benign and malignant lesions. The following metrics were employed to ensure a comprehensive performance evaluation:

**Accuracy** represents the overall effectiveness of the model by calculating the percentage of correctly classified cases among all predictions. For skin cancer classification, it reflects how well the model identifies both benign and malignant lesions accurately.

$$Accuracy = (TP + TN) / (TP + TN + FP + FN) \quad (6)$$

**Precision** measures the proportion of true malignant cases among all instances predicted as malignant. In skin cancer diagnosis, a high precision rate ensures that the model minimizes false positives, reducing unnecessary patient anxiety and avoiding unnecessary biopsies or treatments.

$$\text{Precision} = TP / (TP + FP) \quad (7)$$

**Recall**, also known as sensitivity, reflects the model's ability to correctly identify all malignant lesions. In the context of skin cancer, recall is crucial as missing a malignant case (false negative) could delay treatment, potentially worsening patient outcomes.

$$\text{Recall} = TP / (TP + FN) \quad (8)$$

**The F1 Score** provides a balance between precision and recall, making it an important metric in scenarios where class imbalance exists, such as in the skin cancer dataset. A high F1 Score ensures that the model performs well in both identifying true positives and minimizing false positives.

$$\text{F1 Score} = 2 \times (\text{Precision} \times \text{Recall}) / (\text{Precision} + \text{Recall}) \quad (9)$$

**The Area Under the Curve (AUC)** of the Receiver Operating Characteristic (ROC) curve evaluates the model's capability to distinguish between benign and malignant lesions across different thresholds. A high AUC

value indicates that the model is effective in differentiating between the two classes, which is essential for reliable skin cancer diagnosis.

### 3.1. Training details

The training process for the skin cancer classification model was configured with several key hyperparameters to ensure optimal performance. All input images were resized to 224x224 pixels and normalized by scaling pixel values to the range [0, 1] (1/255.0) to facilitate faster convergence. The Adam optimizer was employed for weight updates due to its adaptive learning rate and ability to handle sparse gradients effectively. As the task involved binary classification, binary cross-entropy was used as the loss function to measure the discrepancy between predicted probabilities and true labels. The model was trained for a maximum of 100 epochs with a batch size of 32, balancing training speed and memory efficiency. To prevent overfitting, early stopping with a patience value of 6 was applied, meaning the training process stopped if the validation performance failed to improve for six consecutive epochs. Table 1 shows the hyperparameters used in the training phase of Transfer learning models used as feature extractors.

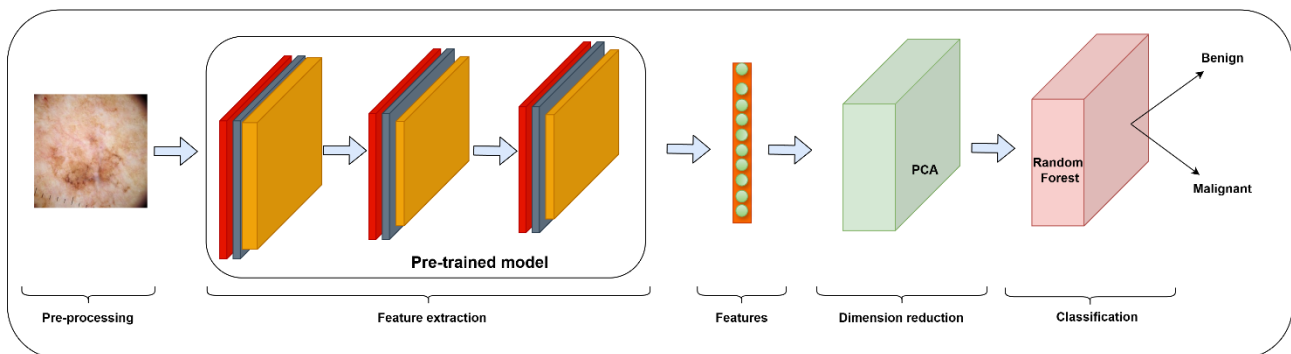


Figure 4: The proposed architecture involves several stages. Initially, preprocessing was applied to the images before training. The DenseNet201 model was then used to extract relevant features. Following this, the most significant feature vectors were selected using the PCA dimensionality reduction algorithm. Finally, the RF algorithm was employed to classify the reduced feature set

Table 1: Hyperparameters

Hyperparameters	Value
Image size	224x224
Normalization	Yes (1/255.0)
Optimization Algorithm	Adam
Early Stopping	Yes (Patience: 6)
Epoch	100
Batch Size	32
Loss Function	Binary cross entropy

## 4. EXPERIMENT AND RESULTS

This section presents the experimental results obtained using the proposed approach. The experiments were conducted with several popular transfer learning models, including Xception, VGG16, ResNet152V2, InceptionV3, InceptionResNetV2, MobileNetV2, EfficientNetB2, and DenseNet201, to evaluate their performance in classifying benign and malignant skin lesions. During preprocessing, only normalization (scaling pixel values to the range [0,1]) was applied, without the use of any additional preprocessing or data augmentation techniques. Feature extraction was performed using the selected transfer

learning models, where only the output layer with a sigmoid activation function was utilized, with no additional hidden layers added. Binary cross-entropy was employed as the loss function for optimization. A 10% subset of the training data was reserved for validation, while the images designated for testing, as outlined in the dataset section, were used for the final evaluation. The summarized performance metrics of each model are presented in Table 2.

Among all the models, DenseNet201 delivered the best performance, achieving a test accuracy of 90.61%, an F1 Score of 90.60%, and an AUC of 90.50%. This model's superior performance can be attributed to its architecture, which effectively captures complex features through densely connected layers, making it highly suitable for differentiating between benign and malignant skin lesions. Additionally, DenseNet201 maintained a well-balanced performance between precision (90.60%) and recall (90.61%), indicating that it effectively identified malignant cases without generating an excessive number of false positives.

The Xception model also demonstrated strong performance, with a test accuracy of 86.97% and an F1 Score of 86.98%, making it the second-best model in the study. The high accuracy of Xception reflects its ability to capture spatial and abstract features efficiently, contributing to reliable classification results. Similarly, ResNet152V2 performed well, achieving 86.52% accuracy, closely trailing Xception. Both models' high F1 Scores (86.98% and 86.49%) confirm their suitability for skin cancer classification tasks.

Other models, such as InceptionV3, VGG16, and InceptionResNetV2, produced moderate performance with test accuracies ranging between 81.52% and 85.00%. While these models demonstrated acceptable precision and recall values, their overall performance was slightly inferior to that of DenseNet201 and Xception. These results suggest that while these models are effective in identifying patterns in skin lesions, they may not generalize as well to unseen data compared to more advanced architectures like DenseNet201.

MobileNetV2 and EfficientNetB2 performed poorly, failing to achieve satisfactory results. MobileNetV2, with a test accuracy of 57.42%, exhibited poor generalization, as reflected in its low F1 Score of 44.65%. Although its precision (76.09%) was relatively high, this model failed to maintain adequate recall, leading to an imbalanced classification performance. EfficientNetB2 demonstrated even weaker performance, with an accuracy of 54.55% and a precision of only 29.75%, despite achieving a recall of 54.55%. The low F1 Score (38.50%) indicates that EfficientNetB2 was not well-suited for this classification task. Figure 5 presents the loss and accuracy graphs for DenseNet201, the model that achieved the highest performance in this study. In contrast, Figure 6 displays the loss and accuracy graphs for EfficientNetB2, which demonstrated the lowest performance. These graphs provide insights into the training process of both models, highlighting DenseNet201's superior ability to converge effectively and achieve high accuracy while illustrating the challenges faced by EfficientNetB2 in learning from the dataset. Figure 7 illustrates the ROC curve for the DenseNet201 model, which provides a visual representation of the model's ability to distinguish between benign and malignant skin lesions at various threshold settings. The area under the curve (AUC) indicates the overall effectiveness of the model in terms of sensitivity and specificity. A higher AUC suggests better performance in correctly classifying positive and negative instances. Figure 8 shows the confusion matrix, which details the model's classification outcomes by comparing the predicted labels with the actual labels. The matrix provides insight into the number of true positives, true negatives, false positives, and false negatives, allowing for a more detailed evaluation of the model's accuracy and error distribution.

In summary, the experimental results highlight that DenseNet201 outperformed all other models, providing the most accurate and balanced classification performance. In contrast, MobileNetV2 and EfficientNetB2 were less effective, suggesting that lightweight architectures may struggle with the complexities of skin cancer detection. These findings underscore the importance of selecting robust deep learning models, such as DenseNet201, for tasks requiring high precision and reliability in medical diagnostic.

Table 2: Performance Results of Transfer Learning Models

Model	Accuracy (%)	Precision (%)	Recall (%)	F1 Score (%)	AUC (%)
Xception	86.97	87.02	86.97	86.98	86.97
VGG16	81.52	84.41	81.52	81.40	82.61
ResNet152V2	86.52	86.51	86.52	86.49	86.25
InceptionV3	84.55	84.78	84.55	84.43	83.94
InceptionResNetV2	85.00	86.02	85.00	84.74	84.03
MobileNetV2	57.42	76.09	57.42	44.65	53.17
<b>DenseNet201</b>	<b>90.61</b>	<b>90.60</b>	<b>90.61</b>	<b>90.60</b>	<b>90.50</b>
EfficientNetB2	54.55	29.75	54.55	38.50	50.00

The performance of deep learning models is largely influenced by their ability to capture and learn complex patterns within the data. While these models excel at extracting deep features, the resulting high-dimensional feature space often leads to increased computational demands, making model training both time-consuming and resource-intensive. To mitigate these challenges, dimensionality reduction techniques like PCA are employed. PCA reduces the number of features by identifying the most significant components that preserve the essential variability in the data, thereby enabling more efficient and faster processing.

In the second phase of the experimental studies, rather than directly training transfer learning models—which can be computationally expensive and time-consuming—features were first extracted using the DenseNet201 model. DenseNet201 was selected based on its superior performance in earlier experiments, as it demonstrated the best classification accuracy and balance across key metrics. The high-dimensional features obtained from this model were then subjected to dimensionality reduction using PCA to minimize complexity without losing critical information.

After reducing the feature set with PCA, two machine learning algorithms, KNN and RF, were used to classify the reduced features. This approach combines the deep feature extraction power of DenseNet201 with the efficiency of traditional machine learning classifiers. The RF and KNN algorithms were chosen for their complementary strengths—RF for its robustness and ability to handle large datasets, and KNN for its simplicity and effectiveness in smaller, well-structured datasets.

The classification results obtained using this approach were evaluated on the test set, and the performance metrics, including accuracy, precision, recall, F1 score, and AUC, are presented in Table 3. The results demonstrate the effectiveness of integrating deep learning with dimensionality reduction and machine learning classifiers, offering a more computationally efficient solution while maintaining high classification performance.

This combined methodology allows for a scalable, efficient approach to skin cancer diagnosis, enabling the model to perform well even when computational resources are limited. The success of this approach highlights the potential of using feature extraction followed by dimensionality reduction and machine learning classification as a viable alternative to fully training deep learning models in resource-constrained environments.

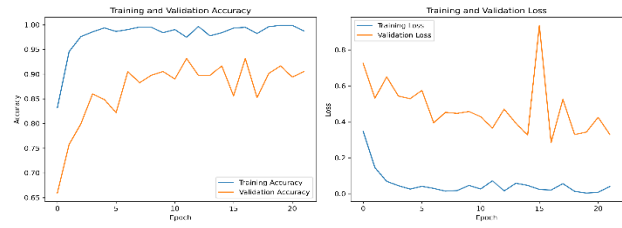


Figure 5: Loss and accuracy graph for DenseNet201

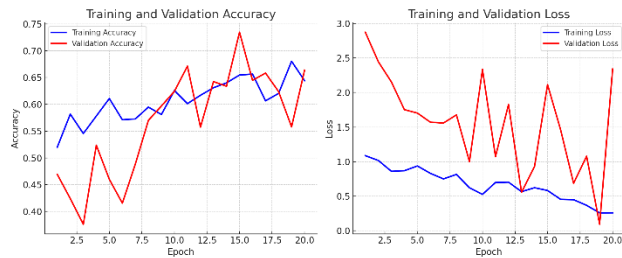


Figure 6: Loss and accuracy graph for EfficientNetB2

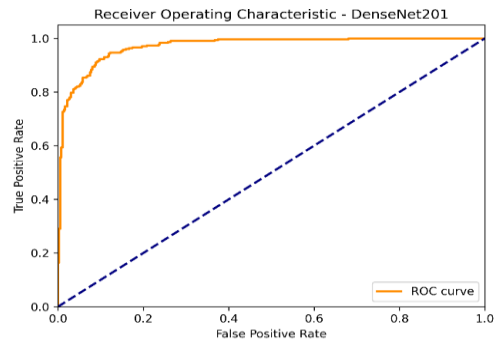


Figure 7: Roc curve for DenseNet201

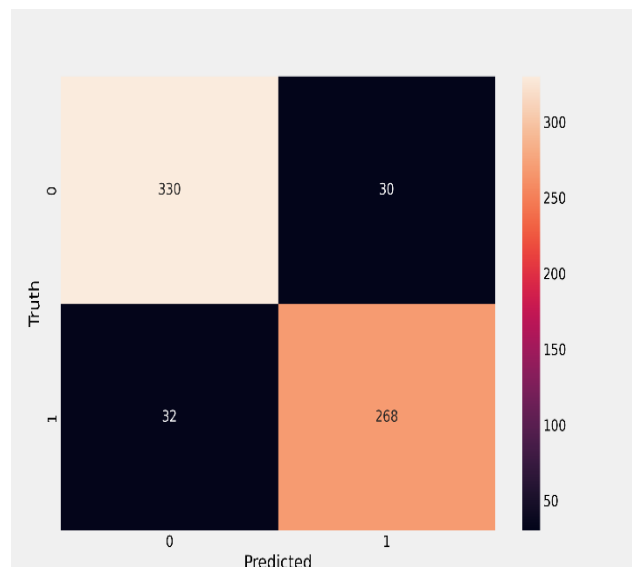


Figure 8: Confusion matrix

Table 3: Results of the Hybrid Model Approach

Model	Test Accuracy (%)	Precision (%)	Recall (%)	F1 Score (%)	AUC (%)
KNN	91.1	91.07	91.1	91.09	90.85
Random Forest	91.28	91.58	91.28	91.32	91.69

Table 3 presents the performance metrics of hybrid models utilizing different machine learning algorithms after feature extraction by a deep learning model. The KNN classifier yielded notable results with a test accuracy of 91.1%, precision of 91.07%, recall of 91.1%, F1 score of 91.09%, and an AUC of 90.85%. In comparison, the Random Forest classifier demonstrated marginally superior performance, achieving a test accuracy of 91.28%, precision of 91.58%, recall of 91.28%, F1 score of 91.32%, and an AUC of 91.69%.

The results indicate that both classifiers when applied to the features extracted from the DenseNet201 model and processed through PCA, exhibit high performance metrics. However, the RF classifier consistently performs better than KNN across all evaluated metrics. This indicates that while KNN delivers solid performance, the Random Forest model demonstrates higher classification accuracy and AUC, proving to be more effective in handling dimensionality-reduced features. These findings underline the efficacy of using a deep learning model for feature extraction combined with PCA for dimensionality reduction, followed by classification with machine learning algorithms. Figure 9 illustrates the qualitative performance of the proposed approach by comparing the predicted outputs with ground truth on randomly selected test images. This figure further validates the effectiveness and reliability of the hybrid model approach.

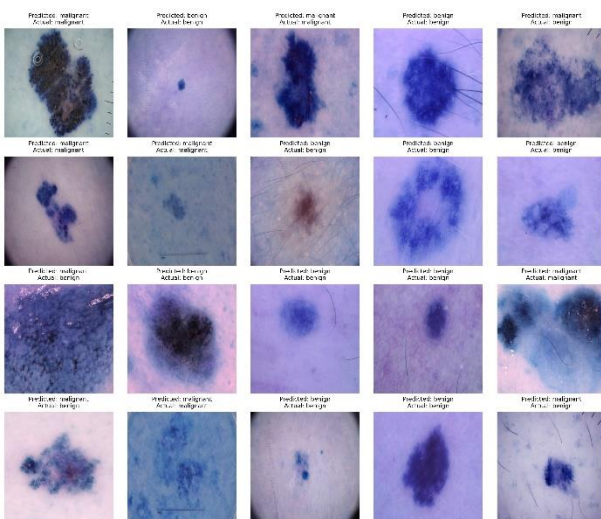


Figure 9: Results of the qualitative assessment of the proposed approach's performance on test images

## 5. CONCLUSION

This study demonstrates the effectiveness of a hybrid approach that integrates deep learning models with machine learning algorithms for the early detection of skin cancer. In the first phase, several transfer learning models were evaluated, with DenseNet201 emerging as the top performer, achieving high accuracy, precision, recall, F1 score, and AUC. Given its superior performance, DenseNet201 was selected for feature extraction in the subsequent phase, and PCA was employed to reduce the dimensionality of the extracted features.

The reduced features were classified using machine learning algorithms, with RF delivering the best overall performance, surpassing the results of KNN. Findings highlight that integrating dimensionality reduction with machine learning classifiers can enhance performance beyond what is achievable by deep learning models alone. Specifically, the RF classifier demonstrated robustness in handling the dimensionality-reduced feature space, providing superior accuracy and balanced performance.

This approach underscores that machine learning techniques, when applied to features extracted from transfer learning models, can offer meaningful improvements in classification tasks. Rather than aiming for state-of-the-art results, the study emphasizes the practicality and efficiency of combining these methodologies to achieve reliable outcomes.

In conclusion, the integration of deep learning with machine learning techniques presents a practical approach for enhancing skin cancer diagnosis, with considerable potential benefits for clinical practice. Future studies will focus on expanding the dataset and incorporating more diverse skin lesion types to improve the model's generalization and robustness.

## REFERENCES

- [1] Alam, T. M., Khan, M. M. A., Iqbal, M. A., Abdul, W., & Mushtaq, M. (2019). Cervical cancer prediction through different screening methods using data mining. *IJACSA International Journal of Advanced Computer Science and Applications*, 10(2).
- [2] Rognoni, E., & Watt, F. M. (2018). Skin cell heterogeneity in development, wound healing, and cancer. *Trends in cell biology*, 28(9), 709-722.
- [3] Fitzmaurice C, Akinyemiju TF, Al Lami FH, Alam T, Alizadeh-Navaei R, Allen C, et al. (2018) Global, regional, and national cancer incidence, mortality, years of life lost, and disability-adjusted life-years for 29 cancer groups, 1990 to 2016: a systematic analysis for the global burden of disease study. *JAMA Oncol.*;4(11):1553–68
- [4] Dinehart, S. M. (2000). The treatment of actinic keratoses. *Journal of the American Academy of Dermatology*, 42(1), S25-S28
- [5] Flohr, C., and Hay, R. 2021. Putting the burden of skin diseases on the global map. *British Journal of Dermatology\**, 184(2): 189-190.
- [6] Skin Cancer Facts & Statistics [Internet]. 2021. Available from: <https://www.skincancer.org/skin-cancer-information/skin-cancer-facts/>
- [7] Ergün, E., & Kılıç, K. (2021). Derin öğrenme ile artırılmış görüntü seti üzerinden cilt kanseri tespiti. *Black Sea Journal of Engineering and Science*, 4(4), 192-200.
- [8] Andrew, T. W., Alrawi, M., & Lovat, P. (2021). Reduction in skin cancer diagnoses in the UK during the COVID-19 pandemic. *Clinical and Experimental Dermatology*, 46(1),145-146.
- [9] Lukaviciute, L., Ganceviciene, R., Navickas, P., Navickas, A., Grigaitiene, J., and Zouboulis, C. C. 2020. Anxiety, depression, and suicidal ideation amongst patients with facial dermatoses (acne, rosacea, perioral dermatitis, and folliculitis) in Lithuania. *Dermatology\**, 236(4): 314-322.
- [10] Arnold, J. D., Yoon, S., and Kirkorian, A. Y. 2019. The national burden of inpatient dermatology in adults. *Journal of the American Academy of Dermatology\**, 80(2): 425-432.
- [11] Feng, H., Berk-Krauss, J., Feng, P. W., and Stein, J. A. 2018. Comparison of dermatologist density between urban and rural counties in the United States. *JAMA Dermatology\**, 154(11): 1265-1271.
- [12] Assiri, A., et al. 2013. Hospital outbreak of Middle East respiratory syndrome coronavirus. *The New England Journal of Medicine\**, 369(5): 407-416.
- [13] W. Yue, Z. Wang, H. Chen, A. Payne, & X. Liu, "Machine learning with applications in breast cancer diagnosis and prognosis", *Designs\**, vol. 2, no. 2, p. 13, 2018.
- [14] Z. Han, B. Wei, Y. Zheng, Y. Yin, K. Li, & S. Li, "Breast cancer multi-classification from histopathological images with structured deep learning model", *Scientific Reports\**, vol. 7, no. 1, 2017.
- [15] C. AKYEL and N. Arici, "A new approach to hair noise cleansing and lesion segmentation in images of skin cancer", *Politeknik Dergisi*, vol. 23, no. 3, p. 821-828, 2020. DOI: 10.2339/politeknik.645395
- [16] A. Karli, "Cilt kanseri görüntüleri kullanılarak eğitilen efficientnet-b3 mimarisinde hiperparametre seçiminin sınıflandırma performansına etkisinin incelenmesi", *Firat Üniversitesi Mühendislik Bilimleri Dergisi*, vol. 36, no. 1, p. 499-507, 2024. DOI: 10.35234/fumbd.1426044
- [17] N. Kausar, A. Hameed, M. Sattar, R. Ashraf, A. Imran, M. Abidinet al., "Multiclass skin cancer classification using ensemble of fine-tuned deep learning models", *Applied Sciences*, vol. 11, no. 22, p. 10593, 2021. DOI: 10.3390/app112210593
- [18] A. Mahbod, G. Schaefer, C. Wang, G. Dorffner, R. Ecker, & I. Ellinger, "Transfer learning using a multi-scale and multi-network ensemble for skin lesion classification", *Computer Methods and Programs in Biomedicine*, vol. 193, p. 105475, 2020. DOI: 10.1016/j.cmpb.2020.105475
- [19] Sivakumar, M. S., Leo, L. M., Gurumekala, T., Sindhu, V., & Priyadarshini, A. S. (2024). Deep learning in skin lesion analysis for malignant melanoma cancer identification. *Multimedia Tools and Applications*, 83(6), 17833-17853.
- [20] Gajera, H. K., Nayak, D. R., & Zaveri, M. A. (2023). A comprehensive analysis of dermoscopy images for melanoma detection via deep CNN features. *Biomedical Signal Processing and Control*, 79, 104186.
- [21] Manimurugan, S. (2023). Hybrid high performance intelligent computing approach of CACNN and RNN for skin cancer image grading. *Soft Computing*, 27(1), 579-589.
- [22] Pham, T.C., Luong, C.M., Visani, M. and Hoang, V.D., 2018, March. Deep CNN and data augmentation for skin lesion classification. In *Asian Conference on Intelligent Information and Database Systems* (pp. 573-582). Springer, Cham.
- [23] Remya, S., Anjali, T., & Sugumaran, V. (2024). A Novel Transfer Learning Framework for Multimodal Skin Lesion Analysis. *IEEE Access*.
- [24] Fahad, N. M., Sakib, S., Raiaan, M. A. K., & Mukta, M. S. H. (2023, February). Skinnet-8: An efficient cnn architecture for classifying skin cancer on an imbalanced dataset. In *2023 International conference on electrical, computer and communication engineering (ECCE)* (pp. 1-6). IEEE.