e-ISSN: 2602-3563

ACTA INFOLOGICA (ACIN)

JUNE, 2024
Volume: 8 | Issue: 1

ISTANBUL UNIVERSITY, INFORMATICS DEPARTMENT

acin.istanbul.edu.tr
https://dergipark.org.tr/tr/pub/acin

İSTANBUL UNIVERSITY PRESS

Acta Infologica
JUNE 2024,
VOLUME 8 ISSUE 1

İSTANBUL
UNIVERSITY
PRESS

Indexing and Abstracting

TÜBİTAK-ULAKBİM TR Dizin
EBSCO Applied Sciences Source Ultimate
Erih Plus
DOAJ
Bielefeld Academic Search Engine (BASE)
OpenAIRE
ResearchBib
ASOS Index

Owner
Prof. Dr. Sevinç GÜLSEÇEN
Istanbul University, Department of Informatics, Istanbul, Turkiye

Responsible Manager
Prof. Dr. Sevinç GÜLSEÇEN
Istanbul University, Department of Informatics, Istanbul, Turkiye

Correspondence Address
Istanbul University, Department of Informatics
Kalenderhane Mahallesi, 16 Mart Şehitleri Caddesi, No: 8
Vezneciler, Fatih, İstanbul, Türkiye
Phone: +90 212 440 00 00/10037
E-mail: acin@istanbul.edu.tr
http://iupress.istanbul.edu.tr/tr/journal/acin/home
https://dergipark.org.tr/tr/pub/acin

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

Authors bear responsibility for the content of their published articles.

The publication language of the journal is English

This is a scholarly, international, peer-reviewed and open-access journal published biannually in June and December.

Publication Type: Periodical

EDITORIAL MANAGEMENT BOARD

EDITORIAL BOARD

Acta Infologica
JUNE 2024,
VOLUME 8 ISSUE 1

İSTANBUL UNIVERSITY PRESS

CONTENTS

RESEARCH ARTICLES

REVIEW ARTICLE

Acta Infologica
acin 2024, 8 (1), 1–10
DOI: 10.26650/acin.1338604

Research Article / Araştırma Makalesi

Transformer-Based Turkish Automatic Speech Recognition

Davut Emre Taşar[1], Kutan Koruyan[2], Cihan Çılgın[3]

[1]Dokuz Eylül University, Graduate School of Social Sciences, Department of Management Information Systems, İzmir, Türkiye
[2]Dokuz Eylül University, Faculty of Economics and Administrative Sciences, Department of Management Information Systems, İzmir, Türkiye
[3]Bolu Abant İzzet Baysal University, Gerede Faculty of Applied Sciences, Department of Management Information Systems, Bolu, Türkiye

Corresponding author : Kutan Koruyan
E-mail : kutan.koruyan@deu.edu.tr

ABSTRACT

Today, businesses use Automatic Speech Recognition (ASR) technology more frequently to increase efficiency and productivity while performing many business functions. Due to the increased prevalence of online meetings in remote working and learning environments after the COVID-19 pandemic, speech recognition systems have seen more frequent utilization, exhibiting the significance of these systems. While English, Spanish or French languages have a lot of labeled data, there is very little labeled data for the Turkish language. This directly affects the accuracy of the ASR system negatively. Therefore, this study utilizes unlabeled audio data by learning general data representations with self-supervised learning end-to-end modeling. This study employed a transformer-based machine learning model with improved performance through transfer learning to convert speech recordings to text. The model adopted within the scope of the study is the Wav2Vec 2.0 architecture, which masks the audio inputs and solves the related task. The XLSR-Wav2Vec 2.0 model was pre-trained on speech data in 53 languages and fine-tuned with the Mozilla Common Voice Turkish data set. According to the empirical results obtained within the scope of the study, a 0.23 word error rate was reached in the test set of the same data set.

Keywords: Wav2vec2, automatic speech recognition, speech-to-text transcription, natural language processing, transformer architecture

Submitted	: 06.08.2023
Revision Requested	: 30.10.2023
Last Revision Received	: 29.11.2023
Accepted	: 30.11.2023
Published Online	: 29.02.2024

1. INTRODUCTION

Speech is the most basic and efficient form of communication between people (Malik, Malik, Mehmood, & Makhdoom, 2021; Padmanabhan & Johnson Premkumar, 2015). However, today's world is developing from an environment where people only communicate with each other to one where they can communicate with sensors and machines. Therefore, in today's world, especially with the development of artificial intelligence in every field, the need for Automatic Speech Recognition (ASR), which aims to provide human-machine or machine-human interaction, has both increased and developed rapidly in parallel with these developments. Yu & Deng, 2016). Although ASR has been a very interesting area for researchers for a long time, it has become much more suitable for use today, although it does not provide effective results compared to many other input tools (keyboard, mouse, touch screen, etc.) (Yu & Deng, 2016). Seen as a science fiction scenario in the past (Ghai & Singh, 2012), ASRs that can only respond to fluent natural language but still have technological barriers to user satisfaction, as pointed out in 2007 by Benzeghiba et al. (2007) and even by Cutajar, Gatt, Grech, Casha, & Micallef (2013) in 2013, are now a reality. In support of this situation, Malik et al. (2021) today described ASRs as the basic communication tools between humans and machines. Recently, the performance of ASR technologies has almost reached that of human transcribers (Amodei et al., 2016; Chiu et al., 2018; Xiong et al., 2017). Despite these developments, ASR is a multidisciplinary field of study, which requires knowledge from disciplines such as linguistics, computer science, signal processing, acoustics, communication theory, statistics, physiology, and psychology (Levis & Suvorov, 2012).

ASR is the process of turning speech into text after the machine recognizes and understands the speech signal, and includes extraction and determination of acoustic features, and acoustic and language models (Shi, 2021). ASR is a critical research area and has a wide range of applications, such as phone applications (Kurian & Balakrishnan, 2009; Tran, Truong, Le, Huh, & Huh, 2023), toys, games, language translations, educational applications such as language learning (Dai & Wu, 2023; Levis & Suvorov, 2012), air control, security and home automation (Cutajar et al., 2013; Annam, Neelima, Parasa & Chinamuttevi, 2023), customer service (Zekveld, Kramer, Kessens, Vlaming, & Houtgast, 2009; Pragati, Kolli, Jain, Sunethra & Nagarathna, 2023), business management (Danis & Karat, 1995; Xie, 2023), evidence gathering (Negrão & Domingues, 2021; Vásquez-Correa & Álvarez Muniain, 2023), healthcare and virtual assistants (Akhilesh, Brinda, Keerthana, Gupta, & Vekkot, 2022).

Parallel to the development of traditional ASR methods, the applications of these ASRs have also become critical for organizational development. Therefore, end-to-end models (E2E) have started to be developed, and more innovative methods have begun to be developed in this area (Yu & Deng, 2016). Although E2E modeling based on machine learning requires large amounts of labelled data (Jain et al., 2023; Yi, Wang, Cheng, Zhou, & Xu, 2021), it does not need complex modeling processes or a manually designed dictionary by humans (Yi, Wang, Cheng, Zhou, & Xu, 2020). Although machine learning based E2E learning offers significant advantages, it faces problems when it is used for ASR purposes, especially in languages with few resources, due to the need for a high number of labeled training sets. Therefore, for machine learning based E2E learning, it is necessary to pre-train the model in a partially supervised or self-supervised learning (SSL) way (Inaguma, Cho, Baskar, Kawahara, & Watanabe, 2019; Schneider, Baevski, Collobert, & Auli, 2019; Yi et al., 2020). In today's modern approaches to ASR, self-supervised learning is a learning method that does not require human annotated data. It is sometimes considered a form of unsupervised learning. In self-supervised learning, some supervised learning tasks are automatically created from unlabeled data. It aims to recover hidden or missing parts or features of input data given an invisible part of the same input. To solve such tasks, the machine is forced to learn strong representations that convey the meaning or structure of the data. These learned representations are expected to be useful in a variety of downstream tasks, usually after fine-tuning with a few tags.

The majority of the research on the Turkish language has adopted classical methods. However, considering the significant success of the Wav2Vec2 architecture in other languages and the lack of any studies for Turkish with the Wav2Vec2 architecture, the potential benefits of analyzing the study results have constituted the study's point of origin. In this study, we developed a transformer-based Turkish speech-to-text conversion model utilizing an XLSR-Wav2Vec2 model (Cross-Lingual Representation Learning for Speech Recognition) pre-trained with speech data from 53 languages and fine-tuned with data from the Mozilla Common Voice Turkish data set with transfer learning. This study has demonstrated the applicability of self-supervised pre-training techniques to acoustic data and open-source E2E ASR systems developed with the Wav2Vec 2.0 (Baevski, Zhou, Mohamed, & Auli, 2020) architecture by using a large corpus for multiple languages or a single language, and language-specific models were constructed through fine-tuning on the target language. We conducted a performance analysis for the model developed, identified its limitations, and discussed potential areas of usage.

In the following parts of this study, the current literature on ASR is given in Chapter 2, the details of the data set and model used in the study are provided in Chapter 3, the findings are presented in Chapter 4, and the discussion is in the last chapter.

2. LITERATURE REVIEW

Since deep learning techniques, which are frequently used for ASR today, are very data-dependent (Malik et al., 2021), the studies carried out within the scope of the literature are both very diverse and can reveal quite different results. In the literature, the amount of ASR research on the Turkish language is limited. Koruyan (2015) proposed a method for automatic caption generation with Google's Web Speech-to-Text API for live online broadcasts, discussing the effects of the speaker's distance to the microphone, ambient noise, and rapid or continuous speech on the model's performance. Yakar (2016) trained speech recognition models for Turkish using a Hidden Markov Model and analyzed their performance. When analyzed with the word error rate (WER), which is used to evaluate the performance of speech recognition systems, the model achieved a success rate of 17.4% with a limited test data set. In addition, to enhance the model's performance, they combined the phonemes predicted by the model and adopted a post-processing algorithm for Turkish word searches. Oyucu, Polat, & Sever (2020) employed the difference in Hirsch histograms to detect noise in speech, designating a threshold value and using sample frequencies to identify noise and silence. After the removal of noise and silence from the data set, the Kaldi library coded in C++ (Povey et al., 2011) was used to develop a speech recognition system. The speech recognition model trained with this approach had a 7.41% higher WER than the model trained without any pre-processing. In more recent studies, Oyucu & Polat (2023) suggested a Language Model (LM) optimization method for the Turkish language with limited resources. As a result of the findings they obtained, lower WER values were obtained in the ASRs applied with the proposed optimized LMs, and the performance of the ASR was improved. Mussakhojayeva, Dauletbek, Yeshpanov, & Varol (2023), in contrast, developed a multilingual ASR by considering ten languages, not only Turkish but also Azerbaijani, Bashkir, Chuvash, Kazakh, Kyrgyz, Sahaca, Tatar, Uyghur and Uzbek. Tombaloğlu & Erdem (2020) developed a Turkish ASR system based on the Deep Belief Network (DBN). As a result of the empirical findings, they found that the DBN-based ASR system outperformed the traditional Gaussian Mixture Method-based Hidden Markov Model. In addition, they supported the study findings that the DBN-based ASR showed superior results when compared to the existing studies in the literature.

Baevski et al. (2020) demonstrated how effective Wav2Vec 2.0 can be compared to other current approaches in speech recognition with a limited amount of labeled data. Schneider et al. (2019) had success with Wav2Vec with a WER of 2.43%, while Baevski et al. (2020) uses 100 times less labeled data with Wav2Vec 2.0, outperforming Wav2Vec2 in a subset of 100 hours. Liu, Yang, Chi, Hsu, & Lee (2020) presented a new speech representation learning approach in which bidirectional transformer encoders are pre-trained on a large amount of unlabeled speech. Chi et al. (2021) proposed Audio ALBERT, a lightweight version of the self-supervised learning speech representation model. Yi et al. (2020) focused on pre-trained Wav2Vec2 implementation using English speaking for the low-resource ASR task in many languages. According to the empirical findings of the study, a relative improvement of more than 20% was achieved in all six languages compared to previous studies. Pham, Waibel, & Niehues (2022) achieved a 44% improvement over fully supervised learning using Wav2Vec2 with a Common Voice corpus. Coto-Solano et al. (2022) tested Wav2Vec2 and alternative models on Cook Islands Maori, an indigenous language spoken by only about 22,000 people in the South Pacific. Their results show that Wav2Vec2 can yield promising results even for extremely low-resource languages such as Cook Islands Maori. Showrav (2022) demonstrated the success of automatic speech recognition with Wav2Vec2 for the Bengali language, similar to the results of Coto-Solano et al. (2022). Shahgir, Sayeed, & Zaman (2022) achieved better performance with Wav2Vec2 than Showrav (2022) for the Bengali language. Akhilesh et al. (2022) produced a simple and computationally cheaper ASR with Wav2Vec2 for the Tamil language. Olev & Alumae (2022) achieved a 6.9% WER for Estonian with an E2E Wav2Vec2 model. Jain et al. (2023), furthermore, obtained successful results by using Wav2Vec2 in children's speech recognition task, which is more difficult than adults. Wills, Bai, Tejedor-Garcia, Cucchiarini, & Strik (2023) went a step further and used the Wav2Vec2 model for a speech recognition task for non-native Dutch children. Although they found that alternative approaches produced better results, they showed that Wav2Vec2 can also be successful in Dutch non-native language children's speech. Hu et al. (2023) used adapted Wav2Vec2 in the task of automatic recognition of elder speech. In addition to these studies, Vaessen & Van Leeuwen's (2022) empirical findings supported that Wav2Vec2 showed successful results in the speaker recognition task besides the speech recognition task. In addition, the study findings revealed that pre-trained weights used to fine-tune the speech recognition task are also useful for fine-tuning speaker recognition.

As a result of the successful results of many studies, Wav2Vec 2.0 has become one of the most preferred neural-based models for ASR today (Vásquez-Correa et al., 2023). Therefore, in this study, a transformer-based ASR is designed using a fine-tuned XLSR-Wav2Vec2 model with data from the Mozilla Common Voice Turkish dataset.

3. MATERIALS AND METHOD

3.1. Data

The study adopted Mozilla Common Voice as the data set. The Common Voice data set is a multilingual transcribed speech collection for speech technology research and development. Common Voice is designed for ASR algorithm development but can also be utilized in other areas, such as language recognition. The Common Voice project employed crowdsourcing for both data collection and data validation, where voluntary users read out and recorded specific texts. The study adopted version 6.1, which contained 76 languages and was later expanded to encompass 85 languages by November 2021. So far, more than 80,000 volunteers have participated in the development of this data set, producing 13,905 and 11,192 hours of recorded and validated data, respectively. The Turkish language set comprises 44 hours of recorded and 39 hours of validated data. It offers one of the most voice-text matches among open-source databases in the field of speech recognition with respect to the number of both hours and languages. Table 1 shows the features of the data set compiled over time. There is a concentration of the voices of male volunteers and people aged 19-39, which creates a bias where the data set could attain the highest possible accuracy with male voices in this age range. However, the study employed the Common Voice data set as it is the most useful data set due to ease of access, being open-source and the high number of hours.

Table 1. Mozilla Common Voice corpus Turkish data set features (Özden, 2021)

Date	Records			Gender		
21.07.21	Validated Hours	Total Hours	Distinct Voices	Male	Female	N/A
	30	37	960	68%	6%	26%
Age Range						
19-29	30-39	40-49	50-59	60-69	Greater than 70	N/A
47%	17%	2%	4%	1%	0%	26%

The Common Voice project was launched in July 2017 with a focus on the English language, and then in June 2018, it was made available for other languages as well. Common Voice, by its nature, has a sustainable data collection pipeline and the collected data is checked by cross-validation. This control system is performed simply by voting whether the text read by one user has been read correctly by other users. Up to three participants listen to each audio clip. Text readings with at least 2 more than the number of correct votes are archived as unverified data, while other data are archived as verified data. Validated training, validation and testing data sets for each language are obtained using at least 2 positive validation voted data. While creating these data sets, repetitive texts are also removed from these data sets, preventing multiplexing.

3.2. Method

Wav2Vec, which is widely used in E2E ASR models, consists of multiple convolution and attention layers (Baevski, Schneider, & Auli, 2019). Convolution layers take the speech input to the algorithm as a sample and produce more compressed hidden representations, and the attention layers allow a more tailored analysis of the input. Strong context dependency modeling capabilities enable the model to make accurate choices during comparative training via inputs. The Wav2Vec2 training procedure consists of two stages: Initially, the model learns the acoustic representations during self-supervised pre-training by using a significant amount of unlabeled speech data. In the second stage, in supervised fine-tuning, the model is trained on labeled speech data to accurately predict sequences of raw audio graphs or characters.

The Wav2Vec2 model structure is presented in Figure 1. Raw speech waveform $x_i \in$ X is normalized to mean and unit variance, transferred to a feature encoder f: X → Z and converted into hidden representations $z_i \in$ Z. The feature encoders in Wav2Vec2 follow the original Wav2Vec and VQ-Wav2Vec design, however, the activation functions are substituted (Hendrycks & Gimpel, 2016).

The Wav2Vec2 network is transformer-based. In other words, it follows the BERT architecture, except for the changes in positional encoding. Hereby, fixed positional embeddings are replaced with relative positional embeddings to learn relative positional information. To that end, a convolution layer is added to the transformer network, similar to Mohamed, Okhonko, & Zettlemoyer (2019). There are two model configurations with different context network setups: Base and large-base models consist of 12 transformer blocks and have 768 hidden sizes and 8 attention heads, while, in large models, hidden sizes increased to 1,024, with 16 attention heads.

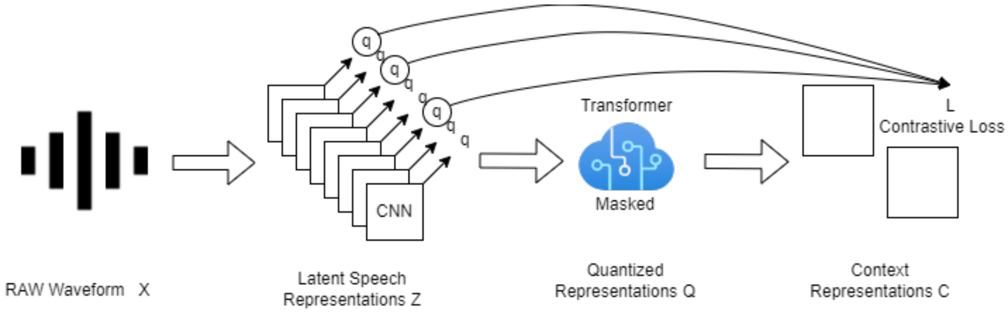

q
q
q
q
q
q
q
Transformer
Masked
CNN
L
Contrastive Loss
RAW Waveform X
Latent Speech Representations Z
Quantized Representations Q
Context Representations C

Figure 1. Wav2Vec 2.0 architecture (Baevski et al., 2020)

The second training stage for Wav2Vec2, fine-tuning for downstream tasks (ASR), begins with the random initiation of a classifier or the projection of a linear layer to C classes on the transformer network. The classes represent the vocabulary comprising characters and a word limit variable. The classifier is trained on labeled speech data and is optimized with a standard Connectionist Temporal Classification (CTC) loss (Graves, Fernández, Gomez, & Schmidhuber, 2006). Particularly, encoder weights are frozen during fine-tuning and therefore not updated. In addition, the quantification module is deactivated.

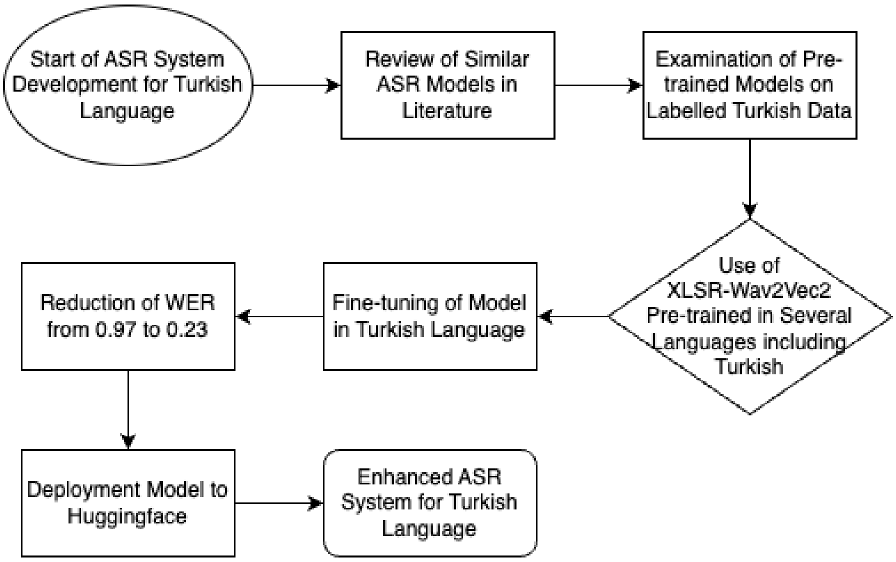

Start of ASR System Development for Turkish Language
Review of Similar ASR Models in Literature
Examination of Pre-trained Models on Labelled Turkish Data
Use of XLSR-Wav2Vec2 Pre-trained in Several Languages including Turkish
Fine-tuning of Model in Turkish Language
Reduction of WER from 0.97 to 0.23
Deployment Model to Huggingface
Enhanced ASR System for Turkish Language

Figure 2. Flow chart of the study

The flow chart in Figure 2 above summarizes the systematic process carried out in the development of the Turkish-specific Automatic Speech Recognition (ASR) system developed in this study. Commencing with a thorough review of extant ASR models in the literature, the research methodically progresses to assess the performance of these models with labeled Turkish data. Central to this endeavor is the employment and fine-tuning of the XLSR-Wav2Vec2 model, which has been pre-trained in multiple languages, including Turkish. This step is pivotal in refining the model's efficacy, as evidenced by the significant reduction in the Word Error Rate (WER) from 0.97 to 0.23. The culmination of this process is an enhanced ASR system, demonstrably more adept at Turkish language recognition, thereby marking a significant advancement in the field of speech recognition technology.

3.3. Word Error Rate

WER is a common measure of the performance of a speech recognition or machine translation system (Klakow & Peters, 2002). WER is a metric used to compute the difference between word-level predicted output and the current output. It can be computed as:

$$WER = (S + D + I) / N,$$

where S is the number of substitutions, D is the number of deletions, I is the number of insertions, and N is the number of measured words.

4. FINDINGS

ASR models convert speech to text; this means that a feature extractor that processes the speech signal into the model's input format, a feature vector, and a specifier that converts the model's output format to text is needed. For this purpose, a token module called Wav2Vec2 CTC Tokenizer was used to split the data into tokens for the Wav2Vec2 model, and a feature extractor called Wav2Vec2 Feature Extractor was used for feature extraction. The Mozilla Common Voice dataset does not only contain speech data and text equivalents. Other than these two features, all fields (e.g., accent, age, client ID, gender, etc.) were removed from the data set within the scope of this study. Afterwards, the characters of the text data in the data set were determined and arranged. In its final form, our dictionary has 40 unique characters. Then, lowercase conversions were made for all text data. The sampling rate needs to be adjusted so that speech signals can be handled by computers. It is pre-trained on the audio data of the XLSR-Wav2Vec2. Since most of these datasets are sampled at 16 kHz, Common Voice sampled at 48 kHz should be downgraded to 16 kHz for training. Therefore, our fine-tuning data was reduced to 16 kHz. Finally, "*Wav2Vec2Processor*" is used to process the data in the format expected by "*Wav2Vec2ForCTC*" for training. For this, the "*map*" function was used. First, by simply calling "*batch['audio']*", the audio data is loaded and resampled, secondly, "*input_values*" values are extracted from the loaded audio file. This only includes normalization, but for other speech models such as Log-Mel filters, this step may correspond to subtraction.

Python programing language was preferred for the entirety of the application, and programing was conducted in Google Colaboratory (Colab). A Google Colab notebook explaining every programing stage in detail that was kept for experimental reproducibility has been provided as an appendix to this manuscript. In contrast to most Natural Language Processing (NLP) models, XLSR-Wav2Vec2 is an open source, multilingual speech-to-text model with a much larger input length than output length. For instance, the output length of a sample with an input length of 50,000 is no more than 100. Due to the large input size, dynamic padding of training groups was preferred to improve model efficiency, which involves padding all the training samples only with the longest sample. Therefore, fine-tuning XLSR-Wav2Vec2 requires a special padding data collator.

Initially, a data collator must be defined. Contrary to common data collators, this data collator should handle "*input_values*" and "*labels*" differently. Subsequently, the XLSR-Wav2Vec2 checkpoint is loaded, and all training parameters are defined. Here, "*group_by_length*" promotes training efficiency via grouping training samples of similar input length as one batch, which can significantly expedite training by vigorously decreasing the total number of useless padding tokens passing through the model. "*learning_rate*" and "*weight_decay*" are heuristically set up until fine-tuning stabilizes. These parameters can be adopted in the present study as they are mostly dependent on the Common Voice data set. The parameters used for training this model are presented in Table 2.

Table 2. Model training hyperparameters

Hyperparameter	**Selected Value**
learning_rate	0.0005
train_batch_size	2
eval_batch_size	8
seed	42
distributed_type	multi-GPU
num_devices	8
total_train_batch_size	32
total_eval_batch_size	16
optimizer	*Adam with betas* = (0.9,0.999)
epsilon	1e-08
lr_scheduler_type	linear lr
scheduler_warmup_steps	500
num_epochs	30.0
mixed_precision_training	Native AMP

Table 3 shows the performance scores and overall performance charts of the model after training for 30 iterations. The trained model was shared by the authors as "wav2vec-tr-lite-AG" on the website huggingface.co, where the performance of the model can be tested through a web interface if desired. The model's WER was computed as 0.23 with the test data set. Individual inquiries yielded results that correlated with these rates. However, as discussed in the Data section of the manuscript, the speech recordings used in model training mainly comprised the voices of male volunteers and people aged 19-39, which might cause the model to produce the result given above in a limited framework.

Table 3. Model performance scores

Training Loss	Iterations	Steps	Validation Loss	Word Error Rate
0.439	3.70	400	1.366	0.970
0.377	7.40	800	0.492	0.537
0.230	11.11	1200	0.393	0.413
0.112	14.81	1600	0.327	0.291
0.147	18.51	2000	0.310	0.267
0.101	22.22	2400	0.259	0.232
0.070	25.92	2800	0.287	0.234
0.054	29.63	3200	0.270	0.231

Table 4 presents sample results of the ASR application developed for the Turkish language within the scope of this study. As can be seen, the audio inputs given as input to the application and the text outputs obtained are exemplified. The examples presented here are randomly selected for testing data. The results obtained appear to be quite successful. In addition, the word errors that occur are generally caused by word suffixes, as seen in Table 4.

Table 4. Model sample results of the ASR application developed for the Turkish language

Input (Turkish and English Translation)	Model Transcription
Bugün hava nasıl? (How is the weather today?)	Bugün hava nasıl?
Akşam yemeği için ne düşünüyorsun? (What are you thinking of making for dinner?)	Akşam yemek için ne düşünüyorsun?
Bu kitabı okudun mu? (Have you read this book?)	Bu kitabı okudun mu?
En yakın hastane nerede? (Where is the nearest hospital?)	En yakın hastane nerde?
Çay mı kahve mi tercih edersin? (Do you prefer tea or coffee?)	Çay mı kave mi tercih edersin?
Pazar günü buluşalım mı? (Shall we meet on Sunday?)	Pazar günü buluşalım mi?
Telefonum nerede acaba? (I wonder where my phone is?)	Telefonum nereye acaba?

Table 4 shows how well our ASR system can transcribe Turkish sentences that people might use in everyday life. These sentences are a bit longer to show the system's ability to handle more than just short phrases. The overall Word Error Rate (WER) of about 0.23 means that most of the time, the system gets the words right, but there are still some small mistakes. For example, it might miss a letter in a word or slightly change a word. These small errors show where we can still make the system better. Even with these mistakes, the table shows that our system does a good job of understanding and writing down Turkish sentences, which is a big step forward for this kind of technology. In addition, all source codes and applications regarding the developed model architecture and the results obtained are presented in the appendix in order to serve to vividly demonstrate the capabilities of the model.

5. DISCUSSION AND CONCLUSION

In the study, an ASR system for Turkish based on a small data set was developed, with Self-Supervised Learning as the primary instrument. Firstly, similar ASR models in the literature were reviewed, the performance of these pre-trained models on labeled Turkish data was examined, and our research was conducted to improve upon the previous results. Considering that ASR systems play an important role in the development of organizations, this study provides a resource for the Turkish language for this need. Especially in recent years, the increase in the use of these systems and the fact that organizations can do the managed work more efficiently and quickly thanks to these systems further emphasizes this importance. In many studies, it has been shown that ASR systems increase business efficiency and speed up business processes in organizations. Researchers such as Filippidou & Moussiades (2020) and Pallett (2003) discussed the increase in work efficiency and the acceleration of work processes with the use of ASR systems. In addition, thanks to ASR systems, organizations can also improve their customer service and customers can easily and quickly solve their problems. As Song et al. (2022) revealed in their study, customer satisfaction increases with the use of ASR systems. As a result, ASR systems play an important role in the development of organizations. Thus, work efficiency increases, and business processes accelerate. Therefore, in parallel with the increase in the use of ASR systems, this study makes significant contributions to the field of ASR for the Turkish language.

The study employed an unsupervised cross-lingual speech representation (XLSR-Wav2Vec2) pre-trained in several languages including Turkish. Subsequently, this pre-trained model was fine-tuned with transfer learning in the Turkish language. Fine-tuning in Turkish was implemented with Wav2Vec 2.0, and the initially higher WER of XLSR-Wav2Vec2 (0.97) was reduced to 0.23 after fine-tuning in the model parameters.

This study showed that Wav2Vec2 models with labeled data outperformed traditional ASR systems. Furthermore, Wav2Vec2 performs an E2E matching of a raw audio file with a sequence of graphs or words, thus eliminating the need to train separate components. However, it must be noted that these models are quite sizable and comprise approximately 317 million parameters, ruling out the possibility of real-time speech-to-text conversion, and require additional hardware, such as a GPU, to ensure an acceptable decoding speed. Therefore, we recommend future studies to adopt larger data sets for the model, reduce the model size or the number of model parameters, or compress Wav2Vec2 models to decrease model training time.

Peer Review: Externally peer-reviewed.

Author Contributions: Conception/Design of Study- D.E.T., K.K., C.Ç.; Data Acquisition- D.E.T.; Data Analysis/Interpretation- D.E.T., K.K., C.Ç.; Drafting Manuscript- D.E.T., K.K., C.Ç.; Critical Revision of Manuscript- D.E.T., K.K., C.Ç.; Final Approval and Accountability- D.E.T., K.K., C.Ç.; Supervision- K.K.

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

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

ORCID IDs of the authors

Davut Emre Taşar	0000-0002-7788-0478
Kutan Koruyan	0000-0002-3115-5676
Cihan Çılgın	0000-0002-8983-118X

REFERENCES

Akhilesh, A., Brinda, P., Keerthana, S., Gupta, D., & Vekkot, S. (2022). Tamil speech recognition using XLSR Wav2Vec2.0 & CTC algorithm. *2022 13th International Conference on Computing Communication and Networking Technologies (ICCCNT)*, 1-6. https://doi.org/10.1109/ICCCNT54827.2022.9984422

Amodei, D., Ananthanarayanan, S., Anubhai, R., Bai, J., Battenberg, E., Case, C., ... & Zhu, Z. (2016). *Deep speech 2: End-to-end speech recognition in English and Mandarin. ICML'16: Proceedings of the 33rd International Conference on International Conference on Machine Learning, Volume 48*, 173-182. https://dl.acm.org/doi/10.5555/3045390.3045410

Annam, S. V., Neelima, N., Parasa, N., & Chinamuttevi, D. (2023, March). Automated Home Life using IoT and Speech Recognition. *In 2023 International Conference on Innovative Data Communication Technologies and Application (ICIDCA)* (pp. 809-813). IEEE.

Baevski, A., Schneider, S., & Auli, M. (2019). vq-wav2vec: Self-supervised learning of discrete speech representations. arXiv. https://doi.org/10.48550/arXiv.1910.05453

Baevski, A., Zhou, Y., Mohamed, A., & Auli, M. (2020). wav2vec 2.0: A framework for self-supervised learning of speech representations. *Advances in neural information processing systems: 34th conference on neural information processing systems (NeurIPS 2020)*, https://proceedings.neurips.cc/paper_files/paper/2020

Benzeghiba, M., De Mori, R., Deroo, O., Dupont, S., Erbes, T., Jouvet, D., ... & Wellekens, C. (2007). Automatic speech recognition and speech variability: A review. *Speech communication, 49*(10-11), 763-786. https://doi.org/10.1016/j.specom.2007.02.006

Chi, P. H., Chung, P. H., Wu, T. H., Hsieh, C. C., Chen, Y. H., Li, S. W., & Lee, H. Y. (2021). Audio albert: A lite bert for self-supervised learning of audio representation. *2021 IEEE Spoken Language Technology Workshop (SLT)*, 344-350. https://doi.org/10.1109/SLT48900.2021.9383575

Chiu, C. C., Sainath, T. N., Wu, Y., Prabhavalkar, R., Nguyen, P., Chen, Z., ... & Bacchiani, M. (2018). State-of-the-art speech recognition with sequence-to-sequence models. *2018 IEEE international conference on acoustics, speech and signal processing (ICASSP)*, 4774-4778. https://doi.org/10.1109/ICASSP.2018.8462105

Coto-Solano, R., Nicholas, S. A., Datta, S., Quint, V., Wills, P., Powell, E. N., ... & Feldman, I. (2022). Development of automatic speech recognition for the documentation of Cook Islands Māori. *Proceedings of the Thirteenth Language Resources and Evaluation Conference*, 3872–3882. https://aclanthology.org/volumes/2022.lrec-1/

Cutajar, M., Gatt, E., Grech, I., Casha, O., & Micallef, J. (2013). Comparative study of automatic speech recognition techniques. *IET Signal Processing, 7*(1), 25-46. https://doi.org/10.1049/iet-spr.2012.0151

Danis, C., & Karat, J. (1995). Technology-driven design of speech recognition systems. D*IS '95: Proceedings of the 1st conference on Designing interactive systems: processes, practices, methods, & techniques*, 17-24. https://doi.org/10.1145/225434.225437

Dai, Y., & Wu, Z. (2023). Mobile-assisted pronunciation learning with feedback from peers and/or automatic speech recognition: A mixed-methods study. *Computer Assisted Language Learning*, 36(5-6), 861-884.

Filippidou, F., & Moussiades, L. (2020). A benchmarking of IBM, Google and Wit automatic speech recognition systems. *IFIP Advances in Information and Communication Technology*, 73-82. https://doi.org/10.1007/978-3-030-49161-1_7

Ghai, W., & Singh, N. (2012). Literature review on automatic speech recognition. *International Journal of Computer Applications, 41*(8), 42-50. http://dx.doi.org/10.5120/5565-7646

Graves, A., Fernández, S., Gomez, F., & Schmidhuber, J. (2006). Connectionist temporal classification: Labelling unsegmented sequence data with recurrent neural networks. *Proceedings of the 23rd international conference on Machine learning - ICML '06,* 369-376. http://dx.doi.org/10.1145/1143844.1143891

Hendrycks, D., & Gimpel, K. (2016). Gaussian error linear units (gelus). arXiv. https://doi.org/10.48550/arXiv.1606.08415

Hu, S., Xie, X., Jin, Z., Geng, M., Wang, Y., Cui, M., ... & Meng, H. (2023). Exploring self-supervised pre-trained ASR models for dysarthric and elderly speech recognition. ICASSP 2023 - 2023 *IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, 1-5. https://doi.org/10.1109/ICASSP49357.2023.10097275

Inaguma, H., Cho, J., Baskar, M. K., Kawahara, T., & Watanabe, S. (2019). Transfer learning of language-independent end-to-end ASR with language model fusion. *ICASSP 2019 - 2019 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, 6096-6100). https://doi.org/10.1109/ICASSP.2019.8682918

Jain, R., Barcovschi, A., Yiwere, M., Bigioi, D., Corcoran, P., & Cucu, H. (2023). A wav2vec2-based experimental study on self-supervised learning methods to improve child speech recognition. *IEEE Access, 11*, 46938-46948. https://doi.org/10.1109/ACCESS.2023.3275106

Klakow, D., & Peters, J. (2002). Testing the correlation of word error rate and perplexity. *Speech Communication, 38*(1-2), 19-28. https://doi.org/10.1016/S0167-6393(01)00041-3

Koruyan, K. (2015). Canlı internet yayınları için otomatik konuşma tanıma tekniği kullanılarak alt yazı oluşturulması [Generating captions using automatic speech recognition technique for live webcasts]. *Bilişim Teknolojileri Dergisi, 8*(2), 111-116. https://doi.org/10.17671/btd.31441

Kurian, C., & Balakrishnan, K. (2009). Speech recognition of Malayalam numbers. *2009 World Congress on Nature & Biologically Inspired Computing (NaBIC*), 1475-1479. https://doi.org/10.1109/NABIC.2009.5393692

Levis, J., & Suvorov, R. (2012). Automatic speech recognition. In *The encyclopedia of applied linguistics.* Retrieved from https://onlinelibrary.wiley.com

Liu, A. T., Yang, S. W., Chi, P. H., Hsu, P. C., & Lee, H. Y. (2020). Mockingjay: Unsupervised speech representation learning with deep bidirectional transformer encoders. *ICASSP 2020 - 2020 IEEE International Conference* on Acoustics, Speech and Signal Processing (ICASSP), 6419-6423. https://doi.org/10.1109/ICASSP40776.2020.9054458

Malik, M., Malik, M. K., Mehmood, K., & Makhdoom, I. (2021). Automatic speech recognition: A survey. *Multimedia Tools and Applications, 80*, 9411-9457. https://doi.org/10.1007/s11042-020-10073-7

Mohamed, A., Okhonko, D., & Zettlemoyer, L. (2019). Transformers with convolutional context for ASR. arXiv. https://doi.org/10.48550/arXiv.1904.11660

Mussakhojayeva, S., Dauletbek, K., Yeshpanov, R., & Varol, H. A. (2023). Multilingual speech recognition for Turkic languages. *Information, 14*(2), 74. https://doi.org/10.3390/info14020074

Negrão, M., & Domingues, P. (2021). SpeechToText: An open-source software for automatic detection and transcription of voice recordings in digital forensics. Forensic Science International: *Digital Investigation, 38*, 301223. https://doi.org/10.1016/j.fsidi.2021.301223

Olev, A., & Alumae, T. (2022). Estonian speech recognition and transcription editing service. *Baltic Journal of Modern Computing, 10*(3), 409-421. https://doi.org/10.22364/bjmc.2022.10.3.14

Oyucu, S., & Polat, H. (2023). A language model optimization method for Turkish automatic speech recognition system. *Politeknik Dergisi*, (Early Access). https://doi.org/10.2339/politeknik.1085512

Oyucu, S., Polat, H., & Sever, H. (2020). Sessizliğin kaldırılması ve konuşmanın parçalara ayrılması işleminin Türkçe otomatik konuşma tanıma üzerindeki etkisi [The effect of removal the silence and speech parsing processes on Turkish automatic speech recognition]. *Düzce Üniversitesi Bilim ve Teknoloji Dergisi, 8*(1), 334-346. https://doi.org/10.29130/dubited.560135

Özden, B. (2021, September 14). Common voice Türkçe'nin durumu [Web blog post]. Retrieved from https://discourse.mozilla.org/t/common-voice-turkcenin-durumu/85895

Padmanabhan, J., & Johnson Premkumar, M. J. (2015). Machine learning in automatic speech recognition: A survey. *IETE Technical Review, 32*(4), 240-251. https://doi.org/10.1080/02564602.2015.1010611

Pallett, D. S. (2003). A look at NIST's benchmark ASR tests: Past, present, and future. 2003 *IEEE Workshop on Automatic Speech Recognition and Understanding (IEEE Cat. No. 03EX721)*, 483-488. https://doi.org/10.1109/ASRU.2003.1318488

Pham, N. Q., Waibel, A., & Niehues, J. (2022). Adaptive multilingual speech recognition with pretrained models. arXiv. https://doi.org/10.48550/arXiv.2205.12304

Povey, D., Ghoshal, A., Boulianne, G., Burget, L., Glembek, O., Goel, N., . . . Vesely, K. (2011). The Kaldi speech recognition toolkit. *IEEE 2011 workshop on automatic speech recognition and understanding*, https://www.fit.vut.cz/research/publication/11196/.en

Pragati, B., Kolli, C., Jain, D., Sunethra, A. V., & Nagarathna, N. (2023, January). Evaluation of Customer Care Executives Using Speech

Emotion Recognition. *In Machine Learning, Image Processing, Network Security and Data Sciences: Select Proceedings of 3rd International Conference on MIND 2021* (pp. 187-198). Singapore: Springer Nature Singapore.

Schneider, S., Baevski, A., Collobert, R., & Auli, M. (2019). wav2vec: Unsupervised pre-training for speech recognition. arXiv. https://doi.org/10.48550/arXiv.1904.05862

Shahgir, H. A. Z. S., Sayeed, K. S., & Zaman, T. A. (2022). Applying wav2vec2 for speech recognition on Bengali common voices dataset. arXiv. https://doi.org/10.48550/arXiv.2209.06581

Shi, Z. (2021). *Intelligence science: Leading the age of intelligence.* Elsevier.

Showrav, T. T. (2022). An automatic speech recognition system for Bengali language based on wav2vec2 and transfer learning. arXiv. https://doi.org/10.48550/arXiv.2209.08119

Song, Y., Lian, R., Chen, Y., Jiang, D., Zhao, X., Tan, C., ... & Wong, R. C. W. (2022). A platform for deploying the TFE ecosystem of automatic speech recognition. *Proceedings of the 30th ACM International Conference on Multimedia*, 6952-6954. https://doi.org/10.1145/3503161.3547731

Tombaloğlu, B., & Erdem, H. (2020). Deep learning based automatic speech recognition for Turkish. *Sakarya University Journal of Science, 24*(4), 725-739. https://doi.org/10.16984/saufenbilder.711888

Tran, D. T., Truong, D. H., Le, H. S., & Huh, J. H. (2023). Mobile robot: automatic speech recognition application for automation and STEM education. *Soft Computing,* 1-17.

Vaessen, N., & Van Leeuwen, D. A. (2022). Fine-tuning wav2vec2 for speaker recognition. *ICASSP 2022 - 2022 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, 7967-7971. https://doi.org/10.1109/ICASSP43922.2022.9746952

Vásquez-Correa, J. C., & Álvarez Muniain, A. (2023). Novel speech recognition systems applied to forensics within child exploitation: Wav2vec2. 0 vs. whisper. *Sensors, 23*(4), 1843. https://doi.org/10.3390/s23041843

Wills, S., Bai, Y., Tejedor-Garcia, C., Cucchiarini, C., & Strik, H. (2023). Automatic speech recognition of non-native child speech for language learning applications. arXiv. https://doi.org/10.48550/arXiv.2306.16710

Xie, T. (2023). Artificial intelligence and automatic recognition application in B2C e-commerce platform consumer behavior recognition. *Soft Computing, 27*(11), 7627-7637.

Xiong, W., Wu, L., Alleva, F., Droppo, J., Huang, X., & Stolcke, A. (2018). The Microsoft 2017 conversational speech recognition system. *2018 IEEE international conference on acoustics, speech and signal processing (ICASSP).* 5934-5938. https://doi.org/10.1109/ICASSP.2018.8461870

Yakar, Ö. (2016). *Sözcük ve hece tabanlı konuşma tanıma sistemlerinin karşılaştırılması* (Master's thesis). Retrieved from https://tez.yok.gov.tr/UlusalTezMerkezi/

Yi, C., Wang, J., Cheng, N., Zhou, S., & Xu, B. (2020). Applying wav2vec2.0 to speech recognition in various low-resource languages. arXiv. https://doi.org/10.48550/arXiv.2012.12121

Yi, C., Wang, J., Cheng, N., Zhou, S., & Xu, B. (2021). Transfer ability of monolingual wav2vec2.0 for low-resource speech recognition. *2021 International Joint Conference on Neural Networks* (IJCNN), 1-6. https://doi.org/10.1109/IJCNN52387.2021.9533587

Yu, D., & Deng, L. (2016). *Automatic speech recognition (Vol. 1).* Berlin: Springer.

Zekveld, A. A., Kramer, S. E., Kessens, J. M., Vlaming, M. S., & Houtgast, T. (2009). The influence of age, hearing, and working memory on the speech comprehension benefit derived from an automatic speech recognition system. *Ear and Hearing, 30*(2), 262-272. https://doi.org/10.1097/aud.0b013e3181987063

How cite this article

Taşar, D. E., Koruyan, K., & Çılgın, C. (2024). Transformer-based Turkish automatic speech recognition. *Acta Infologica, 8*(1), 1-10. https://doi.org/10.26650/acin.1338604

Appendix:

Notebook URL: https://colab.research.google.com/drive/1xcfnBFdtT6HcW6MDLgImj7tJcYafJSJU

Acta Infologica
acin 2024, 8 (1), 11–22
DOI: 10.26650/acin.1349544

Research Article / Araştırma Makalesi

A Model Design Using Blockchain and Smart Contracts Against Cyberattacks in Smart Home Systems

Osman Güler[1]

[1] Tusaş Şehit Hakan Gülşen MTAL, Ankara, Türkiye

Corresponding author : Osman Güler
E-mail : h.osmanguler@gmail.com

ABSTRACT

The use of Internet of Things (IoT)-based smart home systems is rapidly becoming widespread today. The structure of smart devices and the inadequacy of security systems make these systems vulnerable to cyberattacks. Therefore, using a solid security mechanism is important for protecting personal data in smart home systems The main purpose of this study is to present a model that works on the blockchain and smart contract infrastructure to ensure the security of smart home systems against cyberattacks. This study is a design-based research that examines how blockchain and smart contracts can be integrated to increase smart home security. The blockchain technology used in the proposed model protects the integrity of data by providing a decentralized distributed ledger to eliminate possible attack vectors. Additionally, predefined security protocols are automatically executed thanks to the use of smart contracts, thus increasing the overall durability of the system. In this way, the proposed model effectively reduces security vulnerabilities in smart home systems, ensures the immutability of data, prevents unauthorized changes, and offers an effective security solution against possible cyberattacks. As a result, the proposed model can be said to be a robust, efficient security solution for IoT networks and smart home systems.

Keywords: Internet of Things, smart homes, smart contracts, blockchain, cybersecurity

Submitted	: 24.08.2023
Revision Requested	: 27.12.2023
Last Revision Received	: 02.03.2024
Accepted	: 04.03.2024
Published Online	: 25.06.2024

1. INTRODUCTION

With the developments in the field of information and communication technology (ICT), sensor-based technological devices that are able to communicate are being widely used in all areas of our lives. This technology is called the Internet of Things (IoT). IoT technology provides smart physical objects with Internet connection access and power to communicate (Panarello et al., 2018). These objects collect, store, and analyze data from sensors in order to increase efficiency, quality, and production in many areas such as IoT, smart factories, smart home systems, smart agriculture and irrigation, smart cities, smart logistics, and smart health (Gökrem & Bozuklu, 2016). Analyzing data needs to be done with smart and automated methods for performance and efficiency (Savaş et al., 2022b).

The biggest example of IoT technology being used in daily life is smart home systems. Smart homes are a type of housing integrated with IoT that provide comfort, security, convenience, and increased quality of life to their owners (Moniruzzaman et al., 2022). Smart homes refer to private houses that provide automatic smart services through various home devices such as home heating, home lighting, and white goods without human intervention that send and receive data to and from these devices in real-time (Park et al., 2019). Although smart homes provide great benefits to their owners, they are potentially at risk from malicious attacks, as the devices used are constantly connected to a network and offer vulnerable solutions to cyberattacks (Khan et al., 2020). Because IoT devices use a decentralized approach to network connectivity, using standard existing security techniques for inter-device communication is very complex (Alam, 2019). People have to take precautions to make their living spaces, homes, and workplaces safe and to keep their data secure in cyber environments (Savaş & Karataş, 2022a). To address these concerns, the proposal has emerged that using blockchain technology and smart contracts to secure the transactions between IoT devices is crucial for improving system security. This approach offers a robust layer of protection that addresses certain vulnerabilities by connecting to IoT devices using public or private keys and by providing secure identification and authentication rather than adhering to the rules of a central node or intermediary (Hassan et al., 2020). Blockchain technology is important for strengthening the security of communication between IoT devices. Blockchain is a decentralized distributed ledger that provides secure, transparent, and tamper-proof record keeping. Each transaction is encapsulated in a block of information that is cryptographically linked to the previous block, thus creating an unalterable chain of information.

Blockchain technology plays an important role in solidifying a security system, as it works as a distributed ledger and makes data difficult to change. Blockchain enables system components to communicate and share data securely, making difficulty for an attacker trying to modify, delete, copy, or deceive data in a blockchain-enabled system (Tekin et al., 2020; Kodym et al., 2020). Smart contracts are programs that set terms between two or more components and automatically execute certain actions when these conditions are met (Restuccia et al., 2019). When used in conjunction with blockchain technology, smart contracts govern the cooperation and interactions among security system components.

Blockchain and smart contracts are important technologies in the design of smart home security systems. The use of blockchain and smart contracts in smart home systems provides benefits in such areas as secure data sharing by ensuring data integrity, reducing the risk of data manipulation, and increasing system security. Thanks to these technologies, the smart home security has increased, providing homeowners with a safer and more comfortable life.

Various studies have been conducted on the use of blockchain technology for security in smart home systems. Dorri et al. (2016) proposed an architecture based on blockchain technology that includes a smart home, an overlay network, and cloud storage devices. The proposed architecture uses blockchain technology in networked device-to-device transactions and uses reliable distributed methods to ensure the decentralization of the architecture. Because bitcoin requires computational overhead, the proposed method is manageable for low-resource IoTs. In another study, Dorri et al. (2017) equipped their smart home system with a device called a *miner* that is always online and responsible for the communication among all devices in the system. A native private blockchain is used to provide secure access control for mining IoT devices and data. In the proposed system, the miner has a list of communicating devices and gives a key to these devices to ensure user control, thus securing inter-device communication. In addition, the blockchain creates a fixed time-ordered transaction history that can be linked to other layers to provide specific services. Although these blockchain-based approaches are suitable for providing decentralized security and privacy, they are not suitable for use with low-capacity IoT devices because they involve significant amounts of energy, latency, and computational overhead. Dand and Nguyen (2018) proposed an approach using blockchain technology called smart home-based IoT- blockchain. Their proposed architecture was used to create an experimental scenario among the user, the service provider, and the smart home using Ganache, Remix, and Web3.js. This approach proposes a blockchain technology that uses three types of smart contracts (i.e., access control contract [ACC], judge contract [JC], and registration contract [RC]) to ensure secure access control and IoT deployment. According to the test results, the proposed architecture identified and solved challenges in the smart home system such as data privacy, secure access control, and extension capability.

Singh et al. (2019) proposed an architecture for system security in smart homes that uses a multivariate correlation analysis technique to analyze network traffic and determine the relationship between different traffic characteristics. The proposed model consists of four components: the smart home layer, a blockchain network, cloud computing, and service layer. As a result of their tests, the proposed architecture was seen to provide smart homes with a network attack detection and response system. Arif et al. (2020) examined smart home architectures and security situations that use blockchain. They proposed a simple, secure smart home architecture with an improved blockchain called a consortium blockchain, which is a combination of public and private blockchains. The user's role in the blockchain process is eliminated; instead, IoT devices are defined as miners in the system. In this way, previously selected nodes now participate in block creation and consensus. This structure makes the proposed system unique compared to the current state of the art. Zhang and Yan (2021) proposed a blockchain-based smart home access control scheme using Hyperledger Fabric to provide access control with smart contracts. They used a hybrid access control model based on dynamic attribute-based access control and a static access control matrix. This system rejects access requests over the network from malicious attackers or devices not defined in the device list. In this way, user-initiated remote access control and access control between local devices are simultaneously guaranteed. Baucas et al. (2021) proposed a smart home design using proprietary blockchain technology and localization through trilateration based on the received signal strength indicator (RSSI) using the Raspberry Pi 3 model. This system consists of two components and focuses on low-quality access-level implementation. First, it identifies unrecognized devices trying to gain access using the blockchain; secondly, it uses localization to determine the source of the attack and the general location of the device and to obtain more information.

This study proposes a smart home system design using blockchain and smart contracts in order to provide secure communication in IoT and smart home systems and to prevent cyberattacks. The study uses a design-based research method. Design-based research refers to the process of developing a solution-oriented model that can be applied to an existing problem. This model proposal involves design research that examines how blockchain and smart contracts can be integrated to increase smart home security. The theoretical and practical contributions of this study are as follows:

- The study proposes a comprehensive model designed for smart home systems that integrates blockchain technology into the security framework of IoT devices.
- The study's inclusion of smart contracts as part of the model contributes to the theoretical understanding of automated security protocols.
- The proposed model offers a practical solution for strengthening the security of smart home systems against cyberthreats. The research provides practical insights into mitigating the vulnerabilities associated with constant device connectivity and possible cyberattacks through the application of blockchain and smart contract technology.
- As the basis of this research, the combination of blockchain and smart contracts provides a holistic security solution for IoT designed specifically for smart homes.

The second section of the study explains IoT technology, smart home systems, blockchain, and smart contract technologies. The third section presents information about the types of cyberattacks against IoT and the precautions that can be taken. The fourth section explains the proposed blockchain-based smart home automation approach. The fifth section evaluates IoT and Blockchain technologies and makes some suggestions.

2. CONCEPTUAL BACKGROUND AND METHOD

The IoT devices commonly used in smart home systems present a number of security challenges. The fact that these devices are always online, communicate with each other, and exchange data makes these systems vulnerable to cyberattacks. The interactions among blockchain, smart contracts, and IoT form the basis of the security system to be created against cyberattacks in smart home systems. Designed primarily as a decentralized distributed ledger for securing data transactions in IoT systems, blockchain offers a transformative approach. Blockchain is crucial for security, as IoT devices communicate and exchange data automatically, with every transaction being recorded in a cryptographically linked block creating an immutable chain. Integrating blockchain into IoT security preserves the integrity of the data by providing data transactions through a reliable transparent ledger. Every transaction is verified through authentication mechanisms on the network, increasing the security and reliability of the data. Blockchain's decentralized immutable nature reduces the risks associated with unauthorized access, thus ensuring the reliability of the data transmitted between IoT devices in smart home systems. Smart contracts are used to automate security protocols. These programmable contracts perform predetermined actions when certain conditions are met. Smart contracts running on a blockchain ensure that predetermined rules are automatically implemented. This allows security protocols to be automated and human intervention to be reduced.

2.1. The Internet of Things

With the widespread use of Internet technologies, users can access data and communicate whenever and wherever they want. IoT involves a set of systems that allow physical objects consisting of smart objects and sensors to automatically connect and communicate with each other without requiring personal intervention or manual data entry thanks to network connections and Internet access (Can et al., 2016; Gündüz & Daş, 2018). IoT systems consist of four main components: objects, communication, data, and users (Gündüz & Daş, 2018; Oral & Çakır, 2017). Fig. 1 shows the IoT components.

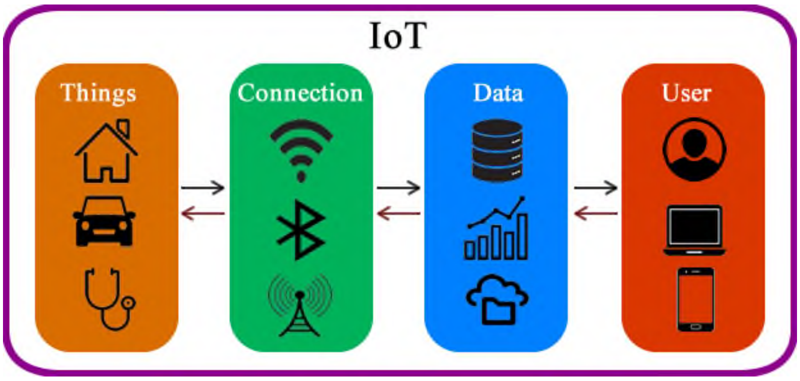


Figure 1. IoT Components

Objects consist of smart devices, sensors, and detectors used in IoT systems. Wired and wireless communication infrastructures such as Bluetooth, infrared, radio frequency identification (RFID), Zigbee, ethernet, and wi-fi are used for inter-object communication and sharing (Gökrem & Bozuklu, 2016). The data component consists of the data collected by the sensors, detectors, and objects in the environment, as well as the processors for these data. The user component consists of the users using the system and the interface programs used for access.

IoT technology is mostly used in daily life. For example, IoT technology can remotely control the ambient temperature, lighting, security systems, and even home appliances such as ovens and washing machines in smart home systems. IoT devices are also used in a variety of industries, including smart cities, smart healthcare, smart agriculture, and smart industrial applications.

2.2. Smart home systems

With the introduction of IoT technologies into daily life, smart technologies have begun being used in every area. One of these is in the home, where people spend their daily lives. Smart home systems are houses equipped with technologies that facilitate the life of the user, increases the comfort environment, provides energy efficiency, and provides a better-quality environment for people (İlkbahar et al., 2021). Electrical appliances, white goods, lighting and heating systems, audio and video systems, and security and camera systems in smart homes can communicate with each other and the user thanks to the wired or wireless network infrastructure (Kuncan & Çaça, 2019). The home owner can use the interface program to connect to smart home systems, control electrical appliances, adjust the temperature of the house, operate white goods, and see inside the house thanks to the cameras. When security threats such as intrusion, gas leakage, fire, or flooding occur, a message can be sent to the landlord and relevant institutions. With technological developments, smart homes now have artificial intelligence (AI) that can learn by themselves, monitor the user's daily life, and redevelop the program according to the user's needs (Avcı, 2022). Data on the user's daily activities at home are collected and used by such things as fuzzy logic, artificial neural networks, and machine learning algorithms. By processing the data with AI algorithms, the user's next behavior can be predicted (Güneş et al., 2019) and thus becomes able to produce unique solutions for home owners.

2.3. Blockchain

Blockchain technology is a decentralized distributed database management system that uses cryptographic techniques and does not require third-party verification (Novo, 2018). This technology involves a network formed of blocks containing encrypted transactions storing interconnected data. The blockchain is created by recording each block and adding it to the previous block; these blocks are linked chronologically and shared among all participants in the network

(Yıldız & Baştuğ, 2018). Once created, a block cannot be deleted or changed. Each block header in the blockchain contains the hash value of the previous block, a nonce (a temporary value used to generate the hash), a Merkle root (hash of all previous hashes), and timestamp information (Wu et al., 2020). Fig. 2 shows the blockchain structure.

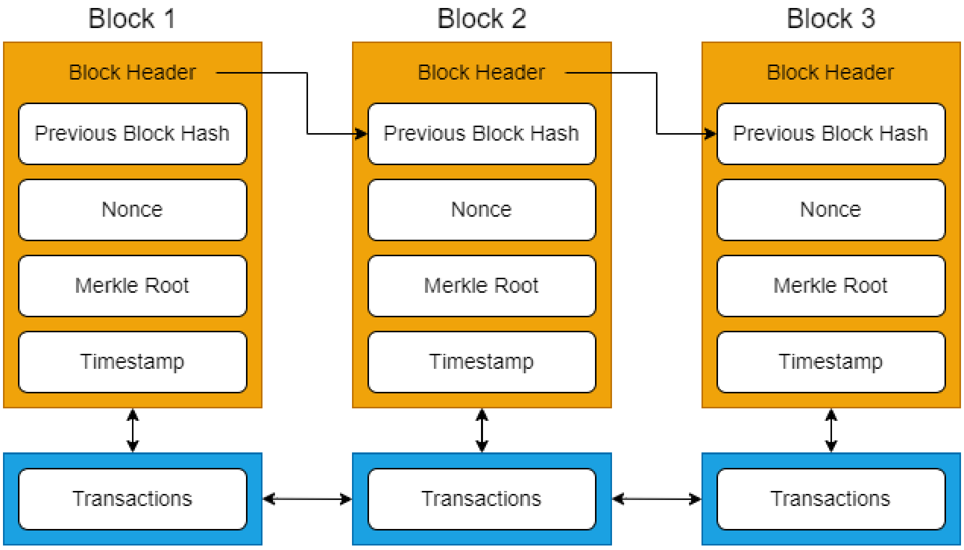


Figure 2. Blockchain architecture

Blockchain technology provides reliable, transparent, and sustainable database management due to the absence of a central authority. In addition, it has a structure that increases security thanks to the encrypted storage and distribution of data. In this way, it protects the accuracy and integrity of the data and prevents data tampering.

2.4. Smart contracts

Smart contracts are self-executing programmable contracts that run on blockchain technology. Thanks to these contracts, code pieces that can automatically run themselves, save data, keep values, and perform various calculations can be added to blocks when a certain situation occurs (Karaarslan & Akbaş, 2017). Smart contracts prevent the errors that occur in classical contracts and facilitate the checks and balances of the system and other contracts. At the same time, smart contracts are more reliable, work faster, have the ability to automatically execute, and save on operating costs compared to conventional contracts.

3. IoT CYBERATTACK DANGER

IoT technology connects billions of devices to the Internet. Many personal and business data are transmitted and stored on the Internet through these devices. As with any device with an Internet connection, IoT devices are a big target for cyberattacks. Information transmitted on IoT networks by unauthorized persons may be able to access and damage the network. Therefore, taking security measures and monitoring traffic continuously for cyberattacks are imperative (Dissanayake, 2021). This section discusses the types of cyberattacks that can be made against IoT devices, as well as the precautions that can be taken against cyberattacks.

3.1. DDoS Attacks

Distributed denial-of-service attacks (DDoS) are attacks that prevent a network from working by connecting many devices to that network. DDoS in an IoT network involves an attack that targets the availability of servers by flooding the communication channel by impersonating requests from the distributed IoT devices (Vishwakarma & Jain, 2020). By targeting IoT devices, these attacks can crash or render IoT applications or devices inoperable.

3.2. Man-in-the-Middle Attacks

One of the most popular attacks against IoT devices involves Man-in-the-Middle (MITM) attacks. MITM (or on-path) attacks are when an attacker gets between two nodes and interrupts communications, thus allowing the attacker to act as a proxy (Kuzlu & Güler, 2021). By appearing as a proxy in this way, the attacker can control incoming and outgoing messages.

3.3. Password Cracking Attacks

Most IoT devices are protected by default passwords that are simple and easy to guess. This facilitates attackers' ability to take control of devices. Attackers crack the current user password using dictionary attacks, which try possible letter and number combinations to guess user passwords, or by brute force attacks, which try all possible password combinations to find valid passwords (Abomhara & Køien, 2015).

3.4. Node Attacks

These attacks can damage a sensor node and involve the attacker physically replacing all or part of a node in order to access and modify sensitive information such as shared cryptographic keys (Islam & Aktheruzzaman, 2020). Attackers can disrupt the functionality of devices by preventing them from communicating with each other.

3.5. Social Engineering

While attackers do target IoT devices, they can also target humans. They do this by collecting information about the target person or institutions, with or without using technological tools (Irmak &Reis, 2018). One example of a social engineering attack involves the method of obtaining people's personal information or passwords by sending an email or message, which is known as phishing.

Many cyber security threats occur on the Internet, such as obtaining personal and corporate data, disclosing private information, and disabling commercial services (Savaş & Savaş, 2022). Because IoT devices in particular are always open to cyberattack, the security of these devices is very important. While using IoT devices, one needs to be aware of security vulnerabilities and take the necessary precautions. For this reason, the first of the measures that can be taken against cyberattacks is to change the default passwords of the devices and use strong passwords. In addition, regularly changing password and using different passwords on different devices are important. Monitoring for updates that eliminate devices' security vulnerabilities and installing manufacturer-provided updates are necessary. One's network system security should be increased, and firewall and antivirus software should be used. Devices that run on the network should be monitored regularly, and when unusual activity is observed, the devices should be turned off and their software and passwords should be updated immediately. Intrusion detection systems can be used to regularly monitor a network. One of the most important measures that can be taken against cyberattacks is to raise user and employee awareness about device security, cyberattacks, and the dangers of cyberattacks.

4. THE BLOCKCHAIN-BASED SMART HOME AUTOMATION APPROACH

This study aims to develop a model that uses smart contracts and blockchain technology to ensure data security against cyberattacks that might occur in smart home systems. The simple design diagram of the proposed system is shown in Fig. 3.

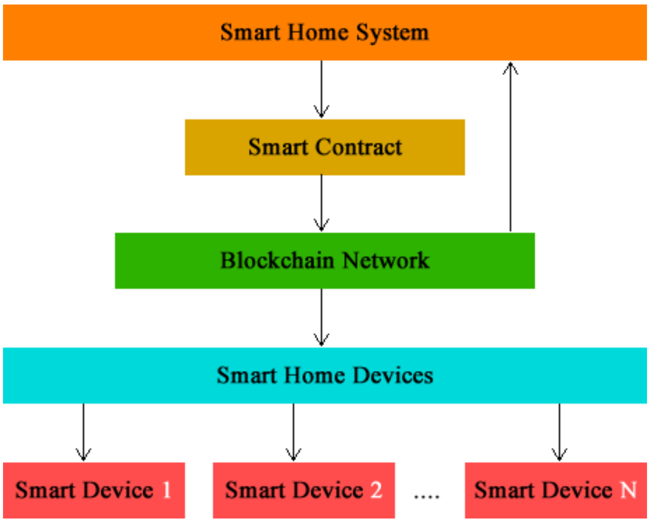


Figure 3. Suggested model design

Figure 3 shows a smart home model consisting of *N* smart devices (e.g., Smart Device 1, Smart Device 2, . . . , Smart Device N). These devices connect to a network system and communicate through a smart contract running on the blockchain network. In this way, the smart devices communicate with the smart contract by means of the blockchain network, exchanging data, receiving instructions, and providing real-time information. With this system, seamless automation and control of smart home devices is ensured based on predefined conditions determined by the smart contract. Fig. 4 presents the flow diagram of the smart home system using blockchain and smart contracts for the proposed system, with Algorithm 1 showing the details.

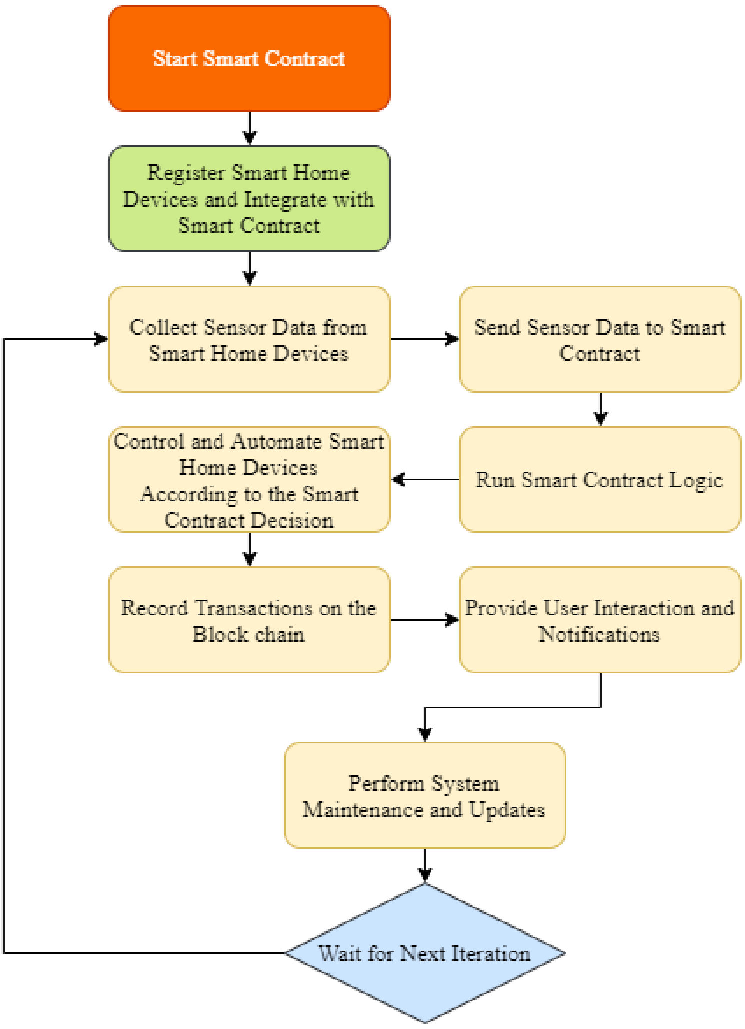

Start Smart Contract
Register Smart Home Devices and Integrate with Smart Contract
Collect Sensor Data from Smart Home Devices
Send Sensor Data to Smart Contract
Run Smart Contract Logic
Control and Automate Smart Home Devices According to the Smart Contract Decision
Record Transactions on the Block chain
Provide User Interaction and Notifications
Perform System Maintenance and Updates
Wait for Next Iteration

Figure 4. Flow diagram of Smart Contract and Blockchain structure

Algorithm 1: Blockchain and Smart Contract-Based Smart Home System

```
Start Smart Contract
Register Smart Home Devices and Integrate with Smart Contract
while(true)
  Collect Sensor Data from Smart Home Devices
  Send Sensor Data to Smart Contract
  Run Smart Contract Logic
  Control and Automate Smart Home Devices Based on Smart Contract Decision
  Record Transactions on the Blockchain
  Provide User Interaction and Notifications
  Perform System Maintenance and Updates
  Wait for Next Iteration
End While
```

- **Start Smart Contract:** A blockchain network is specially designed for the smart home system. The smart contract is deployed in the blockchain network by defining the necessary variables, functions, and conditions within the smart contract for controlling the smart devices.
- **Register Smart Home Devices and Integrate with Smart Contract:** Smart home devices are assigned unique identifiers using blockchain technology. They are registered and integrated into the smart contract, and the device information is stored in the smart contract.
- **Collect Sensor Data from Smart Home Devices:** Sensor data such as temperature, motion, light intensity, and humidity are continuously collected from the smart home devices.
- **Send Sensor Data to Smart Contract:** The sensor data collected from smart home devices is sent to the smart contract for processing and decision making.
- **Run Smart Contract Logic:** Incoming sensor data is evaluated against predefined conditions and rules in the smart contract, triggering appropriate actions or decisions.
- **Control and Automate Smart Home Devices According to the Smart Contract Decision:** Instructions or commands are sent by the smart contract to smart home devices for controlling operations such as turning lights on and off, adjusting thermostat settings, and activating security systems.
- **Record Transactions on the Blockchain:** The smart contract records the interactions, transactions, and decisions on the blockchain to ensure the transparency, immutability, and auditability of the system's activities.
- **Provide User Interaction and Notifications:** User interfaces (e.g., mobile apps, web interfaces) are provided for homeowners to interact with the smart home system. Homeowners are allowed to monitor and control smart home devices, view sensor data, and receive notifications/alerts. Access to devices is controlled using blockchain technology. In this way, only authorized users can access the devices, and in case of any unauthorized access, alarm systems are activated. Access rights to devices at home are managed using blockchain technology. A different access authorization can be defined for each user. For example, only certain devices may be authorized for children.
- **Perform System Maintenance and Updates:** The smart contract and blockchain network is regularly checked and updated to fix bugs and vulnerabilities or to implement system enhancements.
- The same operations are repeated as long as the system is running.

By following this algorithm, the blockchain and smart contract-based smart home system can effectively automate and coordinate the transactions of various devices while providing transparency, security, and privacy to the interactions between devices and homeowners. Some similarities and differences occur between the proposed approach and other studies in the literature. Some studies in the literature have used blockchain alone, with various approaches such as device communication occurring through a miner or through the use of a hybrid access control model. Other studies have used smart contract logic for access control and data privacy and to secure IoT deployment. The current study's proposed model uses blockchain and smart contracts for decentralization and security, similar to the studies in the literature. The model in this study uses blockchain technology for smart contract access control based on sensor data. In this model, the system maintenance and updates step plays an important role for security. Although not explicitly stated in the studies conducted in the literature, regular checks and updates have been implied as being necessary for ensuring system security. Figures of the architectures suggested in the literature are shown in Fig. 5.

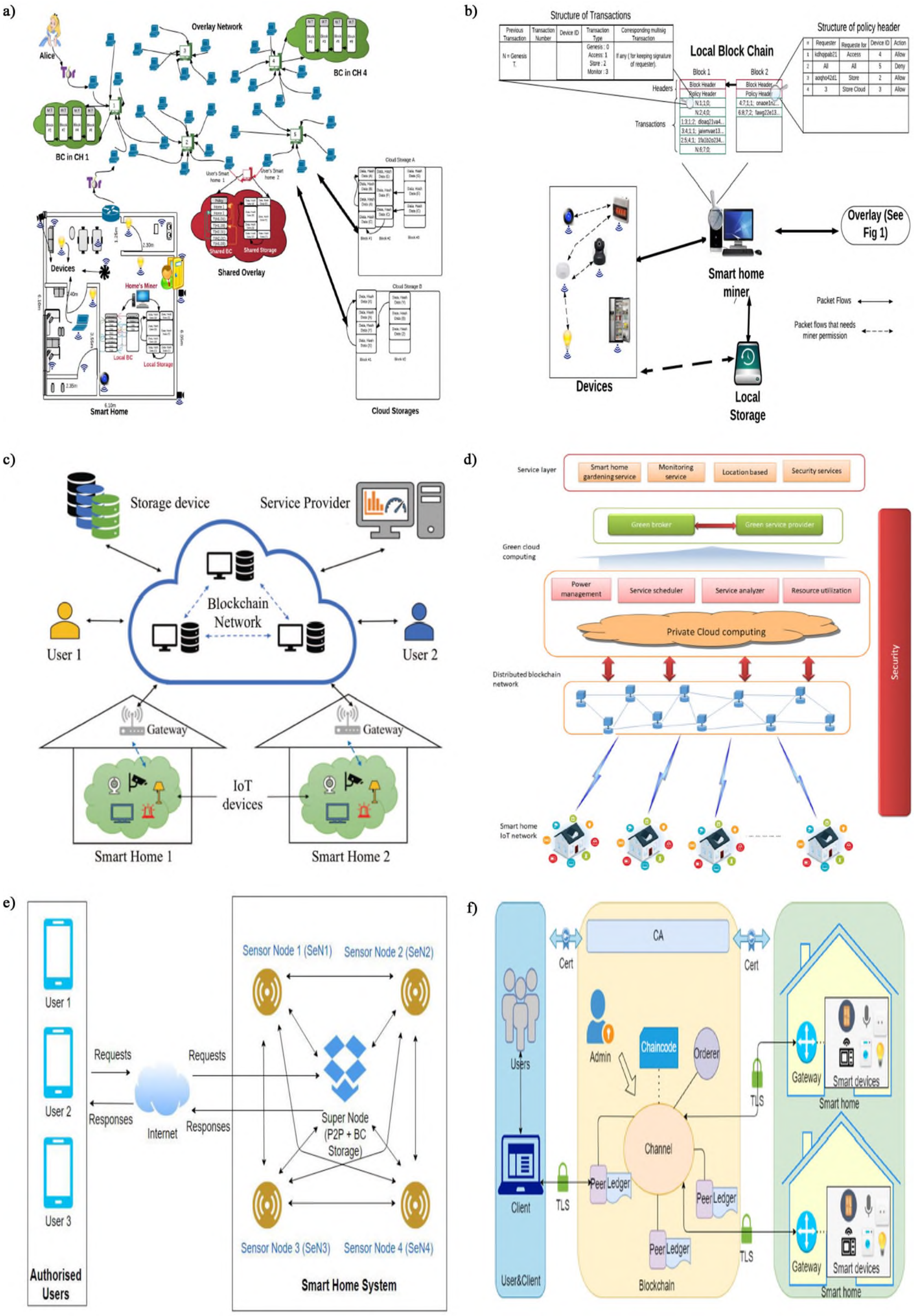


Figure 5. Previously proposed architectures: a) Dorri et al. (2016); b) Dorri et al. (2017); c) Dand & Nguyen (2018); d) Singh et al. (2019); e) Arif et al. (2020); and f) Zhang & Yan (2021)

This study' proposed model appears to be compatible with many concepts presented in the literature, such as the use of blockchain for security, smart contract logic for decision-making, and user interaction. The integration of smart devices with unique identifiers and the continuous monitoring of sensor data can be mentioned as examples of the contributions this study's proposed system presents to the literature.

Thanks to the proposed blockchain and smart contract-based smart home system, malicious access requests and cyberattacks are preventable. The smart contracts that make up the smart home system are resistant to data manipulation attacks because they are recorded on the blockchain in a transparent, unchangeable, and auditable manner. This means that monitoring and modifying any datum in the system are difficult. Blockchain technology protects against unauthorized access attacks by preventing unauthorized changes to the smart contracts and limiting unauthorized access to the smart home devices. Only authorized users can interact with the system through the smart contracts. The blockchain's transparent recording of the order and timestamping of the transactions prevents attackers from violating timestamps or changing the order of transactions. Smart contracts and the blockchain network undergo regular security updates and maintenance. This ensures that any known vulnerabilities are fixed and that the system remains up to date.

The defense mechanisms the proposed model will display against the cyberattacks described in Section 3 are stated below.

- Because data and transactions are spread over a distributed network using blockchain technology, no single point of the system can be focused on, thus effectively preventing DDoS attacks.
- Because security is provided using cryptographic keys and digital signatures, the system is resistant to password cracking attacks.
- The system is resistant to Man-in-the-Middle (on-path) attacks thanks to the transparent and encrypted recording of transactions using blockchain technology, as well as the use of secure encryption protocols in the communication between smart contracts.
- Because blockchain networks work with distributed and consensus algorithms (Proof of Work [PoW]/Proof of Stake [PoS]), the security of the network increases, and attacks against individual nodes become difficult.
- Social engineering attacks generally occur by manipulating the users. In blockchain-based smart home systems, security can be combined with technical measures such as cryptographic keys and digital signatures to ensure resistance to social engineering attacks.

This approach provides resistance to a series of cyberattacks on the system thanks to the features offered by blockchain and smart contracts. However, factors such as implementation, configuration errors, user errors, or vulnerabilities in the system design can affect the security of the system. Conducting ongoing security assessments and updating security measures are important.

5. CONCLUSION

Security in IoT systems is a field of study that has been emphasized in recent years. With the developments in technology, IoT devices have entered every aspect of life. These devices are connected to networks and the Internet and thus are vulnerable to cyberattacks. In order to protect the integrity and immutability of data in smart home networks, this study has proposed a model that uses blockchain and smart contracts for cyber security in smart home systems.

The proposed blockchain and smart contract-based smart home system has been designed to ensure the integrity and immutability of data in smart home networks. The study has also explained the algorithm and flow chart of the system. The proposed smart home system enables automation and coordination of various devices in the smart home environment while also significantly increasing security, privacy, and overall efficiency.

In comparison to prior research efforts in this field, the proposed model addresses the unique challenges posed by smart home systems. While previous studies have explored various blockchain-based approaches, such as those proposed by Dorri et al. (2016, 2017), Dand and Nguyen (2018), Singh et al. (2019), Arif et al. (2020), Zhang and Yan (2021), and Baucas et al. (2021), the current study's model seeks to strike a balance between energy efficiency and security, making it particularly suitable for low-capacity IoT devices.

In future research, the aim will be to test the model's performance against cyberattacks by actually implementing it. In this way, valuable information will be provided about the practical performance and robustness of the proposed blockchain and smart contract-based security framework in the context of smart home systems. The research is thus expected in this way to contribute to the evolving landscape of IoT security, especially in the field of smart home technology.

Peer Review: Externally peer-reviewed.
Conflict of Interest: The author has no conflict of interest to declare.
Grant Support: The author declared that this study has received no financial support.

ORCID IDs of the author

Osman Güler 0000-0003-3272-5973

REFERENCES

Abomhara, M., & Køien, G. M. (2015). Cyber security and the internet of things: vulnerabilities, threats, intruders and attacks. *Journal of Cyber Security and Mobility*, 65-88.

Alam, T. (2019). Blokzincir and its Role in the Internet of Things (IoT). International Journal of Scientific Research in Computer Science, Engineering and Information Technology. Vol 5(1). DOI: 10.32628/CSEIT195137

Arif, S., Khan, M. A., Rehman, S. U., Kabir, M. A., & Imran, M. (2020). Investigating smart home security: Is blockchain the answer?. *IEEE Access, 8*, 117802-117816.

Avcı, İ. (2022). Akıllı evlerde IoT teknolojileri ve siber güvenlik. Avrupa Bilim ve Teknoloji Dergisi, (34), 226- 233.

Baucas, M. J., Gadsden, S. A., & Spachos, P. (2021). IoT-based smart home device monitor using private blockchain technology and localization. *IEEE Networking Letters, 3*(2), 52-55.

Can, O., Sezer, E., Bursa, O., & Ünalir, M. O. (2016). Nesnelerin interneti ve güvenli bir sağlık bilgi modeli önerisi. In *4th International Symposium on Innovative Technologies in Engineering and Science (ISITES2016) 3-5 Nov 2016 Alanya/Antalya-Turkey*.

Dang, T. L. N., & Nguyen, M. S. (2018). An approach to data privacy in smart home using blockchain technology. In *2018 International Conference on Advanced Computing and Applications (ACOMP)* (pp. 58-64). IEEE.

Dissanayake, M. B. (2021). Feature Engineering for Cyber-attack detection in Internet of Things. *International Journal of Wireless and Microwave Technologies, 11*(6), 46-54.

Dorri, A., Kanhere, S. S., Jurdak, R., & Gauravaram, P. (2017). Blockchain for IoT security and privacy: The case study of a smart home. In *2017 IEEE international conference on pervasive computing and communications workshops* (PerCom workshops) (pp. 618-623). IEEE.

Dorri, A., Kanhere, S. S., & Jurdak, R. (2016). Blokzincir in internet of things: challenges and solutions. *arXiv preprint arXiv:1608.05187*.

Gökrem, L., & Bozuklu, M. (2016). Nesnelerin interneti: Yapılan çalışmalar ve ülkemizdeki mevcut durum. *Gaziosmanpaşa Bilimsel Araştırma Dergisi*, (13), 47-68.

Gündüz, M. Z., & Daş, R. (2018). Nesnelerin interneti: Gelişimi, bileşenleri ve uygulama alanları. *Pamukkale Üniversitesi Mühendislik Bilimleri Dergisi, 24*(2), 327-335.

Güneş, H., Bicakcı, S., Orta, E., & Akdaş, D. (2019). Akıllı evlerde kullanılan yapay zekâ teknikleri için simülasyon geliştirilmesi. *Gazi University Journal of Science Part C: Design and Technology, 7*(3), 554-563.

Hassan, M., Chen, J., Iftekhar, A., & Cui, X. (2020). Future of the internet of things emerging with blockchain and smart contracts. *International Journal of Advanced Computer Science and Applications*, 11(6).

Irmak, H., & Reis, Z. A. (2018). Sosyal Mühendislik Saldırılarına Karşı Web Tabanlı Bir Farkındalık Eğitimi. In 7th International Conference on *"Innovations in Learning for the Future": Digital Transformation in Education*, 108.

Islam, M. R., & Aktheruzzaman, K. M. (2020). An analysis of cybersecurity attacks against internet of things and security solutions. *Journal of Computer and Communications, 8*(4), 11-25.

İlkbahar, F., Şeyma, Ü., Karakaya, A. T., & Bayram, E. (2021). Akıllı Ev Sistemleri Üzerine Bir Model Önerisi. *AJIT-e: Academic Journal of Information Technology, 12*(45), 90-105.

Karaarslan, E., & Akbaş, M. F. (2017). Blokzinciri tabanli siber güvenlik sistemleri. *Uluslararası Bilgi Güvenliği Mühendisliği Dergisi, 3*(2), 16-21.

Khan, M.A., Abbas, S., Rehman, A., Saeed, Y., Zeb, A., Uddin, M.I., Nasser, N. & Ali, A. (2020). A machine learning approach for blockchain -based smart home networks security. *IEEE Network, 35*(3), pp.223-229.

Kodym, O., Kubáč, L. & Kavka, L. (2020). Risks associated with Logistics 4.0 and their minimization using Blockchain. *Open Engineering,* 10(1), 74-85.

Kuncan, M. & Çaça, Ö. (2019). Akıllı Ev Teknolojisi için Kablosuz Akıllı Kit. *Avrupa Bilim ve Teknoloji Dergisi*, (17), 271-282.

Kuzlu, M., Fair, C., & Guler, O. (2021). Role of artificial intelligence in the Internet of Things (IoT) cybersecurity. *Discover Internet of things, 1*, 1-14.

Moniruzzaman, M., Khezr, S., Yassine, A. & Benlamri, R. (2020). Blockchain for smart homes: Review of current trends and research challenges. *Computers & Electrical Engineering, 83*, p.106585.

Novo, O. (2018). Blokzincir meets IoT: An architecture for scalable access management in IoT. *IEEE internet of things journal, 5*(2), 1184-1195.

Oral, O., & Çakır, M. (2017). Nesnelerin interneti kavramı ve örnek bir prototipin oluşturulması. *Mehmet Akif Ersoy Üniversitesi Fen Bilimleri Enstitüsü Dergisi, 1*, 172-177.

Panarello, A., Tapas, N., Merlino, G., Longo, F. & Puliafito, A. (2018). Blockchain and iot integration: A systematic survey. *Sensors, 18*(8), p.2575.

Park, J. H., Salim, M. M., Jo, J. H., Sicato, J. C. S., Rathore, S., & Park, J. H. (2019). CIoT-Net: a scalable cognitive IoT based smart city network architecture. *Human-centric Computing and Information Sciences*, 9(1), 1-20.

Restuccia, F., Kanhere, S.D., Melodia, T. & Das, S.K. (2019). Blockchain for the internet of things: Present and future. *arXiv preprint arXiv:1903.07448*.

Savaş, S. & Karataş, S. (2022a). Cyber governance studies in ensuring cybersecurity: an overview of cybersecurity governance. *International Cybersecurity Law Review, 3*(1), pp.7-34. https://doi.org/10.1365/s43439-021-00045-4

Savaş, S., Duraklar, K., Çınar, O.A., Koç, M., Turan, A., Uslu, U., Doğanay, A.S., Özceyhan, O.G., Destan, M.Y. & Duşbudak, H. (2022b). Güneş Enerjisi Sistemlerinde Yenilikçi ve Akıllı Bakım Onarım. *Journal of Information Systems and Management Research, 4*(2), pp.35-49.

Savaş, T., & Savaş, S. (2022). Tekdüzen kaynak bulucu yoluyla kimlik avı tespiti için makine öğrenmesi algoritmalarının özellik tabanlı performans karşılaştırması. *Politeknik Dergisi*, 25(3): 1261-1270.

Singh, S., Ra, I. H., Meng, W., Kaur, M., & Cho, G. H. (2019). SH-BlockCC: A secure and efficient Internet of things smart home architecture based on cloud computing and blockchain technology. *International Journal of Distributed Sensor Networks, 15*(4), 1550147719844159.

Tekin, M., Öztürk, D. & Bahar, İ. (2020). Akıllı lojistik faaliyetlerinde blokzincir teknolojisi. *Kent Akademisi, 13*(3), pp.570-583.

Vishwakarma, R., & Jain, A. K. (2020). A survey of DDoS attacking techniques and defence mechanisms in the IoT network. *Telecommunication systems, 73*(1), 3-25.

Wu, M., Wang, K., Cai, X., Guo, S., Guo, M., & Rong, C. (2019). A comprehensive survey of blockchain: From theory to IoT applications and beyond. *IEEE Internet of Things Journal, 6*(5), 8114-8154.

Yıldız, R. Ö., & Baştuğ, S. (2018). Blok zincir teknolojisi kapsamında elektronik konşimento.(ss. 7-12). IV. Uluslararası Kafkasya–Orta Asya Dış Ticaret ve Lojistik Kongresi, Düzenleyen Adnan Menderes Üniversitesi. Aydın, 7 (8).

Zhang, W., & Yan, H. (2021). A blockchain -based access control scheme for smart home. In *Journal of Physics: Conference Series* (Vol. 1971, No. 1, p. 012049). IOP Publishing.

How cite this article

Güler, O. (2024). A model design using blockchain and smart contracts against cyberattacks in smart home systems.*Acta Infologica, 8*(1), 11-22. https://doi.org/10.26650/acin.1349544

Acta Infologica
acin 2024, 8 (1), 23–33
DOI: 10.26650/acin.1418834

Research Article / Araştırma Makalesi

Sentiment Analysis on GPT-4 with Comparative Models Using Twitter Data*

Mustafa Özel[1], Özlem Çetinkaya Bozkurt[2]

[1]Burdur Mehmet Akif Ersoy University, Social Sciences Institute, Burdur, Türkiye
[2]Burdur Mehmet Akif Ersoy University, Bucak Faculty of Business Administration, Department of Business Administration, Burdur, Türkiye

Corresponding author : Mustafa Özel
E-mail : mozel@mehmetakif.edu.tr

*Bu çalışma 10. Uluslararası Yönetim Bilişim Konferansında (IMISC 2023) sözlü olarak sunulan bildirinin geliştirilmiş halidir.

ABSTRACT

Every day, people from all over the world use Twitter to talk about many different topics using hashtags. Since ChatGPT was launched, researchers have been studying how people perceive it in society. This research aims to find out what Turkish Twitter users think about OpenAI's latest AI model called Generative Pre-trained Transformer 4 (GPT-4). The quantitative data used in this study consist of hashtags on the topic of GPT-4 and involve 2,978 tweets on this topic that were shared on Twitter between March 14-April 9, 2023. The study uses TextBlob sentiment scores to classify the tweets and support vector machines, logistic regression, XGBoost, and random forest algorithms to classify the sentiment of the dataset. The results from the logistic regression, XGBoost, and support vector methods are in close alignment. All parameter findings indicate dependable machine learning, emphasizing the models' success in classifying tweet sentiment.

Keywords: Sentiment analysis, social media, Twitter, X, natural language processing

Submitted	: 12.01.2024
Revision Requested	: 20.03.2024
Last Revision Received	: 15.04.2024
Accepted	: 16.04.2024
Published Online	: 03.06.2024

1. INTRODUCTION

Social media has become an integral part of life as it allows people to communicate, share, learn, and express themselves over common interests in real time. Just like food, water, and home, social media has become a basic need (Dandekar et al., 2018, p. 882). People often share their intentions, troubles, solutions, and moods on social media. Numerous users actively use social media platforms, and through these platforms, users with various opinions express their opinions and thoughts via text (Rahman et al., 2023, p. 1069). Many social media platforms exist and are available today, such as Twitter, Facebook, WhatsApp, Instagram, and TikTok. Among these social media platforms, Twitter is one of the most effective for getting ideas about issues and events. Many people share their perspectives on different issues on Twitter, making social media platforms such as Twitter an open source of data.

Twitter users around the world discuss various topics through hashtags every day. Recently, the most interesting and intriguing of these hashtags has been ChatGPT, with GPT meaning Generative Pre-trained Transformer. Since the introduction of ChatGPT, researchers have started to investigate the public's attitude toward it. Numerous studies have discussed the broad social implications of ChatGPT (Abdullah et al., 2022) and its domain-specific potentials (Munggaran et al., 2023; Botchu & Iyengar, 2023). These studies have predominantly utilized research methods from the social sciences, including interviews, user experience, and expert-based perspectives. Other studies in the literature that are more relevant to the current study have used different computational techniques to explore the public sentiment of ChatGPT using social media data. As an illustration, Liu et al. (2023) endeavored to furnish a thorough examination of extant studies on ChatGPT and its prospective implementations in diverse domains. In pursuit of this objective, they carried out an extensive examination of ChatGPT-related papers in the JarXiv repository and attempted to provide understanding into ChatGPT's skills, potential ramifications, ethical considerations, and prospects for future progress in this domain. The results indicate a notable and increasing fascination with ChatGPT/GPT-4 investigations across several disciplines such as education, history, mathematics, medicine, and physics, that are primarily centered around direct applications of natural language processing. Feng et al. (2023) conducted a study on Twitter and Reddit users to investigate the potential for ChatGPT in code generation and attitudes toward ChatGPT. The study revealed fear as the prevailing emotion linked to ChatGPT code generation, surpassing such emotions as happiness, anger, surprise, and sadness. Lampropoulos et al. (2023) used Twitter data to report on the perspectives, attitudes, emotions, and discourses surrounding the use of ChatGPT for general and educational purposes. Their results demonstrated the broad applicability of artificial intelligence (AI) tools, as well as the versatility of ChatGPT. Li et al.'s (2023) study analyzed Twitter data to ascertain the primary apprehensions regarding the use of ChatGPT in the field of education and found that, while a generally positive sentiment was present, concerns also occurred in five areas: learning outcomes and skill development, academic integrity, skill limitations, political and social impacts, and workforce challenges. Meanwhile, Keskin's (2023) study analyzed news publications to identify the main topics, focusing on how ChatGPT is addressed in Türkiye's Internet agenda. As a result of the analysis, prominent themes were identified such as education, new developments in ChatGPT, business life, information gathering, coding and development, the IT sector, daily life, investment consultancy, and creative content production. Korkmaz et al. (2023) performed a sentiment analysis on Twitter posts related to ChatGPT to thoroughly assess the emotions and opinions expressed over the initial two months following the announcement of ChatGPT. The results showed the majority of users who'd used ChatGPT for the first time to have found the experience successful and to be satisfied with ChatGPT; however, it also aroused negative emotions such as fear and anxiety in some users.

Unlike other research, this study collects and translates into English Turkish tweets about GPT-4, an artificial intelligence application that is on the agenda all over the world before performing the sentiment analyses in an attempt to investigate the general public attitude toward ChatGPT in Türkiye. The study then categorizes these public attitudes according to their sentiment scores using machine learning techniques. In line with this, it has collected tweets sent between March 14-April 9, 2023 and classified the preprocessed tweets based on their sentiment scores using the TextBlob dictionary. The study also subjected the dataset that had been classified according to emotion scores to emotion classification using the logistic regression, support vector machines, and random forest machine learning algorithms. Finally, the study has used Python programming language for data preprocessing and other operations. The findings obtained from this exploratory study can be useful for both the public interested in ChatGPT as well as the developers of ChatGPT-related technology.

Table 1. Some of the Studies on Chat-GPT

Author(s)	Title of the Study	Year
Abdullah et al.	Fundamentals, Applications and Social Impacts	2022
Munggaran et al.	Sentiment Analysis of Twitter Users' Opinion Data Regarding the Use of ChatGPT in Education	2023
Botchu & Iyengar	Will ChatGPT Drive Radiology in the Future?	2023
Liu et al.	Summary of Chat-GPT/GPT-4 Research and Perspective Towards the Future of Large Language Models	2023
Feng et al.	Investigating Code Generation Performance of ChatGPT with Crowdsourcing Social Data	2023
Lampropoulos et al.	A Social Media Data Analysis of General and Educational Use of ChatGPT: Understanding Emotional Educators	2023
Li et al.	ChatGPT in Education: A Discourse Analysis of Worries and Concerns on Social Media	2023
Keskin, E.	Yapay zekâ sohbet robotu ChatGPT ve Türkiye internet gündeminde oluşturduğu temalar	2023
Korkmaz et al.	Analysing the User's Sentiments of ChatGPT Using Twitter Data	2023

2. SOCIAL MEDIA AND DATA ANALYSIS

2.1. Twitter and Chat-GPT

Twitter ranks highly among social media platforms for staying informed about current events and trending topics. Many people share their perspectives on different topics on Twitter, making social media platforms such as Twitter an open source of data. Three basic symbols are used in this communication that are realized through a common universal tweet system. Using the @ symbol followed by a Twitter account name tags (mentions) the person or organization being tagged. Retweet (i.e., RT) is the sharing of an interesting tweet by another Twitter user. A hashtag (i.e., #) is a largely user-generated mechanism for labelling and collating messages (i.e., tweets) on a particular topic. Users who want to send messages add short words, sentences, or abbreviations to their messages, preceded by # to indicate that their messages address certain themes (Aladwani, 2015, p. 16; Bruns & Burgess, 2011, p. 2; Suh et al., 2010, p. 177).

Launched on November 30, 2022, Chat-GPT is an interactive chatbot developed by the AI company OpenAI (Kirmani, 2022, p. 574). Chat-GPT understands what is requested by the user, interprets it, and produces appropriate responses in almost natural human language. Besides practical applications, Chat-GPT's ability to successfully perform complex tasks has made it a major innovation in the fields of natural language processing and AI (Lund & Wang, 2023, p. 26). Finally, OpenAI has introduced GPT-4, the latest member of the GPT family. Many users have praised GPT-4's most recent improvements and distinctive capabilities (Koubaa, 2023, p. 1). Unlike its previous version, GPT-4 is a multimodal and large-scale model that also accepts images as input and can produce text output. GPT-4 has outperformed many traditional natural language processing (NLP) tests, as well as older large language models and more advanced systems (OpenAI, 2023; Aydın & Karaarslan, 2023, p. 4).

2.2. Machine learning

Machine learning (ML) is the realization of knowledge transfer similar to that of humans. In machine learning, a training model is created using data, and the decision-making quality of the system is improved. This learning method is use to try and ensure that the system makes successful predictions or successful classifications against similar data in the future (Doğan, 2022, p. 914).

2.2.1. Logistic Regression

Logistic regression analyzes data to estimate the probability of a certain outcome (dependent variable) based on its relationship with other factors (independent variables; Bircan, 2004, p. 186). It is an algorithm used to solve both regression and classification problems with both numerical and textual data. Three methods are found for applying the logistic regression classification algorithm in real life: binary (binomial), ordinal, and multinominal, with the multiclass (multinomial) approach allowing the dependent variable to have three or more different values (Ulaş & Karabay, 2020, p. 271).

2.2.2. *Support Vector Machines*

Support vector machines are usually divided into linear and non-linear problems. The purpose of using support vector machines in linear problems is to separate the features of the classes as far as possible from each other, with a hyperplane passing through the features (Metlek & Kayaalp, 2020, p. 2217). Nonlinear classifiers are used in non-linear situations. In such cases, the dataset is shifted from a two-dimensional to a three-dimensional space, and mapping is performed. The non-linear mapping approach moves the two-dimensional dataset to the three-dimensional feature space, enabling the linear separation of the dataset (Ayhan & Erdoğmuş, 2014, p. 185).

2.2.3. *Random Forest*

Random forest is an algorithm that creates more than one decision tree during the classification process and thus increases the classification rate. Randomly selected decision trees together form the decision forest (Aydın, 2018, p. 172). Random forest classifier is a prediction tool that uses the average to improve prediction accuracy and prevent overfitting by applying a set of decision tree classifiers to different subsamples of the dataset. The subsample size is always equal to the original input sample size (Veranyurt et al., 2020, p. 279).

2.2.4. *XGBoost*

XGBoost is a decision tree-based machine learning algorithm and a supervised learning algorithm used for classification and regression, with high-value results being obtainable in the shortest amount of time with less resource consumption. XGBoost operates similarly to the random forest algorithm. While bagging is applied in the random forest algorithm, boosting is applied in the XGBoost algorithm (Turan & Polat, 2024, p. 99; Tekin & Yaman, 2023, p. 156).

2.3. Sentiment Analysis

Sentiment analysis is the process of collecting and analyzing people's opinions, thoughts, and impressions on various topics (Wankhade et al., 2022, p. 5731). The beginning and rapid growth of the field coincides with the beginning of social media on the Web, forum discussions, blogs, microblogs, microblogging, Twitter, and social networks. Since the early 2000s, sentiment analysis has emerged as a highly dynamic field of study within the NLP domain. Sentiment analysis has been disseminated beyond the field of computer science to the realm of management sciences and various other disciplines, including marketing, finance, and health. This is because ideas are at the heart of almost all human activities and significantly influence human behavior. The values one believe in, one's reality, and the choices one makes depend to a large extent on how others view and evaluate the world. Therefore, one often considers the opinions of others when making a decision. This is true not only for humans but also for organizations (Zhang et al., 2018, p. 1). Due to the daily increase in user-generated data on the Web, this content needs to be analyzed in order to know users' opinions, thus increasing the demand for sentiment analysis research (Agarwal & Mittal, 2013, p. 14).

Sentiment analysis falls into three main categories: dictionary-based, ML, and hybrid approaches. Dictionary-based methods leverage pre-existing sentiment lexicons for unsupervised classification, while ML methods rely on training data labeled for supervised learning. As the name suggests, hybrid approaches combine elements of both dictionary and ML techniques (Biltawi et al., 2016, p. 339).

3. MATERIALS AND METHODS

This section shows the model of the study (see Figure 1) and explains the study steps, such as obtaining the dataset for sentiment analysis on Twitter using ML, data preprocessing steps, data labeling, data separation and modeling, model comparison s, and performance measurements

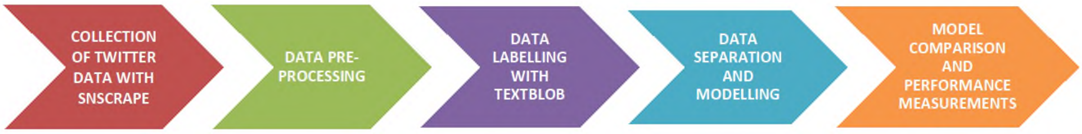
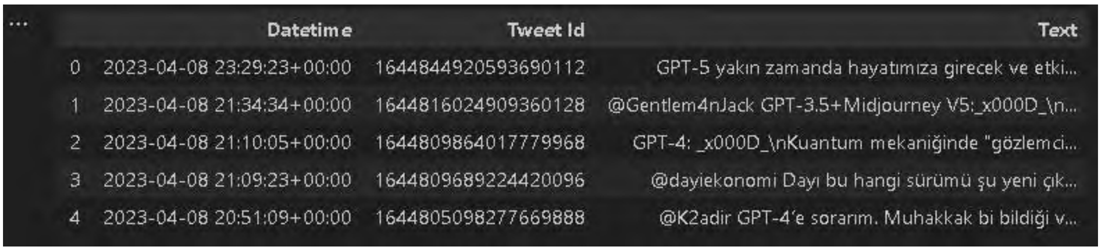

COLLECTION OF TWITTER DATA WITH SNSCRAPE
DATA PRE-PROCESSING
DATA LABELLING WITH TEXTBLOB
DATA SEPARATION AND MODELLING
MODEL COMPARISON AND PERFORMANCE MEASUREMENTS

Figure 1. Application steps.

3.1. Dataset

The study benefits from Turkish tweets containing the keyword "GPT-4" that were shared on Twitter between March 14-April 9, 2023. A total of 3,041 tweets were accessed using the Snscrape Library, and a Python library was used to collect data from Twitter. The obtained dataset contains Datetime, Tweet ID, and Text information. Figure 2 shows examples of the dataset that is used.



```
...                 Datetime             Tweet Id                                                 Text
     0  2023-04-08 23:29:23+00:00  1644844920593690112         GPT-5 yakın zamanda hayatımıza girecek ve etki...
     1  2023-04-08 21:34:34+00:00  1644816024909360128  @Gentlem4nJack GPT-3.5+Midjourney V5:_x000D_\n...
     2  2023-04-08 21:10:05+00:00  1644809864017779968     GPT-4: _x000D_\nKuantum mekaniğinde "gözlemci...
     3  2023-04-08 21:09:23+00:00  1644809689224420096        @dayiekonomi Dayı bu hangi sürümü şu yeni çık...
     4  2023-04-08 20:51:09+00:00  1644805098277669888         @K2adir GPT-4'e sorarım. Muhakkak bi bildiği v...
```

Figure 2. Tweet dataset example.

3.2. Data Preprocessing

The next preprocessing step removes the hashtags, mentions, URLs, and emojis. To remove hashtags (words starting with #), mentions (words starting with @), emojis, and URLs, the study uses the Python Regular Expression Syntax (RE) module, which is a powerful tool for finding, matching, replacing, or removing specific patterns in text. These operations are performed using the function *re.sub().*

After removing hashtags, mentions, URLs, and emojis, the Turkish tweets are converted into normal text and then translated into English in the next step. For this process, the study uses the deep_translator library, which is used for simple translations between different languages. The next step converts uppercase letters in the tweets into lowercase letters. Afterward, the punctuation marks are removed. The conversion from uppercase to lowercase is done with the previously used *lower()* method. To remove punctuation marks, the *string. punctuation* command in the Python string module is used to remove punctuation marks or replace them with specific punctuation marks.

Textual data consist of many redundant words (e.g., this, that, of, or, and, with, the) that are not context-related and will not help in classifying the textual contexts of a tweet but do help humans understand it correctly. These are called stop words (Verma et al., 2019). This next stage filters out the frequently used stop words in a language using the *stop_words* tool, after which the tokenization process is applied. Tokenization involves segmenting the text according to its features, such as spaces and punctuation marks (Küçükkartal, 2020, p. 11; Kahya, 2021, p. 12). Stemming (finding the root) is performed next. Stemming and lemmatization are different methods. Stemming is applied to remove inflectional prefixes or suffixes from words. Words with the same meaning and spelling are considered to be different words according to a prefix or suffix. Stemming is used to avoid this. For instance, words used with different inflections, such as come, coming, and was coming, can all be reduced to the root "come". Lemmatization is the process that takes into account the morphological analysis of words and accordingly separates the meaningful word into its roots (Ağralı & Aydın, 2021, p. 28). For these processes, the Natural Language Toolkit (NLTK) library of the Python programming language is most commonly used in NLP and thus is also used here. For stopwords, the stop_words sub-module of the NLTK library is used, and the NLTK library's *nltk.* stem module is used for stemming.

3.3. Data Labeling

The study has adopted a dictionary-based approach to determine whether the preprocessed tweets contain a positive, negative, or neutral meaning. The dictionary-based approach uses a sentiment dictionary containing opinion words which are then matched with the data to determine polarity. As a result of this matching, sentiment scores are assigned to the opinion words that define the positive, negative, and neutral scores of the words in the dictionary (Hardeniya & Borikar, 2016, p. 318).

This study also uses the TextBlob Library as a dictionary-based approach for labeling the tweets. TextBlob is an open-source Python library built upon NLTK and analyzes textual content to assign polarity scores ranging from -1 (negative) to 1 (positive). It achieves this by meticulously examining each word within a text fragment and assigning semantic scores to individual words. These scores are then meticulously weighted, effectively calculating a weighted average to determine a comprehensive score for the entire sentence based on the polarity contributions of each word (Zahoor & Rohilla, 2020, p. 538). The Textblob library can only classify English texts into three types: positive, negative, and neutral. A polarity value greater than 0 is considered positive, less than 0 is considered negative; and equal to 0 is considered neutral (Diyasa et al., 2021, p. 4). With the TextBlob library, tweets are labeled as positive, negative, or neutral according to their polarity values. As a result of TextBlob's sentiment outputs, 724 (37.17%) of the user tweets were determined to be positive, 304 (15.61%) to be negative, and 920 (47.22%) to be neutral (Figure 3). Thus, the dataset has been prepared for comparing the models.

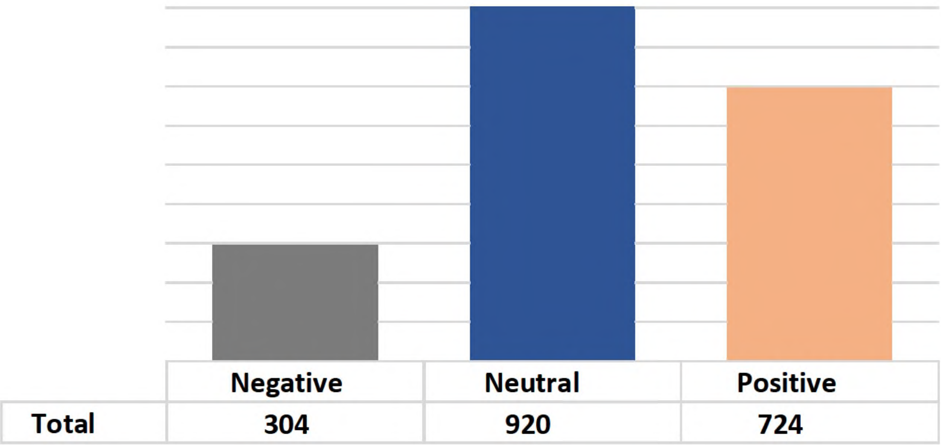


Figure 3. Textblob emotion distributions.

3.4. Word Cloud

This study's tweets labeled as negative, neutral, or positive using sentiment analysis are now visualized with a word cloud. Word clouds are the easiest and most preferred visualization method, as it allows one to visualize the most frequently repeated words in a dataset and to comment on the dataset by looking at these words.

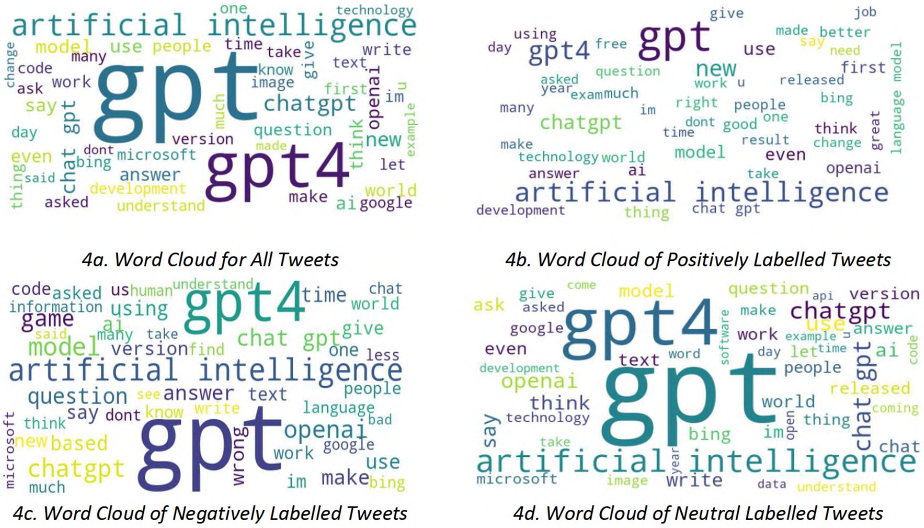


Figure 4. Word clouds of the tweets; a) all tweets, b) positive tweets, c) negative tweets, d) neutral tweets.

When analyzing the distributions of the words used in all the Twitter posts, the most frequently used word is *gpt*, followed by *gpt4, intelligence, artificial, chat, chatgpt, new, model, and openai*. The word clouds containing all the tweets and the sentiments are shown in Figures 4a-4d.

3.5. Data Separation and Modeling

The dataset was prepared for analysis after preprocessing and labeling. The next step splits the dataset into two parts: 80% for training and 20% for testing using the *train_test_split* method in the Python scikit-learn library. For obtaining the optimal hyperparameters, the Grid search technique was additionally applied in order to work with the correct parameters. Hence, the classification performances of the ML algorithms are compared using labeled text data. In this case, the text data should be converted into numerical vectors first. This is done using the *CountVectorizer* class in the *sklearn.feature_extraction.text* module of the scikit-learn library. This allows for the text data to be able to be used with the ML algorithms.

3.6. Comparison Models and Performance Measures

The logistic regression (LR), support vector machines (SVM), XGBoost, and random forest (RF) algorithms have been used to classify the Turkish tweets that were obtained from Twitter and that had been subjected to the GPT-4-themed preprocessing steps. The study uses the confusion matrix, accuracy, precision, sensitivity, and F-1 score performance measures to evaluate the success of the models that have been developed within the scope of the research.

4. FINDINGS

This study has subjected the dataset to sentiment classification using four different ML algorithms: SVM, LR, XGBoost, and RF. The study then examined the performance measures listed in Table 2. The ML techniques applied to the text data labeled with TextBlob were seen to provide successful results. The accuracy rates obtained from the LR, XGBoost, and SVM ML methods were observed to be very close to one another, with XGBoost providing slightly more successful results and RF having the lowest accuracy rate. The same situation is observed when analyzing the other parameters. These values indicate the classification performance of the model to have been successful. The obtained parameter results indicate the ML algorithms that were used to have provided successful classification results. The performances of the ML techniques are listed in Table 2.

Table 2. Performance of Machine Learning Techniques

	Accuracy	Sensitivity	Responsiveness	F-1Score
XGBoost	0.87	0.87	0.87	0.86
LR	0.86	0.87	0.86	0.85
DVM	0.85	0.85	0.85	0.84
RF	0.81	0.83	0.81	0.79

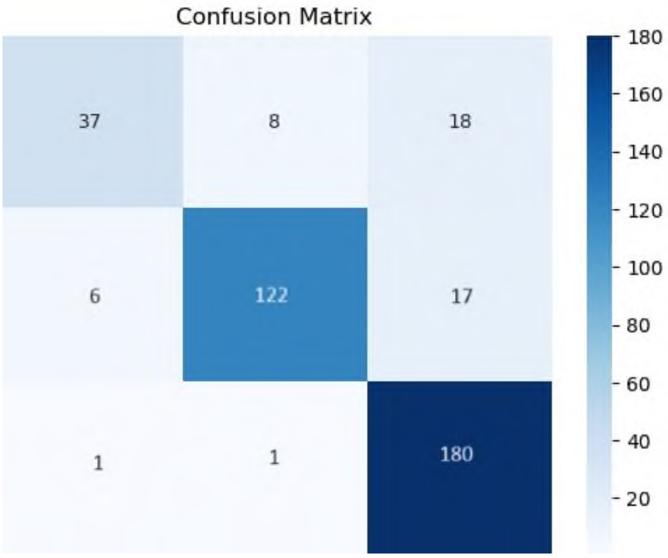


Figure 5. Complexity matrix of the XGBoost algorithm.

Analyzing the complexity matrix of the XGBoost algorithm shows 98.9% (180) of the neutral values, 84.13% (122) of the positive values, and only 58.73% (37) of the negative values to have been correctly classified (see Figure 5).

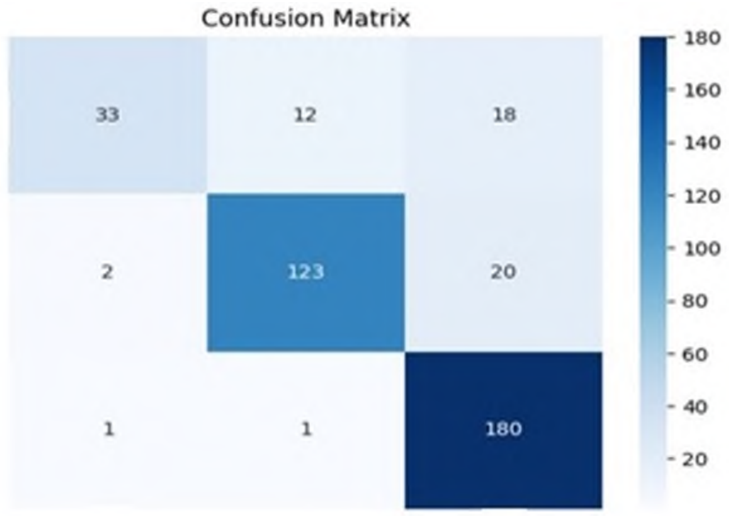


Figure 6. Complexity matrix of the logistic regression algorithm.

Figure 6 depicts the complexity matrix of the LR algorithm. When analyzing the complexity matrix of the LR, 98.9% (180) of the neutral values, 84.83% (123) of the positive values, and 52.38% (33) of the negative values are seen to have been classified correctly.

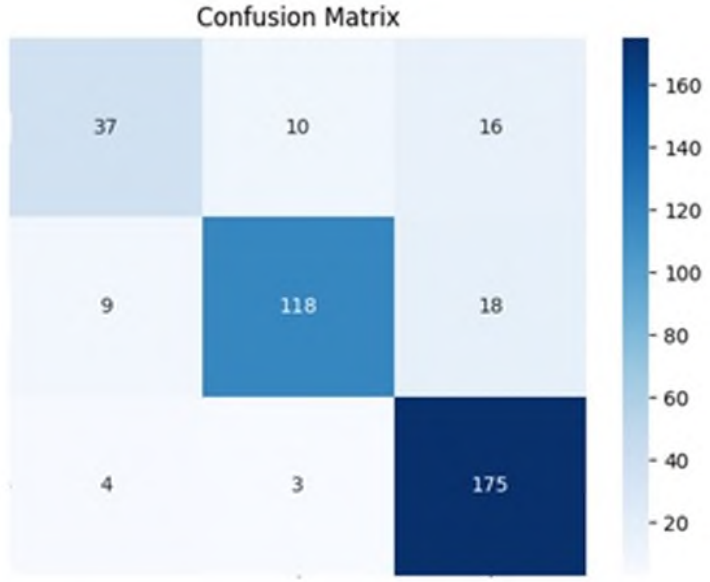


Figure 7. Complexity matrix of the support vector machines algorithm.

Figure 7 represents the complexity matrix of the SVM algorithm. When analyzing the complexity matrix of the SVM, 96.15% (175) of the neutral values, 81.38% (118) of the positive values, and 58.73% (37) of the negative values are observed to have been classified correctly.

When analyzing the complexity matrix of the RF classification, 98.35% (179) of the neutral values, 82.07% (119) of the positive values, and only 28.57% (18) of the negative values were determined to have been classified correctly (see Figure 8).

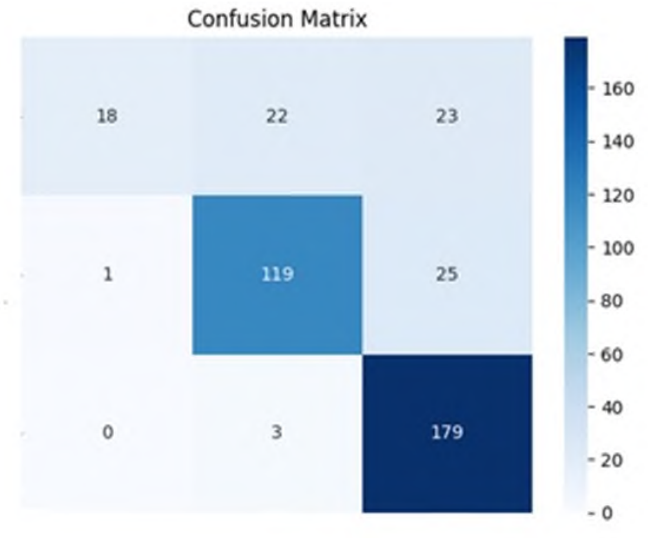


Figure 8. Complexity matrix of the random forest algorithm.

5. DISCUSSION AND CONCLUSION

This study presents the sentiment analysis of Turkish tweets about GPT-4, which was released by OpenAI on March 14, 2023. The study preprocessed the obtained data and then labeled them using the TextBlob dictionary. As a result of the sentiment analysis, 15.61% of the tweets Twitter users posted expressed negative opinions, 37.17% expressed positive opinions, and 47.22% were neutral. The results of the study are consistent with the previous studies by Lampropoulos et al. (2023) and Li et al (2023). The numerically expressed and class-labeled dataset was separated for training and testing. The study also compared the classification performances of the following ML algorithms: SVM, LR, XGBoost, and RF. The results indicate the XGBoost, LR, and SVM algorithms to scored close to one another. However, the XGBoost achieved the highest accuracy rate with an accuracy value of 0.87. All parameter results show that the ML algorithms used in the study to have provided reliable results. Based on these results, the models are seen to have successfully performed the sentiment classification of tweets. When analyzing the complexity matrices related to the results, the XGBoost and LR algorithms were found to have yielded the most successful results when classifying positive and neutral values, while the XGBoost and SVM algorithms were more successful classifying negative values.

Future studies can consider using different ML algorithms, such as Naive Bayes, K-nearest neighbor, and decision trees for the classification of tweets. In addition, tweets on different topics can be analyzed using the same method. Labeled data can also be compared using different sentiment analysis approaches based on a dictionary or on machine learning.

The information obtained from the study can be beneficial for both the general public interested in ChatGPT as well as the developers of ChatGPT-related technology. This information can help these groups create a broader perception of ChatGPT and decide whether they want to use the technology or not. In this way, developers can also understand the social context around ChatGPT and better optimize this technology.

Peer Review: Externally peer-reviewed.

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

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

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

ORCID IDs of the authors

Mustafa Özel	0009-0001-7384-7486
Özlem Çetinkaya Bozkurt	0000-0002-6218-2570

REFERENCES

Abdullah M., Madain A., & Jararweh Y. (2022). ChatGPT: fundamentals, applications and social impacts. *2022 Ninth International Conference on Social Networks Analysis, Management and Security (SNAMS)*. 1-8. http://dx.doi.org/10.1109/SNAMS58071.2022.10062688

Ağralı, Ö., & Aydın, Ö. (2021). Tweet classification and sentiment analysis on metaverse related messages,. *Journal of Metaverse, 1*(1), 25-30.

Aydın, C. (2018). makine öğrenmesi algoritmalari kullanilarak itfaiye istasyonu ihtiyacinin siniflandirilmasi. *Avrupa Bilim ve Teknoloji Dergisi*, (14), 169-175.

Aydın, Ö., & Karaarslan, E. (2023). Is ChatGPT leading generative AI? What is beyond expectations?. 1-23. *Social Science Research Network (SSRN):* https://ssrn.com/abstract=4341500 or http://dx.doi.org/10.2139/ssrn.4341500.

Ayhan, S., & Erdoğmuş, Ş. (2014). Destek vektör makineleriyle siniflandirma problemlerinin çözümü için çekirdek fonksiyonu seçimi. *Eskişehir Osmangazi Üniversitesi İktisadi ve İdari Bilim Dergisi, 9*(1), 175–198.

Agarwal, B., Mittal, N., Bansal, P., & Garg, S. (2015). Sentiment analysis using common-sense and context information. Computational *İntelligence and Neuroscience, 2015*, 1-9. http://dx.doi.org/10.1155/2015/715730

Bayram, İ., & Turan, A. (2022). Türkiye'de kripto para farkindaliği ve tutumu: duygu analizi ve istatistiksel analiz ile bir değerlendirme. *Yönetim Bilişim Sistemleri Dergisi, 8*(2), 20-35. https://dergipark.org.tr/en/pub/ybs/issue/75943/1197985

Bengesi S., Oladunni T., Olusegun R., & Audu H. (2023). A Machine Learning-Sentiment Analysis on Monkeypox Outbreak: An Extensive Dataset to Show the Polarity of Public Opinion From Twitter Tweets, IEEE Access, 11(11811-11826). http://dx.doi.org/10.1109/ACCESS.2023.3242290

Bircan, H. (2004). Lojistik regresyon analizi: Tıp verileri üzerine bir uygulama. *Kocaeli Üniversitesi Sosyal Bilimler Dergisi*, (8), 185-208.

Biltawi, M., Etaiwi, W., Tedmori, S., Hudaib, A., & Awajan, A. (2016). Sentiment classification techniques for arabic language: A survey. *7th International Conference on Information and Communication Systems (ICICS)*, 339-346.

Botchu, R., & Iyengar, K. P. (2023). Will ChatGPT drive radiology in the future?. *Indian Journal of Radiology and Imaging*. https://doi.org/10.1055/s-0043-1769591

Bruns, A., & Burgess, J. E. (2011). The use of Twitter hashtags in the formation of ad hoc publics. *Paper presented at the 6th European Consortium for Political Research General Conference*, 25 - 27 August 2011, University of Iceland, Reykjavik. http://eprints.qut.edu.au/46515/

Dandekar, A. R., Shrotri, A.P., Hargude, N.V., Awate, P.P., & Patil, N.D. (2018,August). The rise of social media and its impact. *Internatıonal Journal of Research And Analytıcal Revıews 5*(3), 882-886.

Diyasa, I. G. S. M., Mandenni, N. M. I. M., Fachrurrozi, M. I., Pradika, S. I., Manab, K. R. N., & Sasmita, N. R. (2021). Twitter sentiment analysis as an evaluation and service base on python textblob. *In IOP Conference Series: Materials Science and Engineering*, 1125(1), 1-12. https://doi.org/10.1088/1757-899X/1125/1/012034

Doğan, G. (2022). Makine öğrenmesi algoritmalari ile betonarme kirişlerin burulma momenti tahmini. *El-Cezeri El-Cezerî Fen ve Mühendislik Dergisi, 9*(2), 912-924.

Feng, Y., Vanam, S., Cherukupally, M., Zheng, W., Qiu, M., & Chen, H. (2023). Investigating code generation performance of chatgpt with crowdsourcing social data. *In Proceedings of the 47th IEEE Computer Software and Applications Conference*, 1-10.

Hardeniya, T., & Borikar, D. A. (2016). Dictionary based approach to sentiment analysis-a review. *International Journal of Advanced Engineering, Management and Science, 2*(5), 317-322.

Kahya, A. N. (2021). Wikipedia'daki verilere metin madenciliği yöntemlerinin uygulanması. *Eskişehir Türk Dünyası Uygulama ve Araştırma Merkezi (ESTUDAM) Bilişim Dergisi, 2*(1), 11-14.

Kaur C., & Sharma A. (2022). Social issues sentiment analysis using python. *2020 5th International Conference on Computing, Communication and Security (ICCCS)*, 1-6. http://dx.doi.org/10.1109/ICCCS49678.2020.9277251

Keskin, E. K. (2023). Yapay zekâ sohbet robotu ChatGPT ve Türkiye internet gündeminde oluşturduğu temalar. *Yeni Medya Elektronik Dergisi, 7*(2), 114-131.

Kirmani, A. R. (2022), Artificial Intelligence-Enabled Science Poetry. ACS Energy Letters, 8, 574–576. https://doi.org/10.1021/acsenergylett.2c02758

Korkmaz, A., Aktürk, C., & Talan, T. (2023). Analyzing the user's sentiments of ChatGPT using twitter data. *Iraqi Journal for Computer Science and Mathematics, 4*(2), 202-214.

Koubaa, A. (2023). GPT-4 vs. GPT-3.5: A concise showdown. Preprint. https://doi.org/10.20944/preprints202303.0422.v1

Küçükkartal, H. K. (2020). Twitter'daki verilere metin madenciliği yöntemlerinin uygulanması. *Eskişehir Türk Dünyası Uygulama ve Araştırma Merkezi Bilişim Dergisi, 1*(2), 10-13.

Lampropoulos, G., Ferdig, R. E., & Kaplan-Rakowski, R. (2023). A social media data analysis of general and educational use of chatgpt: understanding emotional educators. *Available at SSRN*: https://ssrn.com/abstract=4468181 or http://dx.doi.org/10.2139/ssrn.4468181

Li, L., Ma, Z., Fan, L., Lee, S., Yu, H., & Hemphill, L. (2023). ChatGPT in education: A discourse analysis of worries and concerns on social media, 1-35. arXiv preprint arXiv:2305.02201.

Liu, Y., Han, T., Ma, S., Zhang, J., Yang, Y., Tian, J., ... & Ge, B. (2023). Summary of ChatGPT/GPT-4 research and perspective towards the future of large language models. 1-35. arXiv preprint arXiv: 2304.01852. https://doi.org/10.48550/arXiv.2304.01852.

Lund, B.D., & Wang, T. (2023). Chatting about ChatGPT: How may AI and GPT impact academia and libraries? *Library Hi Tech News, 40*(3), 26-29. https://doi.org/10.1108/LHTN-01-2023-0009.

Metlek, S., & Kayaalp, K. (2020). Derin öğrenme ve destek vektör makineleri ile görüntüden cinsiyet tahmini. *Düzce Üniversitesi Bilim ve Teknoloji Dergisi, 8*(3), 2208-2228.

Munggaran, J. P., Alhafidz, A. A., Taqy, M., Agustini, D. A. R., & Munawir, M. (2023). Sentiment analysis of twitter users' opinion data regarding the use of ChatGPT in education. *Journal of Computer Engineering, Electronics and Information Technology, 2*(2), 75-88.

Rachman F. H., Imamah, & Rintyarna B. S. (2022). Sentiment analysis of madura tourism in new normal era using text blob and KNN with hyperparameter tuning. *2021 International Seminar on Machine Learning, Optimization, and Data Science (ISMODE)*, 23-27. http://dx.doi.org/10.1109/ISMODE53584.2022.9742894

Ulaş, M., & Karabay, B. (2020). Terör saldırılarını içeren büyük verinin makine öğrenmesi teknikleri ile analizi. *Fırat Üniversitesi Mühendislik Bilimleri Dergisi, 32*(1), 267-277.

Okoloegbo, C. A., Eze, U. F., Chukwudebe, G. A., & Nwokonkwo, O. C. (2022). Multilingual Cyberbullying Detector (CD) Application for Nigerian Pidgin and Igbo Language Corpus. *In 2022 5th Information Technology for Education and Development (ITED)*, 1-6. https://doi.org/10.1109/ITED56637.2022.10051345

OpenAI (2023). GPT-4 Technical Report. https://cdn.openai.com/papers/gpt-4.pdf.

Rahman, S., Jahan, N., Sadia, F., & Mahmud, I. (2023). Social crisis detection using twitter based text mining-a machine learning approach, BULLETIN OF ELECTRICAL ENGINEERING AND INFORMATICS, 12(2), 1069-1077. https://doi.org/10.11591/eei.v12i2.3957

Santra, A. K., & Christy, C. J. (2012). Genetic algorithm and confusion matrix for document clustering. *International Journal of Computer Science Issues (IJCSI), 9*(1), 322-328.

Shinde, P. P., & Shah, S. (2018). *A review of machine learning and deep learning applications.* In 2018 Fourth İnternational Conference on Computing Communication Control And Automation (ICCUBEA), 1-6. https://doi.org/10.1109/iccubea.2018.8697857

Suh, B., Hong, L., Pirolli, P., & Chi, E. H. (2010). Want to be retweeted? Large scale analytics on factors impacting retweet in Twitter network. *In 2010 Second İnternational Conference on Social Computing*, 177-184.

Tekin, R., & Yaman, O. (2023). Akıllı ev sistemleri için XGBoost tabanlı saldırı tespit yöntemi. *Journal of Intelligent Systems: Theory and Applications, 6*(2), 152-158. https://doi.org/10.38016/jista.1075054

Turan, A. K., & Polat, H. (2024). Yarı denetimli makine öğrenmesi yöntemini kullanarak müzik türlerinin tespiti. *Gazi University Journal of Science Part C: Design and Technology, 12*(1), 92-107. https://doi.org/10.29109/gujsc.1352477

Veranyurt, Ü., Deveci, A., Esen, M. F., & Veranyurt, O. (2020). Makine öğrenmesi teknikleriyle hastalık sınıflandırması: Random forest, k-nearest neighbour ve adaboost algoritmaları uygulaması. *Uluslararası Sağlık Yönetimi ve Stratejileri Araştırma Dergisi, 6*(2), 275-286.

Verma, A., Mittal, V., & Dawn, S. (2019). *FIND: Fake information and news detections using deep learning*. In 2019 Twelfth İnternational Conference On Contemporary Computing (IC3), 1-7. https://doi.org/10.1109/IC3.2019.8844892

Wankhade, M., Rao, A. C. S., & Kulkarni, C. (2022). A Survey on sentiment analysis methods, applications, and challenges. *Artificial Intelligence Review, 55*(7), 5731-5780.

Zahoor, S., & Rohilla, R. (2020). Twitter sentiment analysis using lexical or rule based approach: a case study. *In 2020 8th International Conference on Reliability, Infocom Technologies and Optimization (Trends and Future Directions) (ICRITO)*, 537-542.

Zhang, L., Wang, S., & Liu, B. (2018). Deep learning for sentiment analysis: A survey. Wiley Interdisciplinary Reviews: *Data Mining and Knowledge Discovery, 8*(4), 1-25.

How cite this article

Özel, M., & Çetinkaya Bozkurt, Ö. (2024). Sentiment analysis on GPT-4 with comparative models using Twitter data. *Acta Infologica, 8*(1), 23-33. https://doi.org/10.26650/acin.1418834

Acta Infologica
acin 2024, 8 (1), 34–50
DOI: 10.26650/acin.1444470

Research Article / Araştırma Makalesi

Network Forensics Analysis of Cyber Attacks on Computer Systems using Machine Learning Techniques

Firdevs Yıldız[1], Batuhan Gül[1],Fatih Ertam[1]

[1] Fırat University, Faculty of Technology, Department of Digital Forensics Engineering, Elazığ, Türkiye

Corresponding author : Batuhan Gül
E-mail : b.gul@firat.edu.tr

ABSTRACT

With the rapid development of technology, significant progress has been observed regarding the Internet and interconnected devices, increasing the risk of cyberattacks targeting these platforms. These attacks take diverse and sophisticated forms and pose a serious threat to companies, potentially causing substantial financial losses and service disruptions. In response, the pressing need exists to develop robust defense strategies. This research focuses on analyzing attacks on information systems, specifically concentrating on network forensics using machine learning techniques. The initial phase involves executing various attack scenarios in a virtual environment, recording network packets, and extracting relevant features to create a dataset. A classification framework is then created that includes machine learning algorithms such as random forest, support vector machine (SVM), and Naïve Bayes. Comparing the performance of these algorithms on the study's dataset has revealed the random forest algorithm to achieve the highest accuracy rate at 94.8%, with Naive Bayes having the lowest at 78.9

Keywords: Machine learning, cyberthreat, network forensics, classification algorithms, intrusion detection system

Submitted	: 28.02.2024
Revision Requested	: 05.04.2024
Last Revision Received	: 15.04.2024
Accepted	: 09.05.2024
Published Online	: 11.06.2024

1. INTRODUCTION

In tandem with the continual advancement of technology, computer networks have become an indispensable component in nearly every facet of human life and offer convenience in numerous domains. Termed as structures that facilitate data exchange and communication among computers, computer networks enable seamless interactions among these devices. They serve various beneficial purposes, including data transmission, information exchange, and resource sharing. The increased prevalence of computer network utilization corresponds with a parallel escalation in the significance of network security (Akbal et al., 2019). Due to their ubiquity and paramount importance in human life, computer networks have become a target for cyberattacks.

Cyberattacks directed toward computer networks have transcended the realm of information operations, exerting a broader impact. They not only pose a threat to individual users but also jeopardize the security of institutions, governments, and overall societal well-being. These attacks manifest in various forms, including unauthorized network access, data theft, service disruption, ransomware incidents, and other malicious activities. In addition to inflicting material damages, cyberattacks can lead to broader consequences such as compromised personal privacy, disclosure of trade secrets, and implications for national security (Li & Liu, 2021). Hence, attacks on computer networks have heightened the need for network forensics. Network forensics is a subcategory of digital forensics that employs scientific methodologies to allows electronic evidence related to a crime to be presented in an understandable and unaltered state before judicial authorities. Digital forensics encompasses the entirety of the evidence examination process and employs scientific techniques to facilitate the elucidation of a crime (Başlar, 2020). Meanwhile, network forensics constitutes a digital forensic process encompassing the investigation, analysis, and monitoring of computer networks (Qureshi et al., 2021). Network forensics involves the monitoring of a network for abnormal traffic and the identification of unauthorized entries. An assailant may expunge all log files from a compromised central computer, rendering network-based evidence the sole available proof for forensic analysis in such circumstances (Hunt, 2012). Hence, with the increasing complexity of attacks, the significance of network forensics has further heightened.

With the increasing number and diversity of cyber threats, the field of security is undergoing rapid development. Numerous software and hardware network security tools such as security firewalls, antivirus programs, and intrusion detection systems have been developed (Özekes & Karakoç, 2019). Machine learning is currently gaining popularity due to how it provides a set of methods and techniques that yield high accuracy for detecting attacks. The advancement of machine learning methods and techniques presents new opportunities in the field of network forensics. Machine learning is a subfield of artificial intelligence (AI) that enables computers to learn from data without being explicitly programmed, thus facilitating the resolution of complex problems (Bi et al., 2019). Machine learning algorithms make predictions about new data based on training data without explicitly specifying how models will be applied. Machine learning techniques are garnering significant attention across various industries. In the field of cybersecurity, these techniques are being applied for detecting new and sophisticated attacks, thus contributing to the advancement of cybersecurity(Shaukat et al., 2020). This study focuses on analyzing attacks on computer systems from the perspective of network forensics using machine learning techniques and addresses fundamental concepts in network forensics, machine learning techniques, and studies related to the application of these techniques in the context of network forensics. The study conducts various attacks in a virtual environment and uses different learning methods to run classification processes based on the dataset generated from these attacks for subsequent analysis.

Briefly, the main contributions of this review can be stated as follows:

- The study generates attack scenarios on computer systems, classifies these attack types using machine learning methods, and compares the performances of the classification algorithms by creating and employing different attack scenarios.
- The study then analyzes the attacks on information systems from the perspective of network forensics using machine learning techniques.
- The study also executes diverse attack scenarios in a virtual environment, capturing network packets from the conducted attacks, extracting features, and subsequently constructing a new dataset.
- The study utilizes the created dataset to perform classifications using the random forest, support vector machine (SVM), Naive Bayes, logistic regression, and decision tree classification algorithms. The classification results are compared using performance metrics. In the experiments, the random forest classifier ranked the highest in performance by scoring 94.8% in accuracy, 98% in precision, 91.8% in recall, and 92.4% in F1 score.

The remainder of this paper is organized as follows. Section 2 provides a summary of previous studies conducted in this field. The third section discusses machine learning and the classification algorithms. Section 4 introduces the proposed method and creates the dataset. Section 5 classifies the machine learning classifiers used in the experiments on the dataset. The final section discusses the findings and future directions.

2. RELATED WORKS

Wani et al. (2019) presented an approach aimed at detecting distributed denial-of-service (DDoS) attacks in cloud computing environments. Their study conducted attacks using the Tor Hammer attack tool in the ownCloud environment and created a dataset. They employed SVM, Naïve Bayes, and random forest machine learning algorithms to detect DDoS attacks and compared the success rates using performance metrics. The results of the experiments indicated the SVM model to exhibit higher accuracy and precision compared to Naïve Bayes and random forest.

İnce et al. (2021) addressed and compared machine learning algorithms. They utilized the NSL-KDD dataset to evaluate the comparisons. They then applied different tests on the NSL-KDD dataset and used 5% of, 10% of, and then the entire dataset for these tests. They calculated the models' performances using accuracy and F-score values, with the overfitting machine learning (OML) algorithm achieving the highest performance at 99.8% accuracy and a 99.9% F-score.

Ahmetoğlu and Das (2021) proposed an approach based on intuitive feature selection and machine learning to detect web application attacks using hybrid intrusion detection systems. They combined the web application attacks and normal flow examples from the CSE-CIC-IDS2018 dataset after the data preprocessing steps to create a new dataset. They then used the created dataset for calculating the average mean square error and feature count optimization using the genetic algorithm and logistic regression. They tested this optimization process with five different machine learning algorithms: random forest, SVM, Naïve Bayes, k-nearest neighbors (KNN), and deep neural networks (DNN). The most successful methods for classification were RF, KNN, and DNN, respectively. Upon examining the results, they observed the number of features to be reduced by 85% while keeping the classification success rates at the 99% level.

Radivilova et al. (2019) examined machine learning classification methods for detecting DDoS attacks. They tested the performances of machine learning algorithms using datasets created with different DDoS attack scenarios. They also evaluated the detection performance of the algorithms using performance metrics such as accuracy rate, precision, specificity, and F1 score. Their evaluation results emphasize the successful application of classification methods for detecting DDoS attacks, particularly the random forest algorithm.

Ferrag et al. (2020) presented a study involving deep learning methods for detecting cyberattacks using datasets such as CSE-CIC-IDS2018 and Bot-IoT and examined deep learning methods including recurrent neural networks (RNN), DNN, restricted Boltzmann machines (RBM), deep belief networks (DBN), convolutional neural networks (CNN), deep Boltzmann machines (DBM), and deep autoencoders (DA). They categorized the literature's 35 attack detection datasets and analyzed the performance of deep learning methods on these datasets.

Dina et al. (2021) investigated the use of machine learning methods for detecting unauthorized entries. They employed various machine learning techniques such as decision trees, SVM, artificial neural networks (ANN), random forest, and Naïve Bayes classifiers for unauthorized entry detection and examined their advantages and disadvantages. Additionally, they utilized datasets created to represent various types of attacks and performed evaluations using performance metrics.

Shafiq et al. (2020) discussed the importance of identifying cyberattack traffic for IoT security. They proposed an effective machine learning algorithm selection framework and a hybrid algorithm for identifying anomalies and unauthorized entries in IoT network traffic. The dataset comprises normal traffic, attacks, and Internet of Things (IoT) Botnet attacks for identifying and detecting the best attacks in the IoT network. They selected and accurately applied five machine learning algorithms (i.e., Naïve Bayes, Bayes Net, C4.5 decision tree, random forest, and random tree). They then evaluated the performance of the applied machine learning algorithms and stated the performance of Naïve Bayes and random tree to be highly effective compared to the other machine learning algorithms. When comparing the performance results from the Naïve Bayes and random tree machine learning algorithms, they indicated Naïve Bayes to be the most effective machine learning algorithm.

Shaukat et al. (2020) presented an evaluation of commonly used machine learning methods for detecting certain cyber threats by examining three fundamental machine learning techniques (i.e., deep belief network, decision tree, and SVM). Based on frequently used datasets considered as references, they briefly evaluated the performance of these machine learning techniques in the areas of unwanted email detection, unauthorized entry detection, and malware detection.

Zhang et al. (2019) proposed a network attack detection method based on a deep hierarchical network and original flow data for the CICIDS2017 and CTU datasets. Their method used CNN classification to learn spatial features and long short-term memory (LSTM) classification to learn temporal features. As a result of the classification process with

CNN+LSTM, they detected data in the CICIDS2017 dataset with 99.8% accuracy and in the CTU dataset with 98.7% accuracy.

Aamir et al. (2021) conducted a study on a subset of the CICIDS2017 dataset, specifically the Juma-working hours-afternoon dataset, which includes Benign, PortScan, and DDoS labels. The first stage of the study removed features with infinite or calculated values from the dataset. Correlations between features were calculated, and features with correlations below 20% were eliminated. They used the remaining features to create a new dataset and then applied standard scaler normalization . Approximately 70% of the obtained dataset was allocated for training, and 30% for testing. Classification algorithms and some ensemble classifiers were used, resulting in accuracy values of 60.6%, 97.1%, 99.0%, 68.7%, and 85.5%. These results provide an analysis of different algorithm performances and variability regarding classification accuracy.

Karaman et al. (2021) used ANN to analyze attack detection models using the CSE-CIC-IDS 2018 dataset. Their study created five separate sub-datasets and selected features for each sub-dataset. They used the created models to determine whether attacks were DDoS, BruteForce, Botnet, or denial-of-service (DoS) attacks and to identify the type of attack, achieving accuracy values of 99.11%, 99.31%, 99.26%, 93.23%, and 92.26% for each sub-dataset.

Aslan and Yilmaz (2021) proposed a deep learning-based architecture capable of classifying malware derivatives effectively using a hybrid model. They first collected data and then designed the DNN to be used in the study. They then tested the proposed model on the Malimg, Microsoft BIG 2015, and Malevis datasets and reached high performance rates compared to previous examples in the literature, with testing on the Malimg dataset achieving an accuracy of 97.78%.

Al-Zubi et al. (2021) presented a model for the security and privacy of patient data. Using machine learning models, this approach predicted cyberattack behavior in the healthcare field and facilitated the processing of this data. The proposed approach is based on a patient-centric design, allowing users to control data sharing access by protecting information on trusted devices such as their mobile phones. Experimental results indicated the proposed model to provide a higher attack prediction rate (96.5%), accuracy rate (98.2%), and efficiency rate (97.8%) with lower latency (21.3%) and reduced communication costs (18.9%) compared to other existing models.

Pallathadka et al. (2023) proposed a method that compares the performance of students in an institution using machine learning methods. They investigated the performance of such machine learning methods as Naïve Bayes, ID3, C4.5, and SVM using the UCI machinery student performance dataset. As a result of the comparison, the SVM algorithm was seen to have the highest accuracy rate.

Suryadevara (2023) proposed a new method to detect whether a person has diabetes or not using machine learning techniques. The study used the diabetes dataset and compared such classification algorithms as KNN, logistic regression, and random forest over this dataset. As a result of comparing five machine learning methods, decision tree was seen to achieve the highest accuracy rate at 99

Ashton et al. (2023) investigated the advantage and potential of using machine learning in the field of pediatrics. As a result, their research predicted machine learning to be able to greatly help people manage pediatric conditions over the next 5-10 years.

Noella and Priyadarshini (2023) proposed a novel machine learning-based method that analyzes Parkinson's and Alzheimer's diseases. They compared the performances of bagged ensemble, ID3, Naïve Bayes, and multiclass support vector machine classifiers using the Positron Emission Tomography (PET) dataset. In the comparison, bagged ensemble showed a higher performance than other machine learning algorithms with an accuracy of 90.3%, sensitivity of 89%, specificity of 92%, and precision of 87%.

The majority of previous studies have been on health issues, with a very limited number of studies observed to have occurred on analyzing the network forensics of attacks on computer systems using machine learning methods in the past years.

3. MACHINE LEARNING

Machine learning is a subfield of AI and had its foundations laid in the 1950s. In 1959, Arthur Samuel coined the term machine learning and conducted studies related to algorithms used in computer games. With the advancements in technology and data science over the years, machine learning has evolved into a significant scientific discipline that possesses the ability to self-learn through algorithms. The data provided to a system is recognized through the employed algorithms, and the system responds accordingly in the output, thus becoming intelligent over time without human intervention (Sharma et al., 2021).

Currently, machine learning is being applied in various fields, and new algorithms and models are continually being developed. Machine learning algorithms have been instrumental at addressing various issues and contributing to the advancement of fields such as cybersecurity and digital forensics. The capabilities of machine learning enable the detection of cyber threats, the creation of predictions, and the identification of anomalies in a network. The prediction steps using machine learning methods are illustrated in Fig. 1.

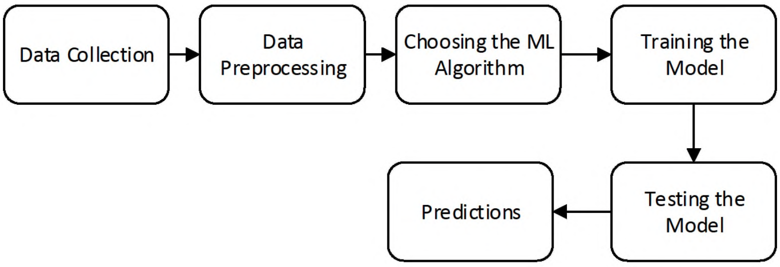


Figure 1. The steps of machine learning.

3.1. Classification Algorithms

This study creates its own dataset to measure the performance of classification algorithms and then compares these algorithms with each other to obtain an algorithm that shows the best performance. This part of the study provides brief information about the classification algorithms used herein.

3.1.1. Decision Trees

Decision tree algorithms divide a dataset into small subsets based on various features and aim to classify the data within each subset. Like branches of a tree, classes are further divided into sub-branches. The data to be classified reach the root of the tree, progress toward the sub-branches based on conditions, and become members of the closest class according to their values. The accuracy of the rules forming the classes affects the algorithm's performance.

Information gain is calculated during root selection. Information gain is utilized to determine how much information a feature provides in the classification problem. The feature with the highest information gain becomes the root. This process continues until all features are exhausted. Information gain is calculated using entropy. Entropy represents the probability of uncertainty or disorder occurring. Although decision tree algorithms have many advantages, they also have disadvantages. Decision tree algorithms can be unreliable because they have high variance, especially when the data set is too noisy. Even the slightest change in the data set can lead to large differences in tree structures. In addition, these algorithms are prone to overfitting and can capture noise in the data set as if it were true data.

3.1.2. Naïve Bayes

The Naïve Bayes algorithm is a machine learning algorithm commonly used in statistical classification problems and provides a probabilistic approach. The Naïve Bayes algorithm is quite advantageous as it does not require too much data for training (Nurdina & Puspita, 2023). Training data are calculated according to Bayes' theorem, and probability percentages are determined using the concept of probability to calculate the relationship between the features and classes in the dataset. New data are then classified based on these probability values. The assumption of the conditional independence of features in the Naïve Bayes algorithms may cause this algorithm to achieve poor performance, considering that real-life datasets may be related to each other. In addition, these algorithms can show high performance over datasets with simple distributions (e.g., Gaussian distribution) while being unable to provide optimum performance over complex data.

3.1.3. Random Forest

The random forest algorithm is a powerful machine learning method commonly used for classification or regression problems. It achieves classification by employing multiple decision trees with the goal of improving classification accuracy. Random decision trees are selected from the dataset to form a forest, thereby enhancing adaptability to the data and obtaining more accurate results over datasets. These algorithms are less prone to overfitting compared to decision tree algorithms. However, using datasets with a large number of data may result in a lot of memory consumption, because these algorithms store more than one decision tree in their memory.

3.1.4. Support Vector Machine (SVM)

SVM is a classifier used for classifying data or performing a regression process. The SVM algorithm draws a boundary to separate two different groups on a plane. This boundary is drawn as far as possible from the members of both groups and is referred to as the margin. The margin represents the gap between the classes and is adjusted in classifications based on the drawn boundaries. Test data belong to the group that is closer to the boundaries. When drawing the boundary, two lines are drawn close to each group. By bringing these drawn boundaries closer to each other, a common boundary is obtained with the aim of classify the data the most accurately. However, these algorithms may experience problems in datasets where each parameter has a different amount of data. In these cases, SVM can classify in favor of the parameter with the most data. This may cause poor performance in correctly classifying parameters with a small amount of data. In addition, SVM algorithms may consume a lot of memory when classifying data sets that contain many features and data. Therefore, it may not be a good option for classifying data sets with high amounts of data.

3.1.5. Logistic Regression

Logistic regression is a supervised learning algorithm that is used when the outcome can only take one of two values. It can be used for both classification and clustering and is employed to predict the probability of data points belonging to two different classes, typically represented as 0 and 1. Logistic regression is commonly used to address binary classification problems. Because the logistic regression algorithm assumes a linear relationship between the independent variables and the log ratios of the obtained result, it may not show high performance when classifying data with complex decision boundaries. Despite this drawback, the logistic regression algorithm is very suitable for use in cases where the outcome variable has two different categories. In addition, having low variance makes this algorithm more suitable for use regarding complex data by reducing the risk of overfitting.

The most interpretable classification algorithms are decision tree and Naïve Bayes algorithms, but they do not perform as well as other algorithms at predicting missing data. The random forest algorithm offers the most advanced prediction performance, and SVMs are more powerful than other algorithms at finding complex decision boundaries. Logistic regression may be the best option for binary classification scenarios but is not the best option for nonlinear relationships. In general, all classification algorithms have their strengths and weaknesses, and this study compares and analyzes all the mentioned classification algorithms over the study's own data set.

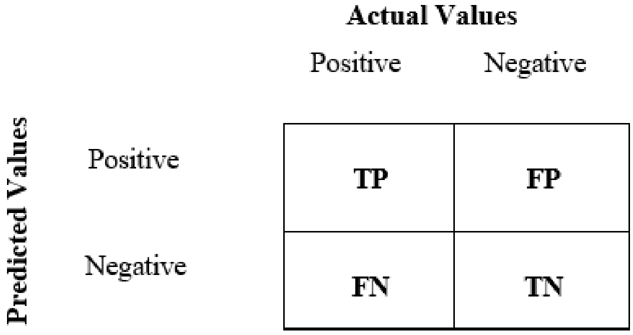
3.2. Performance Metrics

Performance metrics and the confusion matrix are crucial tools used to evaluate the performance of classification models and to analyze classification results in more detail. The confusion matrix is employed to provide a more detailed analysis of classification results by visualizing the relationship between true classes and the classes predicted by the model. These tools help assess the classification capabilities of a model and to identify the measures necessary for improving performance.

After the classification process, the study utilizes such performance metrics as the confusion matrix, accuracy, precision, recall, and F1 score to evaluate the performance of the employed methods, with the chosen performance metrics being explained as follows.

3.2.1. Confusion Matrix

The performance measurement of real and predicted data for a classification problem is represented by a confusion matrix consisting of four cells representing four different values, as illustrated in Fig. 2.



		Actual Values	
		Positive	Negative
Predicted Values	Positive	TP	FP
	Negative	FN	TN

Figure 2. The confusion matrix.

True Positive (TP): Positively predicted and actually positive. Correctly predicted.
False Positive (FP): Predicted as positive but actually negative. Incorrectly predicted.
False Negative (FN): Predicted as negative but actually positive. Incorrectly predicted.
True Negative (TN): Negatively predicted and actually negative. Correctly predicted.

3.2.2. Accuracy

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN} \tag{1}$$

3.2.3. Precision

Precision is found by dividing the total predicted positive rate of the classification model with the true positives. A model with a high precision value shows that the model can make predictions with high accuracy. This is particularly useful for reducing the number of false positives. The precision rate is calculated as shown in Equation 2.

$$Precision = \frac{TP}{TP + FP} \tag{2}$$

3.2.4. Recall

Recall, also known as sensitivity or true positive rate, is the ratio of true positive predictions to the total number of actual positive instances. It measures how many of the actual positive instances were correctly identified. The recall rate is calculated as shown in Equation 3.

$$Recall = \frac{TP}{TP + FN} \tag{3}$$

3.2.5. F1 Score

The F1 score is calculated by taking the harmonic mean of the precision and recall metrics. This is used to achieve a balanced classification performance. The F1 score is calculated as shown in Equation 4.

$$F1 = \frac{Precision * Recall}{Precision + Recall} \tag{4}$$

4. METHOD

This study conducts attacks on computer systems, classifies the executed attacks using machine learning methods, and subsequently compares the performance results of the employed classification methods. The study has tested the executed attacks and the algorithms on a computer running an Intel(R) Core(TM) i7-6700HQ CPU @ 2.60GHz 2.59 GHz. Two virtual machines, one running Linux and the other Windows 7, were installed on the computer described above. The system on which the attacks were executed is a virtual machine running the Kali Linux operating system, while the target machine for the attacks is a virtual machine running the Windows 7 operating system.

The packets of the attacks on the virtual Windows 7 machine were intercepted using Wireshark and saved in .pcap format. The CICFlowMeter tool was utilized to analyze and extract features from the .pcap files obtained with Wireshark, and then the data were converted to .csv format for classification purposes. The Python programming language was employed to perform the classification using machine learning algorithms. The detailed information about the used software is provided in Table 1.

This study classifies the attacks on computer systems using machine learning algorithms and compares their performances. The flowchart for the proposed method is illustrated in Fig. 3.

The first stage carried out attacks in a virtual environment and obtained the ,pcap files using Wireshark. The CICFlowMeter tool was utilized to extract features from the obtained .pcap files which were then convert to .csv format. The resulting .csv files were merged to create a single .csv file forming the dataset. Attacks with a low number of instances in the dataset were removed, infinite values were converted to not-a-number (NaN) values, and then NaN

Table 1. *The Software Programs Used in the Study*

Software	Developer	Purpose of Use	Output Format	Source
Wireshark	Gerald Combs	Captures data packets by listening to network traffic, analyzes them, and saves them as .pcap files.	.pcap	https://www.wireshark.org/
CICFlowMeter	Canadian Institute for Cybersecurity	It has been used to analyze network traffic and extract its features.	.csv	https://github.com/ISCX/CICFlowMeter
Python	Guido van Rossum	It is an open-source programming language. It has been used for the classification process.		https://www.python.org/

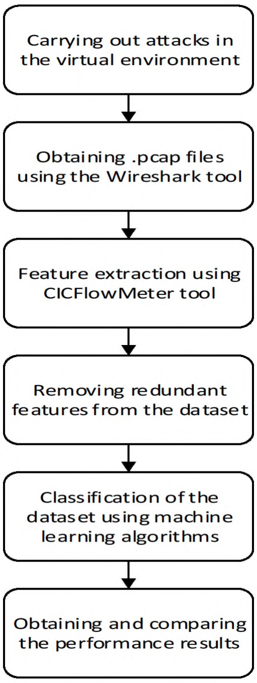

Carrying out attacks in the virtual environment
Obtaining .pcap files using the Wireshark tool
Feature extraction using CICFlowMeter tool
Removing redundant features from the dataset
Classification of the dataset using machine learning algorithms
Obtaining and comparing the performance results

Figure 3. The flowchart for the proposed method.

values were also eliminated from the dataset. This process ensured the correction of unnecessary, corrupted, or missing data. The dataset contains 80 features.

To perform the classification process on the dataset, 20% of the data was set aside for testing and 80% for training. Classification was carried out using classification algorithms. The performance results are compared using performance metrics.

4.1. Attacks Performed on the Virtual Machine

Cyberattacks were carried out to create a data set. Six different types of attacks were executed (i.e., Brute Force, DoS ICMP Flood, DoS Syn Flood, Syn Scan, UDP Scan, and Man in the Middle). Before initiating the attacks, Wireshark was run on the target machine. The Hydra tool was used for executing the Brute Force attack, and the Bettercap tool was employed for implementing the Man in the Middle attack. Hping3 was utilized for conducting the DoS attacks, and Nmap was employed for executing the Scan attacks. In addition to the attack data, normal traffic was included as a separate class in the dataset, resulting in a total of seven classes (Table 2) with numerical labels assigned accordingly.

Table 2. *Class Labels and Descriptions*

Labels	Classes	Descriptions	Software Programs Used
0	Normal	Normal network traffic has been used.	The network was eavesdropped on using Wireshark.
1	BruteForce	Session login information was obtained using the user information method over FTP, RDP, SSH protocols.	Session information was obtained using the Hydra tool.
2	Man in The Middle	Communication network was eavesdropped on and manipulated.	The Bettercap tool was used to eavesdrop on and manipulate communication on the network.
3	DoS_Syn_Flood	Intensive SYN packets were sent to the target to deplete resources.	Network traffic was manipulated using the Hping3 tool.
4	DoS_ICMP_Flood	Intensive ICMP packets were sent to the target to deplete resources.	Network traffic was manipulated using the Hping3 tool.
5	SYN_Scan	Port scanning was performed by sending SYN packets to determine the TCP connection status of the target.	The Nmap tool was used to determine the open ports of the target system.

4.2. Dataset

This stage collects the raw data by intercepting network traffic packets. Before initiating any cyberattacks, the general network traffic was monitored through Wireshark, and the recorded .pcap files were labeled as normal. Six different attacks were executed targeting specific objectives, and Wireshark was used to intercept the network attack packets. The obtained .pcap files were labeled according to the type of attack. Features were extracted from all obtained .pcap files using CICFlowmeter software, and then the .csv files were created for the dataset. The obtained .csv files were merged to create a single .csv file to form the dataset. Low-occurrence attacks were removed from the dataset, infinite values were converted to NaN values, and then the NaN values were also eliminated from the dataset. The final dataset comprises 8,094 instances, with the quantity of data for each class provided in Table 3.

Table 3. *Class-Based Data Quantities*

Class	Amount of Data
DoS_Syn_Flood	2,559
UDP_Scan	1,916
normal	1,131
BruteForce	1,118
SYN_Scan	1,002
DoS_ICMP_Flood	315
Man in The Middle	53

Initially, the dataset contained a total of 84 features; however, features were removed that could negatively impact or that were ineffective for the classification process. The removed features were generally string values. This process aimed to address unnecessary, corrupted, or missing data. In the final version of the dataset, 80 features were retained. To perform the classification over the dataset, 20% of the data were designated as the test set, and 80% as the training set. The features extracted from the dataset are based on the work of Kilincer et al. (2022).

4.3. Classification

The machine learning algorithms used for classifying the attacks against the target system include the random forest, SVM, Naïve Bayes, logistic regression, and decision tree methods. Performance metrics were employed to evaluate the models' effectiveness. Multiclass classification was used for classification and created a total of seven classes comprising six attack types and one normal class. The selection of these classification methods is based on their different capabilities that are tailored to specific features and requirements. Machine learning algorithms are widely utilized in tasks such as classification and regression and are particularly effective when dealing with high-dimensional and complex data. These methods often come with error-correction capabilities and are resistant to overfitting the data, making them suitable for applications in various domains. The study now evaluates the differences between classification models and determines which model demonstrates the best performance.

5. EXPERIMENTAL RESULTS

This study has involved creating a dataset, splitting it into training and testing sets, and applying the specified classification models to the dataset. After successful training, the study evaluated the results from the selected machine learning algorithms.

The random forest model exhibits high accuracy when predicting the normal class. It shows good performance with high accuracy at predicting attacks for the BruteForce, Man in the Middle, DoS_Syn_Flood, and SYN_Scan classes. The model demonstrates high accuracy at predicting UDP_Scan attacks. However, for DoS_ICMP_Flood attacks, the model exhibits lower performance compared to the other classes of attacks. The random forest model's overall accuracy rate has been calculated as 94.87%, indicating a high ability to make accurate classifications. The performance characteristics of the random forest classification method are presented in Table 4 and Fig. 4.

Table 4. *The Performance Metrics of the Random Forest Algorithm*

Labels	Classes	Precision	Recall	F1-Score
0	Normal	0.98	0.85	0.91
1	BruteForce	1.00	1.00	1.00
2	Man in The Middle	1.00	1.00	1.00
3	DoS_Syn_Flood	1.00	1.00	1.00
4	DoS_ICMP_Flood	0.61	0.61	0.61
5	SYN_Scan	1.00	1.00	1.00
6	UDP_Scan	0.86	0.93	0.89

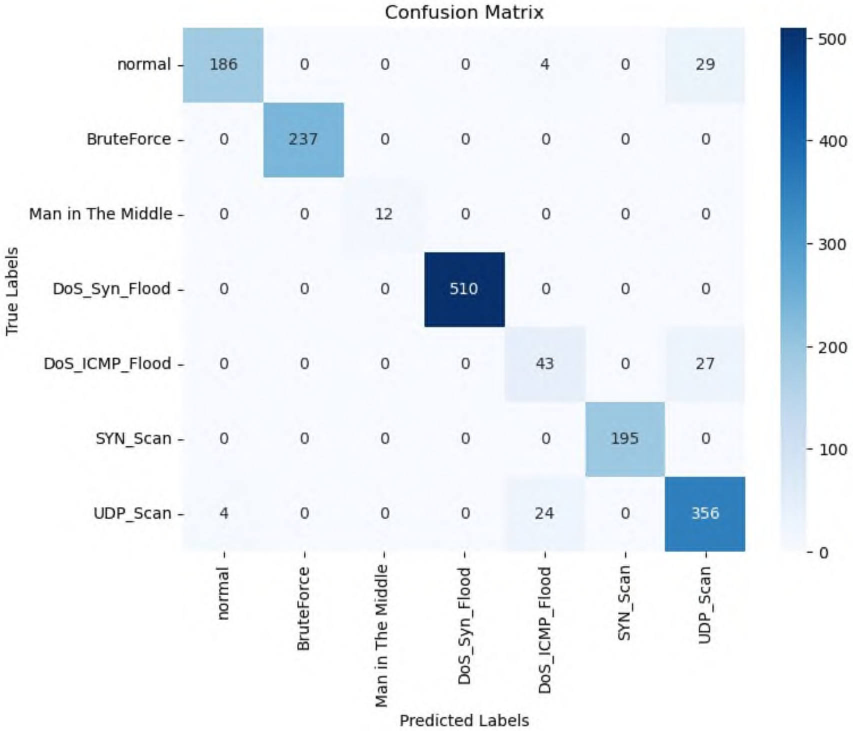


Figure 4. The confusion matrix of the random forest algorithm.

The Naïve Bayes model demonstrates a moderate accuracy when predicting the normal class. Precision, recall, and F1-score for DoS_Syn_Flood and SYN_Scan classes are 1.00, indicating the Naïve Bayes model to perform well at predicting these attacks. The model shows a moderate accuracy at predicting DoS_ICMP_Flood and UDP_Scan attacks. However, it exhibits lower performance at predicting Man in the Middle attacks compared to other classes. The overall accuracy rate of the Naïve Bayes model is calculated as 79%, indicating a moderate level of performance. The Naïve Bayes algorithm's performance values are presented in Table 5 and Fig. 5.

Table 5. *Naive Bayes Algorithm's Performance Values*

Labels	Classes	Precision	Recall	F1-Score
0	normal	0.44	0.79	0.57
1	BruteForce	0.95	0.93	0.94
2	Manin The Middle	0.14	0.91	0.24
3	DoS_Syn_Flood	1.00	1.00	1.00
4	DoS_ICMP_Flood	0.45	0.76	0.56
5	SYN_Scan	1.00	1.00	1.00
6	UDP_Scan	0.80	0.21	0.33

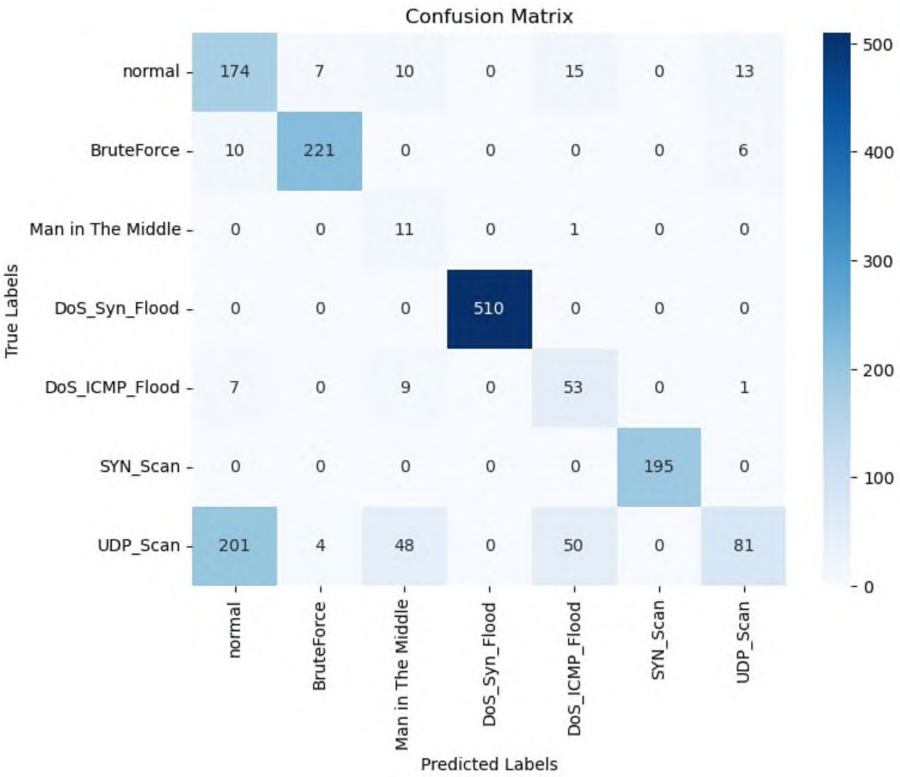


Figure 5. Naive Bayes algorithm's confusion matrix.

The logistic regression model exhibits a moderate level of performance when predicting the normal class. For Brute Force, DoS_Syn_Flood, and SYN_Scan classes, the model appears to perform well at predicting these attacks. The model shows low performance at predicting DoS_ICMP_Flood and Man in the Middle classes and demonstrates a moderate level of performance at predicting UDP_Scan attacks. The overall accuracy rate of the logistic regression model has been calculated as 88%, indicating a good level of performance. The performance values of the logistic regression classification method are presented in Table 6 and Fig. 6.

Table 6. *The Performance Values of the Logistic Regression Algorithm*

Labels	Classes	Precision	Recall	F1-Score
0	Normal	0.73	0.61	0.67
1	BruteForce	0.96	0.99	0.98
2	Man in The Middle	0.88	0.58	0.70
3	DoS_Syn_Flood	0.99	0.99	0.99
4	DoS_ICMP_Flood	0.53	0.38	0.43
5	SYN_Scan	0.99	1.00	1.00
6	UDP_Scan	0.72	0.82	0.77

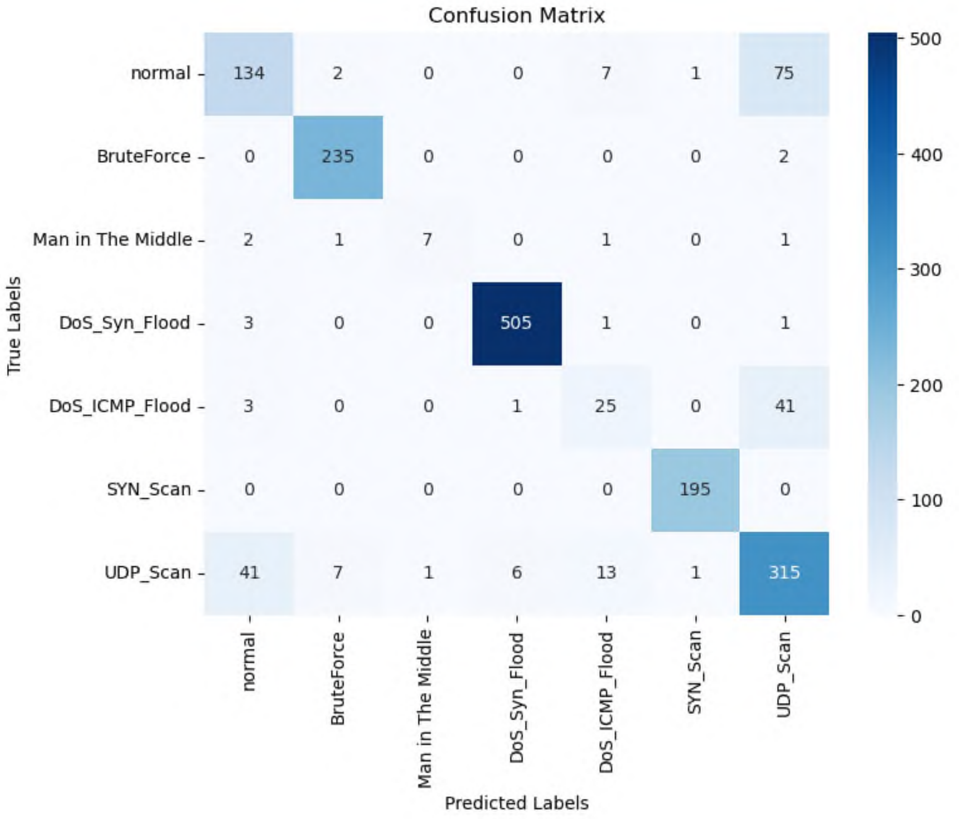


Figure 6. The confusion matrix of the logistic regression algorithm.

The SVM model exhibits a moderate level of performance when predicting the normal class. For Brute Force, DoS_Syn_Flood, and SYN_Scan classes, the SVM model seems to perform well at predicting these attacks. The model shows low performance at predicting the DoS_ICMP_Flood class and demonstrates a moderate level of performance at predicting UDP_Scan attacks. The SVM model made no correct predictions for the Man in the Middle attacks. The overall accuracy rate of the SVM model has been calculated as 85%, indicating the model to generally perform well. The performance values of the SVM method are presented in Table 7 and Fig. 7.

Table 7. *The Performance Values of the SVM Algorithm*

Labels	Classes	Precision	Recall	F1-Score
0	Normal	0.74	0.55	0.63
1	BruteForce	0.88	1.00	0.94
2	Man in The Middle	0.00	0.00	0.00
3	DoS_Syn_Flood	0.97	1.00	0.98
4	DoS_ICMP_Flood	0.74	0.37	0.50
5	SYN_Scan	1.00	1.00	1.00
6	UDP_Scan	0.70	0.81	0.75

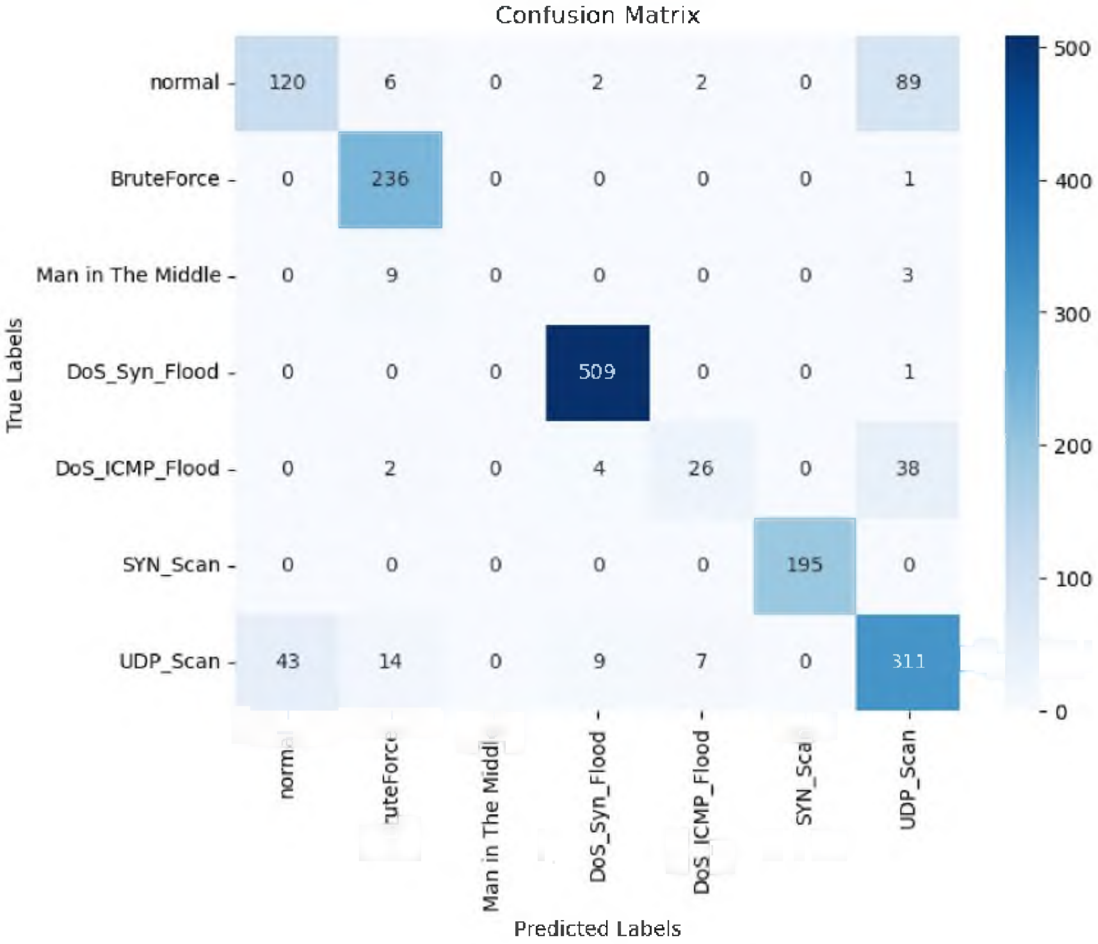


Figure 7. The confusion matrix of the SVM algorithm.

The decision tree model demonstrates high-level performance when predicting the normal class. For Brute Force, DoS_Syn_Flood, SYN_Scan, and UDP_Scan classes, the decision tree model appears to perform well at predicting these attacks. The model shows moderate-level performance at predicting the DoS_ICMP_Flood and Man in the Middle classes. The overall accuracy rate of the decision tree model has been calculated as 94%, indicating the model to generally perform well. The performance values of the decision tree method are presented in Table 8 and Fig. 8.

Table 8. *The Performance Values of the Decision Tree Algorithm*

Labels	Classes	Precision	Recall	F1-Score
0	Normal	0.84	0.88	0.86
1	BruteForce	1.00	1.00	1.00
2	Man in The Middle	1.00	0.92	0.96
3	DoS_Syn_Flood	1.00	1.00	1.00
4	DoS_ICMP_Flood	0.60	0.54	0.57
5	SYN_Scan	1.00	1.00	1.00
6	UDP_Scan	0.88	0.88	0.89

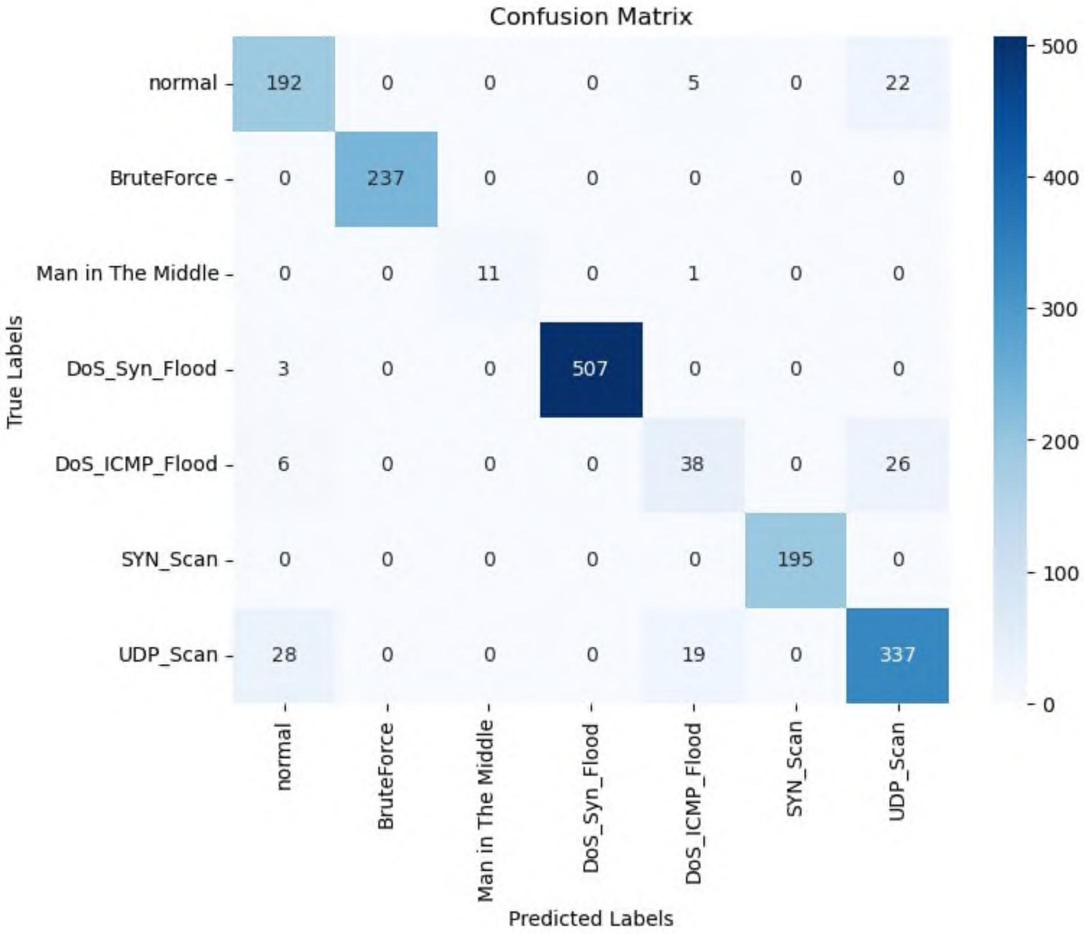


Figure 8. The confusion matrix of the decision tree algorithm.

When comparing the used classification algorithms and looking at the results in Table 9, the random forest classifier is sene to have achieved the highest accuracy rate and to also exhibit good performance in terms of precision, recall, and F1 score. The Naïve Bayes classifier appears to have the lowest performance compared to the other models.

Table 9. *Comparison of the Classifiers Used*

Classifiers	Accuracy	Precision	Recall	F1-Score
Random Forest	0.948	0.930	0.918	0.924
Naive Bayes	0.789	0.853	0.789	0.789
Logistic Regression	0.881	0.877	0.881	0.876
SVM	0.856	0.851	0.856	0.846
Decision Tree	0.941	0.941	0.941	0.941

6. CONCLUSIONS AND FUTURE DIRECTIONS

The field of network forensics addresses issues such as detecting cybercrimes, analyzing attacks, and ensuring network security. The utilization of machine learning techniques in this domain demonstrates significant advantages for obtaining, analyzing, detecting, classifying, and implementing security measures and can lead to the development of robust products. This study discusses threats and attacks on computer systems, with some of these attacks being simulated in a virtual environment. The study executed various types of cyberattacks on a simulated target computer in a virtual environment. Feature extraction from network packets was performed to prepare a dataset for classification. Different machine learning classification algorithms were applied to the created dataset to compare their abilities at accurately classifying attacks.

While some algorithms were successful at classifying certain attacks, others did not achieve the same level of success. Generally, lower performance was observed at classifying DoS_ICMP_Flood and Man in the Middle attacks. The imbalance in dataset numbers is considered a potential cause for the lower performance in these attacks. Collecting more data for such attacks or employing different feature extraction methods could enhance the classification performance. When evaluating the performances of the classification algorithms, the random forest classifier stands out for having achieved the highest accuracy rate and demonstrating good precision, recall, and F1 score.

To assess efficiency, measuring the training and testing times of the selected machine learning algorithms would be beneficial. This approach allows for the evaluation of critical factors such as accuracy and processing speed, thus facilitating an objective comparison for determining the most suitable method.

This study aims to contribute to detecting attacks on computer systems and to developing security strategies. The obtained results can help identify which algorithms perform best on a dataset and which algorithms more accurately detect specific cyberattacks. As a result, the study provides guidance to network forensic experts and security professionals on how to effectively detect, analyze, and take preventive measures against attacks. This study may contribute to enhancing security measures in the field of network forensics and the formulation of strategies to combat cybercrimes.

In order to secure computer systems, stay ahead of industry innovations and tailor security policies is imperative. Contemporary machine learning techniques that have been widely adopted in the realms of cybersecurity and digital forensics play a pivotal role in detecting anomalies and are crucial for data detection and analysis.

High-performance classification algorithms such as random forest are preferable for analyses in digital forensics within network security. To enhance the performance of machine learning algorithms with suboptimal efficiency, further research and improvement efforts are essential. Some attacks exhibited low performance regarding classification, suggesting that collecting more data or employing different feature extraction methods could enhance the applications used for feature extraction.

Future studies are recommended to involve datasets with increased data volume and to conduct analyses for various operating systems. Alongside the continuous development of security technologies and the reinforcement of security measures, this study proposes holding awareness campaigns and training programs on cybersecurity to augment awareness among individuals.

This study's approach has contributed to the perpetual advancement of security technologies and the overall reinforcement of cybersecurity measures.

Peer Review: Externally peer-reviewed.

Author Contributions: Conception/Design of Study- F.Y., B.G.; Data Acquisition- F.Y., F.E.; Data Analysis/Interpretation- F.Y., B.G.; Drafting Manuscript- B.G., F.Y.; Critical Revision of Manuscript- F.E.; Final Approval and Accountability- F.Y., B.G., F.E.; Supervision- F.E.

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

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

ORCID IDs of the author

Firdevs Yıldız	0000-0002-6101-9798
Batuhan Gül	0009-0007-1772-5373
Fatih Ertam	0000-0002-9736-8068

REFERENCES

Aamir, M., Rizvi, S. S. H., Hashmani, M. A., Zubair, M., & Usman, J. A. . (2021). Machine Learning Classification of Port Scanning and DDoS Attacks: A Comparative Analysis. *Mehran University Research Journal of Engineering and Technology*. https://doi.org/10.22581/muet1982.2101.19

Ahmetoğlu, H., & Daş, R. (2021). Makine Öğrenmesi Yöntemleri Kullanarak Web Uygulama Saldırılarının Tespitinde Genetik Öznitelik Seçimi Yaklaşımı. *Türkiye Bilişim Vakfı Bilgisayar Bilimleri ve Mühendisliği Dergisi.* https://doi.org/10.54525/tbbmd.1018465

Akbal, E., Doğan, Ş., Tuncer, T., & Atalay, N. S. (2019). Adli Bilişim Alanında Ağ Analizi. *Bitlis Eren Üniversitesi Fen Bilimleri Dergisi*. https://doi.org/10.17798/bitlisfen.479303

AlZubi, A. A., Al-Maitah, M., & Alarifi, A. (2021). Cyber-attack detection in healthcare using cyber-physical system and machine learning techniques. *Soft Computing*. https://doi.org/10.1007/s00500-021-05926-8

Ashton, J. J., Young, A., Johnson, M. J., & Beattie, R. M. (2023). Using machine learning to impact on long-term clinical care: principles, challenges, and practicalities. *Pediatric Research*. https://doi.org/10.1038/s41390-022-02194-6

Aslan, O., & Yilmaz, A. A. (2021). A New Malware Classification Framework Based on Deep Learning Algorithms. *IEEE Access.* https://doi.org/10.1109/ACCESS.2021.3089586

Başlar, Y. (2020). Adli Bilişim Sürecinde Karşılaşılan Sorunlar ve Çözüm Önerileri. *Türkiye Barolar Birliği Dergisi, 32*(148), 47–76. Retrieved from https://app.trdizin.gov.tr/makale/TXpZeU5EUXpNdz09/adli-bilisim-surecinde-karsilasilan-sorunlar-ve-cozum-onerileri

Bi, Q., Goodman, K. E., Kaminsky, J., & Lessler, J. (2019). What is machine learning? A primer for the epidemiologist. *American Journal of Epidemiology*. https://doi.org/10.1093/aje/kwz189

Dina, A. S., & Manivannan, D. (2021). Intrusion detection based on Machine Learning techniques in computer networks. *Internet of Things (Netherlands)*, 16(August). https://doi.org/10.1016/j.iot.2021.100462

Ferrag, M. A., Maglaras, L., Moschoyiannis, S., & Janicke, H. (2020). Deep learning for cyber security intrusion detection: Approaches, datasets, and comparative study. Journal of Information Security and Applications. https://doi.org/10.1016/j.jisa.2019.102419

Hunt, R. (2012). New developments in network forensics-Tools and techniques. *IEEE International Conference on Networks*, ICON. https://doi.org/10.1109/ICON.2012.6506587

İnce, C., İnce, K., Hanbay, D., Üniversitesi, İ., İşlem, B., Başkanlığı, D., . . . Bölümü, M. (2021). Saldırı Tespit Sistemlerinde Sınıflandırma Yöntemlerinin Kıyaslanması. *Dergipark.Org.Tr*, (1), 1–11. Retrieved from https://dergipark.org.tr/en/pub/bbd/issue/59753/791939

Karaman, M. S., Turan, M., & Aydin, M. A. (2021). Yapay Sinir Ağı Kullanılarak Anomali Tabanlı Saldırı Tespit Modeli Uygulaması. *European Journal of Science and Technology.* https://doi.org/10.31590/ejosat.1115825

Kilincer, I. F., Ertam, F., & Sengur, A. (2022). A comprehensive intrusion detection framework using boosting algorithms. *Computers and Electrical Engineering.* https://doi.org/10.1016/j.compeleceng.2022.107869

Krishna Suryadevara, C. (2023). Issue 4 Diabetes Risk Assessment Using Machine Learning: A Comparative Study of Classification Algorithms. *International Engineering Journal For Research & Development, 8*(4). Retrieved from www.iejrd.com

Li, Y., & Liu, Q. (2021). A comprehensive review study of cyber-attacks and cyber security; Emerging trends and recent developments. *Energy Reports.* https://doi.org/10.1016/j.egyr.2021.08.126

Nancy Noella, R. S., & Priyadarshini, J. (2023). Machine learning algorithms for the diagnosis of Alzheimer and Parkinson disease. Journal of Medical Engineering and Technology. https://doi.org/10.1080/03091902.2022.2097326

Nurdina, A., & Puspita, A. B. I. (2023). Naive Bayes and KNN for Airline Passenger Satisfaction Classification: Comparative Analysis. *Journal of Information System Exploration and Research.* https://doi.org/10.52465/joiser.v1i2.167

Özekes, S., & Karakoç, E. N. (2019). Makine Öğrenmesi Yöntemleriyle Anormal Ağ Trafiğinin Tespit Edilmesi. *Düzce Üniversitesi Bilim ve Teknoloji Dergisi*. https://doi.org/10.29130/dubited.498358

Pallathadka, H., Wenda, A., Ramirez-Asís, E., Asís-López, M., Flores-Albornoz, J., & Phasinam, K. (2023). Classification and prediction of student performance data using various machine learning algorithms. *Materials Today*: *Proceedings*. https://doi.org/10.1016/j.matpr.2021.07.382

Qureshi, S., Tunio, S., Akhtar, F., Wajahat, A., Nazir, A., & Ullah, F. (2021). Network Forensics: A Comprehensive Review of Tools and Techniques. *International Journal of Advanced Computer Science and Applications.* https://doi.org/10.14569/IJACSA.2021.01205103

Radivilova, T., Kirichenko, L., Ageiev, D., & Bulakh, V. (2019). Classification methods of machine learning to detect DDoS attacks. *Proceedings of the 2019 10th IEEE International Conference on Intelligent Data Acquisition and Advanced Computing Systems: Technology and Applications, IDAACS 2019.* https://doi.org/10.1109/IDAACS.2019.8924406

Shafiq, M., Tian, Z., Sun, Y., Du, X., & Guizani, M. (2020). Selection of effective machine learning algorithm and Bot-IoT attacks traffic identification for internet of things in smart city. *Future Generation Computer Systems.* https://doi.org/10.1016/j.future.2020.02.017

Sharma, N., Sharma, R., & Jindal, N. (2021). Machine Learning and Deep Learning Applications-A Vision. *Global Transitions Proceedings.* https://doi.org/10.1016/j.gltp.2021.01.004

Shaukat, K., Luo, S., Chen, S., & Liu, D. (2020). Cyber Threat Detection Using Machine Learning Techniques: A Performance Evaluation Perspective. *1st Annual International Conference on Cyber Warfare and Security, ICCWS 2020 - Proceedings*. https://doi.org/10.1109/ICCWS48432.2020.9292388

Shaukat, K., Luo, S., Varadharajan, V., Hameed, I. A., & Xu, M. (2020). A Survey on Machine Learning Techniques for Cyber Security in the Last Decade. *IEEE Access.* https://doi.org/10.1109/ACCESS.2020.3041951

Wani, A. R., Rana, Q. P., Saxena, U., & Pandey, N. (2019). Analysis and Detection of DDoS Attacks on Cloud Computing Environment using Machine Learning Techniques. *Proceedings - 2019 Amity International Conference on Artificial Intelligence, AICAI 2019.*

https://doi.org/10.1109/AICAI.2019.8701238

Zhang, X., Chen, J., Zhou, Y., Han, L., & Lin, J. (2019). A Multiple-Layer Representation Learning Model for Network-Based Attack Detection. *IEEE Access.* https://doi.org/10.1109/ACCESS.2019.2927465

How cite this article

Yıldız, F., Gül, B., & Ertam, F. (2024). Network forensics analysis of cyber attacks on computer systems using machine learning techniques. *Acta Infologica, 8*(1), 34-50. https://doi.org/10.26650/acin.1444470

Acta Infologica
acin 2024, 8 (1), 51–59
DOI: 10.26650/acin.1431443

Research Article / Araştırma Makalesi

The Precarious Pirouette: Artificial Intelligence and Environmental Sustainability

Ronald Manhibi[1], Kudzayi Tarisayi[2]

[1]Bindura University of Science Education, Bindura, Zimbabve
[2]Stellenbosch University, Stellenbosch, South Africa

Corresponding author : Kudzayi Tarisayi
E-mail : kudzayit@gmail.com

ABSTRACT

The exponential ascension of artificial intelligence (AI) prompts profound inquiries concerning equitable access to its advantages versus environmental externalities. While trailblazing economies relish AI's benefits such as economic expansion and technological eminence, the colossal energy required to train and operate AI systems exacts a hefty toll on the environment, disproportionately burdening marginalized nations. This imbalanced paradigm epitomizes disparities of the digital divide, with impoverished nations bearing externalities while lacking access to innovations. This study aims to explore the intricate relationship between AI and environmental sustainability through a qualitative methodology encompassing a literature review and document analysis of industry practices and viewpoints. The findings unveil AI as a double-edged sword, with empirical analyses exposing its striking carbon emissions and resource depletion, which if left unchecked, could impede global decarbonization initiatives. However, AI also demonstrates strong potential for optimizing energy systems, predictive modelling, and advancing climate solutions if conscientiously developed. The study elucidates this conundrum and proposes responsible innovation pathways involving renewable energy adoption, enhanced efficiency, optimized hardware, carbon accounting, transparency, and legislative mindfulness. Integrating climate justice and digital divide perspectives illuminates avenues for steering AI's trajectory towards environmental stewardship and inclusive accessibility through proactive collaboration across sectors. Ultimately, collective wisdom will determine whether AI ushers in climate justice or injustice.

Keywords: Artificial intelligence, climate change, energy efficiency, carbon emissions, climate justice

Submitted	: 05.02.2024
Revision Requested	: 29.04.2024
Last Revision Received	: 01.05.2024
Accepted	: 21.05.2024
Published Online	: 28.06.2024

1. INTRODUCTION

The escalating power demands of artificial intelligence (AI) present an environmental justice dilemma: those least responsible for the consequences disproportionately bear the greatest burden. This study examines this inequity through the converging lenses of climate justice and digital divide frameworks. It contributes to the existing literature by providing a comprehensive analysis of AI's environmental impact, addressing the gap in understanding the implications of unchecked AI development on global decarbonization efforts and proposing solutions to mitigate the adverse effects.

The recent explosion of interest in AI, fuelled by user-friendly tools like ChatGPT, has led to surging demand for AI infrastructure and computing power. This growing adoption across industries carries a heavy energy cost that could soon overburden existing power grids. Projections portend that by 2030, AI could claim over 10% of the world's electrical bounty (Luccioni, 2020), its accompanying emissions imperilling crucial efforts to relinquish carbon. The veil of opacity shrouding developmental practices further derails any attempts at accountability (Dhar, 2020). To accurately assess impact and navigate promising trajectories, a thorough and all-encompassing inquiry is imperative.

While the theoretical framework explores climate justice and digital divide perspectives, the literature review delves into empirical analyses unveiling the striking carbon emissions and voracious resource consumption entwined with AI systems. For instance, training a single AI model can emit as much carbon as five cars in their lifetimes (Hao, 2019). The study aims to elucidate the potential environmental repercussions from unbridled AI progression and propose solutions to harmonize AI innovation with ecological boundaries.

Fostering this harmonization necessitates a multifaceted approach involving enhanced efficiency, renewable energy procurement, optimized hardware, carbon accounting, and supportive policies. However, technical solutions alone are insufficient; a collective shift in mindset prioritizing environmental stewardship over narrow self-interest is essential. As experts emphasize, progress should be redefined as holistic advancement benefiting humanity through climate justice and just transitions, rather than exclusive gains for the technocratic elite (Dobbe & Whittaker, 2019).

Promoting open access to intellectual capital over proprietary ownership offers potential pathways, as does sustainable investment in developing nations for inclusive participation (Gichuki, 2022). Nonetheless, avoiding another extractive paradigm depends on recognizing our shared future within planetary limits. With thoughtful intentions and wisdom, AI could unveil solutions to issues of inequity and ecological constraints if stewardship prevails over self-interest. As Rolnick et al. (2021) summarize, "With responsible innovation, AI can become integral to an energy future that balances decarbonisation, resilience, and accessibility." Achieving this necessitates transparency, accountability, and global cooperation centered on climate justice.

2. THEORETICAL FRAMEWORK

The burgeoning energy demands of artificial intelligence present an environmental justice dilemma: those least responsible for the consequences disproportionately bear the greatest burden. This paper examines this inequity through the converging lenses of climate justice and digital divide frameworks.

Climate justice perspectives emerge from environmental justice research examining how climate change detrimentally and disproportionately affects marginalized communities despite their negligible greenhouse gas contributions (Holifield et al., 2017). Historical exploitation and socioeconomic disenfranchisement leave these communities especially vulnerable with limited climate resilience resources. Consequently, some populations experience unequal environmental hazard exposure or unjust denial of environmental benefits, raising environmental justice concerns (Mohai et al., 2009). The climate justice movement arose from civil rights activism responding to these disparities, upholding fair treatment, meaningful participation, and the universal right to a healthy environment regardless of race, ethnicity, national origin, or income (MPCA, 2022).

Scholars have developed theoretical frameworks elucidating environmental justice notions to inform research and policy. Gee and Payne-Sturges (2004) delineate how social and environmental factors interact across levels to produce environmental health inequities. Individual factors like genetics and behaviours shape vulnerability, intersecting with socioeconomics, racism, and power imbalances, influencing differential hazard exposure through land use patterns. Grace et al. (2018) present four climate justice dimensions highly relevant to the ethical AI evolution: procedural, distributive, restorative, and social. Those most impacted by AI's ecological effects warrant enhanced participation in development decisions (procedural). AI's environmental costs levied on vulnerable communities require redress through climate financing and resource exchange (distributive). We must acknowledge and remedy damages from unchecked AI progress exacerbating climatic perils (restorative). Equitable sustainability demands reimagining progress as shared prosperity within ecological limits, not disproportionate gains for the technocratic class (social).

The digital divide theory further contextualizes these AI disparities. Dewan and Riggins (2005) probe technological, economic, and social access divides. AI leaders luxuriate in expansive digital capital while developing nations with minimal AI infiltration shoulder relatively cumbersome burdens from surging energy consumption and waste. van Dijk and Hacker's (2018) contemporary model expands divides across motivational, material, skills, and usage dimensions. Profit-driven AI evolution, resource-challenged infrastructures in marginalized communities, limited technical prowess beyond industrial centres, and usage discrepancies from algorithmic bias collectively perpetuate injustice. If left unchecked, artificial intelligence risks unravelling progress towards sustainable, clean energy. Resolving this quandary necessitates industry, government and civil society interventions championing socially responsible, climate just AI, rather than economic myopia. Integrating climate justice and digital divide perspectives illuminates avenues for AI to promote responsible ecological stewardship.

3. METHODOLOGY

In this scholarly exploration, a qualitative methodology encompassing a desk-based literature review and document analysis of purposively selected texts has been employed to delve into the environmental ramifications of artificial intelligence. A thorough examination of contemporary interdisciplinary literature on AI and sustainability has been conducted by engaging with academic databases. Publicly accessible corporate and NGO documents were scrupulously assessed utilizing a coding technique to evaluate industry practices and viewpoints concerning AI's ecological repercussions.

These harmonious research methods collectively cultivate an all-encompassing comprehension of AI's multifaceted potential—ranging from exacerbating emissions and energy demands to devising innovative solutions for climate change. The intention of this inquiry is to contribute to the inception of judicious policies and pioneering avenues that harness the inherent merits of AI while safeguarding against its unbridled progression undermining crucial decarbonisation endeavours.

Justification of the study

In this pivotal exploration, we delve into the profound environmental reverberations of artificial intelligence, whose unrestrained expansion may subsume ongoing sustainability initiatives without our deliberate guidance. Projections portend that by 2030, AI could claim over 10% of the world's electrical bounty (Luccioni, 2020), its accompanying emissions imperilling crucial efforts to relinquish carbon. The veil of opacity shrouding developmental practices further derails any attempts at accountability (Dhar, 2020). To accurately assess impact and navigate promising trajectories, a thorough and all-encompassing inquiry is imperative. This study weaves together the drapery of state-of-the-art understanding from multifarious experts, probing emergent priorities within the industry. Thus, it spawns an array of illuminating insights poised to shape policy and shepherd collective action. As humanity stands betwixt the gargantuan potential of AI and the relentless march of climate upheaval, discerning trade-offs becomes a vital endeavour. Guided by moral compass and sagacity, AI contains within it the power to manifest equitable abundance whilst harmonizing with Earth's ecological boundaries. This scholarly examination endeavours to illuminate pathways for sustainable progress by demystifying pathways conducive to sustainable development.

4. FINDINGS

The Enigmatic Dance of AI and Climate Change: A Double-Edged Sword

Beneath the burgeoning canopy of artificial intelligence lies an intricate tapestry of promise and apprehension, embroidered with questions of sustainability. A growing corpus of empirical analyses unveils the striking carbon emissions and voracious resource consumption entwined with the creation and deployment of AI systems (Hao, 2019; Hutson, 2022). This revelation stirs disquietude in the face of potential environmental repercussions from unbridled AI progression. As the boundless potential of artificial intelligence unfurls like the petals of a blossoming flower, it concurrently births an immense responsibility. We must vigilantly ensure that the beguiling allure of analytic prowess does not eclipse our unyielding commitment to fostering a harmonious existence upon this celestial sphere we fondly refer to as home.

In the intricate dance of artificial intelligence's endless evolution, three pivotal catalysts interweave to form the ever-evolving fabric of artificial intelligence: ground-breaking advancements in machine learning algorithms, an inexorable accrual of training data, and burgeoning computational power devoted to neural network optimization. Though indispensable for expansion, these driving forces demand staggering energy outlays—estimations suggest that training a lone natural language model expels carbon dioxide (CO2) on par with multitudes of transcontinental flights. Hao (2019) opines that training a single AI model can emit as much carbon as five cars in their lifetimes and that includes the manufacture of the car itself. The geographic locale of AI facilities bears resoundingly significant implications; renewable energy grids emit a mere fraction in comparison to their fossil fuel-reliant counterparts.

Moreover, the vast data centres cradling colossal learning models and complex neural networks upon which AI thrives generate an extraordinary thermal output. This profound heat signature is quelled through the practice of evaporative cooling—a technique capable of dissipating intense heat yet yielding copious amounts of water in return (Li, Wang, Shi, & Wang, 2023). The dependence upon water-based cooling methodologies intensifies the global demand for a resource in increasingly scarce supply. It becomes vitally imperative that we contemplate the technological advancements birthed by the artificial intelligence revolution within the contextual framework of our planet's overall well-being. Consequently, invoking the insightful wisdom of Large Language Model (LLM)-driven chatbots such as ChatGPT entails virtually summoning forth 500ml of freshwater. In 2022, this dynamic interplay culminated in Microsoft and Google witnessing a startling escalation in water consumption—34% and 20% respectively—translating to an astonishing 6.4 billion litres for Microsoft alone within that year.

The unsettling opacity shrouding the industry's sustainability practices grips our attention; however, glimmers of hope proliferate like stars awaiting discovery against the expanse of night. These guiding lights manifest as computable carbon accounting, energy-efficient hardware, and legislative mindfulness focused on emissions tracking and transparency. It behoves us to thoughtfully scrutinize the ecological footprints rent upon the Earth by AI systems, encourage responsible innovation, and wield a clarion call to compel corporations toward prioritizing environmental stewardship over myopic measurements devoid of tangible consequences.

A seminal 2019 study pierced the veil of energy usage and carbon emissions inherent to the cultivation of common natural language processing models, surmising that one such creation exhales an astounding 626,000 pounds of carbon dioxide equivalent – aligned with the lifetime exhalations of five average passenger vehicles (Strubell et al., 2019). Generative AI entities, such as ChatGPT, which can weave human-like textual tapestries, impose even greater resource voracity. The birth of ChatGPT proclaimed an emission of over 550 tons of carbon dioxide (CO2), resonating with the ecological impact of 550 roundtrip sojourns betwixt New York and San Francisco (Saenko, 2022). The carbon footprint permeating AI's utilization is similarly substantial; a single AI query emits a carbon dioxide (CO2) cloud four to five times vaster than its internet search counterpart – approximating 1gramme of carbon dioxide (CO2) (Jennifer, 2023). Bearing witness to over 1.5 billion ChatGPT queries in March 2023 alone, these emissions accumulate with bewildering rapidity (Jennifer, 2023). As corporations interweave AI into search engines and multifarious products, queries and commensurate emissions may propagate exponentially. Lying at the crux of these staggering digits is AI's prodigious hunger for computational power, heightening the call for energy resources. Calculating the labyrinthine algorithms of AI necessitates specialized hardware such as graphics processing units (GPUs), which voraciously consume 10-100 times more power than their conventional counterparts (Dhar, 2020).

The recent explosion of interest in artificial intelligence, fuelled by user-friendly tools like ChatGPT, has led to surging demand for AI infrastructure and computing power. This growing adoption across industries carries a heavy energy cost that could soon overburden existing power grids. For example, training a single AI model can consume as much electricity as 120 households use in an entire year (Freeman, 2023). Leading AI firms require more energy than major cities just to train their algorithms. Current GPUs and CPUs are designed for gaming, not optimized for AI's parallel computing needs. Training an AI model may require hundreds or thousands of servers operating in parallel and presenting an immense energy challenge. Data centres focused on AI already consume around 3% of global electricity, with cooling accounting for 40% of their power draw. Experts forecast the growth rate of processing power for AI to double from 6-7% to 15% annually as adoption expands. Yet energy is not the only bottleneck - network bandwidth to transfer massive training data between processors also strains capacity. According to Bill Haskell, CEO of Innventure, AI computing demand doubles every 3.4 months, outpacing Moore's Law (Lu, 2017). This exponential growth could overload power grids if left unchecked. Sustainable solutions are needed to supply sufficient energy and cooling for AI's voracious appetite.

The environmental toll also goes beyond electricity use. Manufacturing AI hardware and disposing of obsolete models creates substantial electronic waste (e-waste). For example, training a large neural network can produce over 626,000 pounds of carbon dioxide emissions, equivalent to flying about 650 roundtrips from New York to San Francisco. Energy consumption also has financial costs - estimates show that training complex AI models can incur millions in cloud computing bills. Companies must weigh these planetary and economic impacts against AI's benefits. Some firms are reducing power usage through efficiency, while innovators are developing optimized chipsets for AI's specialized computing needs. But much work remains to ensure AI fulfils its potential responsibly and sustainably.

Rapid AI adoption further exacerbates potential emissions. The AI Index Report revealed that from 2012 to 2018, the computational power needed for AI training increased by over 300,000 times - a rate surpassing efficiency improvements (Amodei & Hernandez, 2018). Consequently, by 2025, AI could produce up to 5.5% of global emissions (Rolnick et al., 2021), directly conflicting with urgent climate targets that call for rapid decarbonization.

The thriving societies engendering and profiting from artificial intelligence possess highly evolved economies with measures in place to buffer the effects of climate change or at least provide alternatives and mitigatory strategies. This digital chasm between developed and developing nations serves as a stark reminder that marginalized communities may not yet be poised to capitalize on the artificial intelligence revolution whilst remaining inevitably subject to its consequences on the environment. From an environmental justice standpoint, one cannot overlook that the ramifications of artificial intelligence systems will disproportionately impact already marginalized populations around the globe. It is disconcerting to acknowledge that these very communities, which shall bear the ecological repercussions of artificial intelligence most heavily, reside on the periphery of reaping its myriad benefits.

AI driven opportunities for resolving the climate crisis

Paradoxically, artificial intelligence unveils the potential to address the pressing climate crisis, through means such as enhancing energy efficiency, predictive modelling of extreme weather events, and optimizing transportation systems (Rolnick et al., 2019). For example, machine learning may cultivate "greener AI" by developing more energy-efficient neural network architectures and hardware devices (Cai et al., 2017). AI systems trained with climatic data, could more accurately predict the advent of floods, droughts, and additional calamities, thereby fostering adaptation and resilience. Likewise, intelligent grids powered by AI can streamline energy distribution and storage pathways. While AI's genesis inevitably carries environmental costs, its judicious application holds the key to expedite ecological remedies.

The orchestration of AI as an instrument for mitigating and adapting to climate change necessitates an interwoven collaboration among technology firms, governmental bodies, and the scientific community. The establishment of legislative frameworks that champion transparency and sustainability in AI systems is indispensable. Equally crucial is the provision of robust training data and subject matter acumen by climate researchers dedicated to cultivating ecologically specialized AI. The adoption of environmentally cognizant innovation practices, coupled with a commitment to open data exchange and computable carbon accounting, will further bolster ecologically beneficial AI development. Through prudent creative processes and mindful applications, artificial intelligence holds boundless promises in combating climatic challenges (Rolnick et al., 2019).

Perceived Solutions from the Literature

Nevertheless, there are solutions to alleviate AI's environmental impact. Such approaches encompass computable carbon accounting and auditing, which monitor emissions throughout the machine learning supply chain (Lacoste et al., 2019). Energy-efficient chipsets designed for AI tasks are under development, accompanied by optimized algorithms requiring less intensive training (Cai et al., 2017). Legislation mandating sustainability reporting and practices could also encourage tech companies to prioritize emissions reduction and clean energy procurement. Ultimately AI presents a double-edged sword for energy. While it facilitates remarkable efficiency and grid enhancements, unregulated progress may lead to increased electricity consumption and emissions. However, conscientious development with a focus on ecological consequences can guide AI's path towards climate solutions. Companies should balance techno-economic advancements against external factors, collaborating across industries to ensure AI's positive potential does not overshadow environmental responsibility. With responsible innovation, AI can become a crucial part of an energy future that harmonizes decarbonisation, resilience, and accessibility.

Moreover, feasible strategies exist to address AI's emissions issue. These methods involve enhancing data centre energy efficiency, creating optimized AI hardware, and researching energy-efficient algorithms and neural architectures (Lacoste et al., 2019). Companies can procure renewable energy and prioritize carbon-neutral facilities while implementing carbon accounting to track and disclose emissions across operations. Policies requiring transparency and emission reductions in tech sectors can fortify corporate accountability.

Fundamentally, AI developers must achieve equilibrium between rapid progress and ecological repercussions. If implemented conscientiously, AI offers immense potential to improve energy efficiency, grid management, and climate modelling. However, unrestrained development could exacerbate unsustainable emissions. To prevent surpassing a 1.5°C increase in temperature – the threshold for a climate disaster – requires mindful innovation and cross-sector collaboration. Ultimately, AI can either significantly combat or substantially aggravate today's environmental crises; our collective decisions will dictate its course.

Implications on the Global South

The escalating power demands of AI could strain African grids already burdened by reliability issues and rising demand. South Africa, possessing the continent's second-largest economy, experiences frequent rolling blackouts due to generation shortfalls (Eskom, 2022). Without substantial investments in new capacity, increased AI adoption could exacerbate these deficits. The halted construction of two coal plants in 2022 (Nzimande, 2022) highlights the challenges, while renewables like solar and wind offer promise but face storage and transmission barriers. Efficient, optimized AI systems that align with South Africa's energy masterplan could foster sustainable growth. However, unchecked expansion of data centres and GPU farms may drastically increase consumption, which is projected to nearly double by 2040 (DoE, 2019). Targeted policies and public-private collaboration are crucial for maximizing AI's benefits while minimizing environmental impact.

Furthermore, AI's mounting energy consumption endangers carbon reduction targets and compromises climate commitments such as South Africa's net-zero goal by 2050 (UNFCCC, 2021). Critics argue that the country relies excessively on carbon-intensive coal power, generating over 200 million tonnes of energy-related CO2 annually (Eskom, 2022). Introducing vast AI infrastructure could raise emissions further unless clean energy procurement and carbon offsets are adopted decisively. Energy-efficient AI systems may help mitigate these effects. African tech hubs like Kenya's Konza Technology City should prioritize renewable energy to limit emissions when expanding AI adoption (Gichuki, 2022). Colocation in more efficient grids like Ethiopia's dam-powered system can also reduce environmental impact. Unrestrained AI growth without optimizing efficiency and energy sources risks negating sustainability benefits from economic development. To ensure that AI supports resilience and inclusivity across Africa, proactive policies must align with each nation's climate objectives through comprehensive impact assessments.

5. DISCUSSION AND CONCLUSION

This study has illuminated artificial intelligence's (AI) complex, double-edged relationship with environmental sustainability. AI offers tremendous potential for optimizing energy systems, enhancing efficiency, and propelling climate solutions through sophisticated techniques like machine learning forecasting and automated management. However, the escalating computational demands, energy consumption, and carbon footprint of developing and operating AI also risk grave, unintended ecological consequences.

The empirical evidence exposing AI's environmental toll is disconcerting. Training complex language models like GPT-3 generates staggering CO2 emissions equivalent to hundreds of flights. AI hardware production yields concerning electronic waste levels. Currently, energy-intensive AI data centres consume nearly 3% of global electricity, predominantly from polluting fossil fuels due to opaque energy sourcing. Projections indicate AI could account for over 5% of worldwide emissions within years, conflicting with urgent climate goals. These alarming statistics illustrate how unchecked AI advancement prioritizing efficiency gains could prove catastrophic for sustainability. The immense energy requirements intrinsic to the machine learning pipeline risk overwhelming planetary boundaries if emissions continue unabated. While AI optimizations may yield localized environmental benefits, the existential climate change threat far outweighs limited efficiencies considering the sector's escalating emissions holistically.

Yet promising solutions could mitigate AI's toll if comprehensively adopted: enhanced energy efficiency, renewable energy procurement, optimized AI hardware, sustainable neural architectures, carbon accounting with transparent disclosure, and environmental sustainability policies and incentives. However, technology companies must prioritize responsibility and stewardship over profitability. Robust regulations focused on AI's ecological impacts, not just economic benefits, are essential for accountability. All stakeholders should holistically evaluate AI's advantages and externalities.

Individually, evaluating AI's overall impact is crucial beyond just efficiency gains. For sustainable mobility, emphasizing shared resources over private autonomy minimizes environmental degradation. More broadly, ethical AI prosperity within ecological boundaries requires abandoning extractive, consumption-driven models for responsible, socially conscious innovation centred on environmental justice. Indeed, intentionally developing AI prioritizing these principles demonstrates immense potential for advancing climate justice and equity. AI could empower marginalized communities in climate activism through locally sourced data analysis, support vulnerable nations' loss and damage claims by attributing extreme events, reduce bias excluding marginalized voices in climate science, prioritize equitable clean energy access, integrate Indigenous knowledge systems, and forecast climate migration patterns.

However, such transformative AI applications require grounding innovation in environmental stewardship, social responsibility, and sustainable development principles. Society faces a crossroads - will pursuing AI catalyse ecological renaissance and an equitable carbon-neutral transition? Or exacerbate climate catastrophe, perpetuating unjust impacts on marginalized populations? The path depends on prioritizing wisdom and conscience over accomplishment alone. In many ways, AI presents a modern Faustian bargain offering expedited advancement but potentially at the cost of ecological endurance. Anthropogenic climate change poses an existential threat, and dedicating AI's immense capabilities towards sustainability, restorative justice, and human rights rather than perpetuating extraction offers hope. Yet balance requires holistic, multistakeholder approaches championing ethics and transparency over self-interest and unsustainable growth.

AI's breathtaking possibilities also risk irreversible degradation when decarbonization remains humanity's preeminent imperative. How the global community develops and implements AI over coming decades will test whether our species can responsibly wield technological prowess for environmental restoration. With moral clarity supporting responsible, planetary-conscious development, pathways can emerge towards climate justice, ecological regeneration, and prosperity elevating our entire species equitably. Though hurdles remain, rallying collaboration towards ethical, existentially aware AI innovation represents our era's greatest opportunity for realizing harmonious prosperity within true ecological limits. The future awaits judicious stewardship.

Way Forward

Addressing both climate and equity challenges posed by artificial intelligence requires multi-faceted approaches, such as enhancing transparency, increasing renewable energy procurement (Dhar, 2020), optimizing efficiency, and implementing carbon accounting (Lacoste et al., 2019). However, technical solutions alone are not enough. A collective shift in mind set is essential. As experts emphasize, progress should be redefined as holistic advancement benefiting humanity through climate justice and just transitions, rather than exclusive gains for the technocratic elite (Dobbe & Whittaker, 2019). Promoting open access to intellectual capital over proprietary ownership offers potential, as does sustainable investment in developing nations for inclusive participation (Gichuki, 2022). Nonetheless, avoiding another extractive paradigm depends on recognizing our shared future within planetary limits. With thoughtful intentions and wisdom, AI could unveil solutions to issues of inequity and ecological constraints if stewardship prevails over self-interest. As Rolnick et al. (2021) summarize, "With responsible innovation, AI can become integral to an energy future that balances decarbonisation, resilience, and accessibility." This necessitates transparency, accountability, and global cooperation.

Conclusion

This study illuminated the double-edged implications of artificial intelligence in relation to climate change – serving as both a potential solution and a threat. AI can optimize energy systems, but its growing training requirements and carbon footprint challenge sustainability. Our analysis examined AI's emissions impact, from hardware production to power-intensive data centres, which could hinder crucial decarbonisation efforts without intervention. However, solutions exist such as efficiency enhancements, renewable energy procurement, transparent reporting, and policy-driven innovation with an emphasis on environmental stewardship.

Fundamentally, AI development should be a collaborative process for mutual benefit and adhere to planetary boundaries. If responsibly harnessed, AI could reveal pathways for an equitable transition within ecological limits. This necessitates recognizing our interconnected fates and existential stakes. As governments devise national AI strategies, research must continually direct progress towards climate justice. While challenging questions persist, our era demands conscientiousness and wisdom. With purposeful intent, humanity has the potential to shape AI's next chapter, steering our world towards justice in harmony with a sustainable planet.

Peer Review: Externally peer-reviewed.

Author Contributions: Conception/Design of Study- R.M., K.T.; Data Acquisition- R.M., K.T.; Data Analysis/Interpretation- R.M., K.T.; Drafting Manuscript- R.M., K.T.; Critical Revision of Manuscript- R.M., K.T.; Final Approval and Accountability- R.M., K.T.; Supervision- R.M., K.T.

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

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

ORCID IDs of the authors

Ronald Manhibi 0000-0001-6262-0297
Kudzayi Tarisayi 0000-0003-0086-2420

REFERENCES

Amaya, A., Bach, N., Phillips, C., Tejedor, A., Steinhaeuser, K., Lakkaraju, H., & Kammen, D. M. (2021). Social biases in solar geoengineering research. Nature Climate Change, 11(12), 1063-1067.

Amodei, D., & Hernandez, D. (2018). AI and Compute. OpenAI Blog.

Danish, S. S. (2023). AI-enabled energy policy for a sustainable future. Sustainability, 15, 7643.

Dhar, V. (2020). The carbon impact of artificial intelligence. Nature Machine Intelligence, 2, 423-425. https://doi.org/10.1038/s42256-020-0219-9

Dobbe, R., & Whittaker, M. (2019). AI and climate change: How they're connected, and what we can do about it. Climate Research.

Eskom. (2022). Environmental impact. https://www.eskom.co.za/OurCompany/SustainableDevelopment/EnvironmentalImpact/Pages/CDM_Projects.aspx

Eskom. (2022). Load shedding data. http://loadshedding.eskom.co.za/LoadShedding

Freeman, K. S. (2023). AI and energy consumption: Are we headed for trouble? Imore News.

Gee, G. C., & Payne-Sturges, D. C. (2004). Environmental health disparities: A framework integrating psychosocial and environmental concepts. Environmental Health Perspectives, 112, 1645–1653. https://doi.org/10.1289/ehp.7074

Gichuki, C. (2022). Konza Technopolis eyes AI, manufacturing and agriculture. The Exchange.

Hao, K. (2019). Training a single AI model can emit as much carbon as five cars in their lifetimes: Deep learning has a terrible carbon footprint. MIT Technology Review.

Holifield, R., Chakraborty, J., & Walker, G. (Eds.). (2017). The Routledge Handbook of Environmental Justice. Routledge.

Hutson, M. (2022). Measuring AI's Carbon Footprint: New tools track and reduce emissions from machine learning. IEEE Spectrum.

Jennifer, L. (2023). How big is the CO2 footprint of AI models? ChatGPT's emissions.

Kumari Rigaud, K., de Sherbinin, A., Jones, B., Bergmann, J., Clement, V., Ober, K., Schewe, J., Adamo, S., McCusker, B., Heuser, S., & Midgley, A. (2018). Groundswell: Preparing for Internal Climate Migration. World Bank.

Lacoste, A., Luccioni, A., Schmidt, V., & Dandres, T. (2019). Quantifying the carbon emissions of machine learning. Workshop on Tackling Climate Change with Machine Learning.

Li, R., Wang, W., Shi, Y., & Wang, P. (2023). Advanced Material Design and Engineering for Water-Based Evaporative Cooling. Advanced Materials, 2209460. https://doi.org/10.1002/adma.202209460

Lu, C. (2017). AI, native supercomputing and the revival of Moore's Law. APSIPA Transactions on Signal and Information Processing, 6, E9. doi:10.1017/ATSIP.2017.9

Luque-Ayala, A., Chapman, A., Scuriatti, C., Maia, L., Hunter, C., & Peres, W. (2021). Digital territories: Google Maps as a political technique in the reorganization of urban energy systems. Energy Research & Social Science, 79, 102138. https://doi.org/10.1016/j.erss.2021.102138

Minnesota Pollution Control Agency (MPCA). (2022). Environmental justice framework.

Mohai, P., Pellow, D., & Roberts, J. T. (2009). Environmental justice. Annual Review of Environment and Resources, 34, 405-430. https://doi.org/10.1146/annurev-environ-082508-094348

Morello-Frosch, R., & Lopez, R. (2006). The riskscape and the color line: Examining the role of segregation in environmental health disparities. Environmental Research, 102(2), 181-196. https://doi.org/10.1016/j.envres.2006.05.007

Nzimande, B. (2022). Minister Blade Nzimande: Abandonment of Kusile units 5 and 6 construction projects. Department of Higher Education, Science and Innovation.

Risser, M. D., & Wehner, M. F. (2017). Attributable human-induced changes in the likelihood and magnitude of the observed extreme precipitation during Hurricane Harvey. Geophysical Research Letters, 44(24), 12,457-12,464. https://doi.org/10.1002/2017GL075888

Rolnick, D., Donti, P.L., Kaack, L.H., Kochanski, K., Lacoste, A., Sankaran, K., Ross, A.S., Milojevic-Dupont, N., Jaques, N., Waldman-Brown, A., Luccioni, A., Maharaj, T., Sherwin, E.D., Mukkavilli, S.K., Körding, K.P., Gomes, C., Ng, A.Y., Hassabis, D., Platt, J.C., ... Bengio, Y. (2021). Tackling Climate Change with Machine Learning. arXiv.

Rolnick, D., Philip, L., Kaack, L., Lacoste, A., Luccioni, A. (2019). Trends and applications in climate informatics. Journal of Parallel and Distributed Computing, 134, 141-150. https://doi.org/10.1016/j.jpdc.2019.08.006

Saenko, K. (2022). The huge carbon footprint of AI algorithms: Machine learning has a disastrous environmental impact. Boston Globe.

Schwartz, R., Dodge, J., Smith, N.A., & Etzioni, O. (2019). Green AI. arXiv preprint arXiv:1907.10597.

Strubell, E., Ganesh, A., & McCallum, A. (2019). Energy and policy considerations for modern deep learning research. Thirty-Fourth AAAI Conference on Artificial Intelligence.

Strubell, E., Ganesh, A., & McCallum, A. (2019). Energy and Policy Considerations for Deep Learning in NLP. arXiv.

Whyte, K. P. (2018). What do Indigenous knowledges have to offer climate change research? In S. Díaz, J. Settele, E. Brondízio, & H. T. Ngo (Eds.), The IPCC and Indigenous Peoples (pp. 57-59). Intergovernmental Panel on Climate Change.

Wylie, S., Jalbert, K., Dosemagen, S., & Ratto, M. (2022). Advancing climate justice with community-based air quality monitoring and machine learning in California's Imperial County. Geo: Geography and Environment, 9(1), e00107.

Zhou, H., Liu, Q., Yan, K., & Du, Y. (2021). Deep learning enhanced solar energy forecasting with AI-driven IoT. Wireless Communications and Mobile Computing, 2021. https://doi.org/10.1155/2021/9249387

How cite this article

Manhibi, R., & Tarisayi, K. (2024). The precarious pirouette: Artificial intelligence and environmental sustainability. *Acta Infologica, 8*(1), 51-59. https://doi.org/10.26650/acin.1431443

İSTANBUL UNIVERSITY PRESS

Acta Infologica
acin 2024, 8 (1), 60–70
DOI: 10.26650/acin.1475658

Research Article / Araştırma Makalesi

Diagnosis of Internal Frauds using Extreme Gradient Boosting Model Optimized with Genetic Algorithm in Retailing

Aytek Demirdelen[1], Pelin Vardarlıer[2], Yurdagül Meral[2], Tuncay Özcan[3]

[1]İstanbul Medipol University, Institute of Social Sciences, PhD Program in Business Administration, İstanbul, Türkiye
[2]İstanbul Medipol University, Faculty of Business and Management Sciences, Department of Human Resources Management, İstanbul, Türkiye
[3]İstanbul Technical University, Faculty of Business Administration, Department of Management Engineering, İstanbul, Türkiye

Corresponding author : Tuncay Özcan
E-mail : tozcan@itu.edu.tr

ABSTRACT

Fraud is one of the most vital problems that can lead to a loss of organizational reputation, assets and culture. It is beneficial for companies to anticipate possible fraud in order to protect both culture and company assets. The aim of this study is to provide a fraud detection model using classification and optimization algorithms. For this purpose, this study proposes a novel hybrid model called XGBoost-GA to enhance the prediction quality for cashier fraud detection in retailing. In the proposed model, the genetic algorithm (GA) is used to optimize the parameters of extreme gradient boosting (XGBoost) model. The proposed XGBoost-GA model is compared with XGBoost, logistic regression (LR), naive bayes (NB) and k-nearest neighbor (k-NN) algorithms. The performance comparison is presented with a case study with the actual data taken from a grocery retailer in Turkey. Numerical results showed that the proposed hybrid XGBoost-GA model produces higher accuracy, recall, precision and F-measure than other classification algorithms. In this context, the use of proposed model in fraud detection will be beneficial for companies to use their resources effectively. Classification algorithms will also accelerate organizations in terms of detecting the possible damage of fraud to company assets before it grows.

Keywords: Fraud Detection, Retailing, Machine Learning, Extreme Gradient Boosting, Genetic Algorithm

Submitted : 01.05.2024
Accepted : 14.05.2024
Published Online : 03.06.2024

1. INTRODUCTION

The short story Minority Report, written by Philip K. Dick in 1956, was brought to the big screen by Steven Spielberg in 2002 with the same name. The film reached large audiences with its detection, prediction, actors and special effects of crime and criminals and raised questions in the minds. In this movie, the crimes are seen by the oracles before they happen and can be prevented beforehand. In this context, we are depicted in a world where not only the detection of crime but also foresight is dominant in the fight against crime and the detection is thought to be made before the action. This depiction is partially possible today, thanks to the advanced computer. In this context, early detection of fraud plays a critical role. A proactive approach in early detection is beneficial in protecting companies' assets (Erol, 2016).

The retail sector, which has a high transaction volume, also takes measures with audit activities to detect abuses it faces. Loss prevention and internal audit processes, which are among the business processes of the retail industry, have also undergone digital transformation and had to be reshaped. This transformation has gained importance for companies in terms of early detection of fraud and reduction of possible damage.

The impact of digital transformation on business processes is naturally reflected in audit methodologies. Audit processes are being reshaped, data analytics, automation etc. test methods have started to be used in field applications of process controls. The world of auditing resulting from Industry 4.0 will automate existing procedures, expand audit scopes, save time and increase the quality of audit assurance (Esmeray, 2018).

In this context, artificial intelligence, machine learning and advanced data analytics applications, which are part of the change created by the elements of Industry 4.0, have enabled the application of a proactive approach in the detection of frauds. Therefore, this study aims to present a hybrid model using extreme gradient boosting (XGBoost) and genetic algorithm (GA) for cashier fraud detection in the retail sector.

The remainder of this study is as follows: In Section 2, a review of the literature studies on the fraud detection problem is provided. In Section 3, gradient boosting algorithm, genetic algorithm and the proposed XGBoost-GA model are introduced. In section 4, the performance analysis of the developed models for the cashier fraud detection are given. In Section 5, the findings and conclusions are presented.

2. LITERATURE REVIEW

Fraud detection is an interesting research topic for both practitioners and academics. In the literature, there are many studies on credit card, telecommunication, tax/customs and insurance fraud. These studies can be summarized as follows:

Hanagandi et al. (1996) created a scoring system for credit card fraud detection by combining a radial basis function network with a density-based clustering algorithm. In this study, artificial neural networks were used to create the model. Shen et al. (2007) investigated the effectiveness of applying classification models to credit card fraud detection problems. In this study, the performance of decision tree, artificial neural network and logistic regression algorithms in fraud detection was tested. Seyedhossein & Hashemi (2010) proposed a method based on the creation of customer profiles for credit card fraud detection. The focus of this study is on cases of fraud that are not detected at the transaction level. In the proposed method, daily amounts spent on an individual credit card account were examined by time series analysis. Bhattacharyya et al. (2011) compared the performance of logistic regression, random forest, and support vector machines for credit card fraud detection. Sahin & Duman (2011) analyzed the decision tree and support vector machines (SVM) for credit card fraud detection in their study. Perols (2011) used six popular statistical and machine learning models to detect financial statement fraud. Numerical results showed that logistic regression and support vector machines performed well according to artificial neural network, C4.5 and stacking. Mahmoudi & Duman (2015) used a linear separator called Fisher Discriminant Function (FDA) in their study to detect credit card frauds. Vlasselaer et al. (2015) proposed a new approach called APATE (Anomaly Prevention using Advanced Transaction Exploration) to detect fraudulent credit card transactions in online stores. In the proposed approach, the characteristics of incoming transactions and the time since the last shopping date, shopping frequency and shopping amount derived from customer spending history are combined. Then, using this data from the network of credit card holders and businesses, a time-dependent risk score was derived for each network object. Renjith (2018) used the support vector machines method to detect fraudulent sellers in online sales areas. Additionally, it was stated that the algorithm used would not be sufficient to make a decision for a new seller whose historical data is not available. Shukur & Kurnaz (2019) used Logistic Regression, Artificial Neural Networks and K-Nearest Neighbor methods for credit card fraud detection. In this study, numerical results showed that Logistic Regression had the best classifier performance and K-Nearest Neighbor algorithm had the worst classifier performance. Nadim et al. (2019) compared different machine learning methods for credit card fraud detection according to performance criteria such as accuracy, precision, sensitivity and specificity. As a result of this comparison, it was revealed that logistic regression, random forest and XG-Boost algorithms gave the best results according to the accuracy rate, and random forest and XG-Boost algorithms gave the best results according to the cost criterion. Niu et al. (2019) performed the performance analysis of supervised and unsupervised learning techniques for credit card fraud detection using AUC-ROC curves. As a result of the study, among supervised learning techniques, the XG-Boost classifier was the most successful method with an accuracy rate of 98.94%, while the Decision Tree classifier was the least successful method with an accuracy rate of 95.42%. Pehlivanli et al. (2019) used support vector machine and artificial neural network methods to detect fraudulent purchases in the retail industry. In this study, different kernel functions were tested for the support vector machine and it was revealed that the support vector machine performed better. Varmedja et al. (2019) used logistic regression, Naive Bayes and random forest algorithms for the detection of credit card fraud with an original data set and compared the performance of these methods according to precision, sensitivity and accuracy values. Performance analysis showed that the most successful method among the methods used was the random forest algorithm. Walke (2019) compared the performance of supervised learning techniques and unsupervised learning techniques in solving the fraud detection problem. As a result of this study, it was observed that supervised learning techniques were more successful than unsupervised learning techniques. Askari & Hussain (2020) proposed a hybrid algorithm based on fuzzy logic and decision tree for fraud detection in online transactions. Parmar et al. (2020) used many different classification algorithms in detecting credit card fraud and tested the performance of these algorithms with the accuracy rate and F-score obtained from the confusion matrix. As a result of this analysis, it was concluded that the best results were obtained with the K-Nearest Neighbor method and the worst results were obtained with the Logistic Regression method. Roseline et al. (2022) performed pattern recognition on the card transaction database to detect credit card fraud and used machine learning algorithms to identify suspicious transactions. In this study, the class imbalance problem was addressed and machine learning algorithms such as Naive bayes, SVM, ANN and LSTM were used. Yi et al. (2023) presented a machine learning method integrated with the Egret Swarm Optimization Algorithm (ESOA), a meta-heuristic algorithm for financial fraud detection. Huang et al. (2024) proposed a machine learning-based K-means clustering method to improve the accuracy and efficiency of financial fraud detection.

When the literature studies are examined, it is seen that fraud detection is considered as a binary classification problem in many studies and machine learning-based algorithms are widely used to solve the problem. At the same time, a significant part of the literature studies focuses on the problem of credit card fraud in sectors such as banking and insurance. On the other hand, studies on fraud detection in the retail and e-commerce sector are limited. In this direction, this study aims to develop a hybrid model called XGBoost-GA to improve the classification accuracy for cashier fraud detection in retailing.

3. METHODOLOGY

In this section, extreme gradient boosting (XGBoost), genetic algorithm (GA), the proposed hybrid XGBoost-GA model and the performance metrics used to compare classification models are introduced.

3.1. EXTREME GRADIENT BOOSTING (XGBOOST)

Extreme gradient boosting (XGBoost) is a high-performance classification algorithm based on decision trees. XG-Boost classifier introduced by Chen and Guestrin (2016). This model combines weak classifiers with stronger classifiers and at each iteration.

Suppose that $D = (x_i, y_i)$ denotes a data set with n samples and m attributes. Here, x_i denotes the input data and y_i denotes the class label for the ith sample. The predicted class labels can be calculated using Equation (1):

$$\hat{y}_i = \Sigma_{k=1}^{K} f_k(x_i),\ f_k \in F \tag{1}$$

In Equation (1), $\hat{y}_i$ indicates the predicted class label for the ith sample, K is the number of trees, $f_k\ (x_i)$ denotes the predicted score of the kth tree, F is denoted by the space of all regression trees.

The objective function of XGBoost is described using Equation (2) (Chen et al., 2018).

$$\begin{gathered} F_{obj}(\theta) = L(\theta) + \Omega(\theta) \\ where\ L(\theta) = l(\hat{y}_i, y_i),\ \Omega(\theta) = \alpha T + 1/2\varepsilon w^2. \end{gathered} \tag{2}$$

In Equation (2), the objective function consists of two components. The first component reflects is differentiable convex loss function, while the second term is a regularized term that penalizes complex models. Additionally, T

indicates the number of leaves in the tree, α denotes the learning rate, ε is a regularized parameter and w is the weight of the leaves.

The objective function can be rewritten using Equation (3).

$$L(\theta) = \Sigma_{i=1}^{n} l\left(y_i, \hat{y}_i^{(t-1)} + f_i(x_i)\right) + \Omega(\theta) \tag{3}$$

The optimization goal is to construct a tree structure that minimizes the objective function (prediction error) in each iteration.

3.2. GENETIC ALGORITHM

Genetic algorithm (GA) is a population-based stochastic metaheuristic algorithm. This algorithm was inspired by Darwin's theory of evolution and was first introduced by Holland (1975). Genetic algorithm tries to find good solutions for NP-Hard optimization problems in a reasonable time by using parent selection, crossover and mutation operators. In this direction, GA is widely used in solving the parameter optimization problem of classification algorithms.

The basic principle of GA is the survival of stronger individuals to reach better solutions and the creation of better individuals from these individuals using crossover and mutation operators. In this algorithm, each individual represents a candidate solution. The steps of basic GA can be summarized as follows:

Step 1: Creating the coding structure that represents individuals Depending on the decision variable of the optimization problem, binary, discrete, permutation or real value coding can be used.
Step 2: Generating the random initial population
Step 3: Calculating the fitness value of each individual
Step 4: Selection of parents to create new individuals
Step 5: Creating new individuals using crossover and mutation operators
Step 6: Repeating Steps 3-5 until stopping criterion such as maximum number of iterations, maximum duration or target objective function value is satisfied.

3.3. PROPOSED APPROACH: XGBOOST-GA HYBRID MODEL

The XGBoost algorithm, like other classification algorithms, has a large number of parameters. The optimization of these parameters significantly improves the classification accuracy compared to the default parameters. In the parameter optimization problem, the parameters of the XGBoost model are decision variables.

These variables are presented in Table 1.

Performance metrics for classification accuracy can be used as the objective function of the optimization model. The performance metrics of the classifier can be calculated using the confusion matrix, whose general form is presented in Table 2.

Table 1. The parameters of XGBoost used in parameter optimization

Parameter	Data Type	Description
alfa (a)	Real	Learning rate
gamma (g)	Integer	The minimum loss value to make a partition on a leaf node of the tree
max depth (*md*)	Integer	The maximum number of splits
min child weight (*mcw*)	Integer	The minimum sum of sample weight in a child node
max delta step (*mds*)	Real	A measure of regularization
subsample (*ss*)	Real	A parameter used to control overfitting
n estimators (*n*)	Integer	The number of trees

Table 2. General form of confusion matrix for binary classification problem

		Predicted Class	
		C_0	**$\sim C_0$**
Actual Class	**C_0**	True Positive (TP)	False Negative (FN)
	$\sim C_0$	False Positive (FP)	True Negative (TN)

At this point, the most important of these performance metrics and their calculation formulas are as follows.

$$Accuracy\ Rate = \frac{TP+TN}{TP+TN+FN+FP} \tag{4}$$

$$Precision = \frac{TP}{TP+FP} \tag{5}$$

$$Recall = \frac{TP}{TP+FN} \tag{6}$$

$$F_{measure} = \frac{2 * Precision * Recall}{Precision + Recall} \tag{7}$$

In the parameter optimization problem, F-measure is used as the objective function due to the unbalanced distribution of the class label. The genetic algorithm is used to solve the parameter optimization problem. The pseudocode of the proposed hybrid XGBoost-GA model is presented in Table 3.

Table 3. The pseudo-code of the proposed XGBoost-GA hybrid method

```
Load the dataset
Divide data into training and testing datasets
Initialize GA parameters (max number of iteration=1000, population size=20, elit ratio=0.1, mutation probability= 0.05, crossover probability=0.8, parents portion= 0.3)
Define a,γ, md, mcw, mds, ss and n parameters of XGBoost randomly
Define $\text{Fmeasure}_{best}$=0
Set $i$=1
While ($i$ < max number of iteration)
    Calculate F-measure of XGBoost by using training data
    Set fitness function=$\text{MAPE}_i$
    Calculate $\text{F-measure}_i$
    if ($\text{Fmeasure}_i$ >$\text{Fmeasure}_{best}$) then
        Update $\text{Fmeasure}_{best}$= $\text{Fmeasure}_i$
        Update a,γ, md, mcw, mds, ss and n
    end if
    i=i+1
end while
Create a NGBM(1,1) model with finalized a,γ, md, mcw, mds, ss and n parameters
Calculate Fmeasure value of the testing data
```

4. APPLICATION

In this section, the application steps and performance analysis of the proposed approach are presented with dataset taken from a retail chain in Turkey.

4.1. DATA SET

In this study, real-life data from a retail chain in Turkey is used to detect cashier fraud. This dataset consists of 10 attributes and 13520 cashier transactions. The attributes included in the data set are presented in Table 4.

Table 4. The overview of the attribute in the dataset

Attribute	Type	Distinct Value
Transaction Type	Nominal	3
City	Nominal	5
Time Period	Nominal	3
Gender	Nominal	2
Age	Numeric	31
Seniority	Numeric	87
Position	Nominal	7
Marital status	Nominal	3
Category	Nominal	5
IsFraud	Binary	2

The descriptive statistics of numeric and nominal attributes in the dataset are given in Table 5 and Table 6, respectively.

Table 5. Descriptive statistics of numeric attributes in the dataset

Attribute	Average	Std. Dev.	Minimum	Q1	Median	Q3	Maximum
Age (year)	30.559	6.766	18	26	29	35	52
Seniority (month)	51.320	43.790	4	16	44	67	226

Table 6. Descriptive statistics of nominal attributes in the dataset

Attribute	Possible Values and Percentages
Transaction Type	Price Check (60%), Cancel Line (21%), Cancel Receipt (19%)
City	Istanbul (44.5%), Ankara (6%), Izmir (6%), Adana (4.7%) and Other (38%)
Time Period	Midday (51.5%), Night (30.5%), Morning (18%),
Gender	Female (59%), Male (41%)
Position	Cashier (42%), Staff (26%) and Other (32%)
Marital status	Single (45%), Married (27%) and Uknown (22%)
Category	Food (69%), Fresh food (13%), Bazar (12%) , Textile (4%) and Electronics (2%)
IsFraud	Fraud (10%), Non-Fraud (90%)

4.2. DETECTION OF CASHIER FRAUD WITH PROPOSED APPROACH

This study aims to propose a hybrid model using extreme gradient boosting (XGBoost) optimized with the genetic algorithm (GA) for cashier fraud detection in the retail sector. Additionally, the performance of this proposed model is evaluated by comparing it with XGBoost, logistics regression, naïve bayes, k-nearest neighbor (k-NN). The dataset is divided into 80% training data and 20% testing data for validation of the classification models.

In the parameter optimization model, F-measure is used as the fitness function of the GA due to the unbalanced distribution of the class label, as can be seen in Table 6. The classification models and the parameter optimization problem are programmed using the Python language and related packages. The developed codes are run on a PC with Intel ® Core™ i5-7200U CPU at 2.71 GHz, 8GB RAM, and Windows 10 Pro.

The parameters of genetic algorithm for the optimization model are as follows: the maximum iteration number is 1000, the population size is 20, elitism ratio is 0.1, the probability of crossover is 0.8 and the probability of mutation is 0.05. The default values are used for the other variables such as crossover and selection function. 10 independent replications are carried out using this parameter set. In the genetic algorithm, the F-measure converges very fast to a stationary point, as can be seen in Fig. 1. Fig. 1 shows the value of fitness function according to the number of iterations.

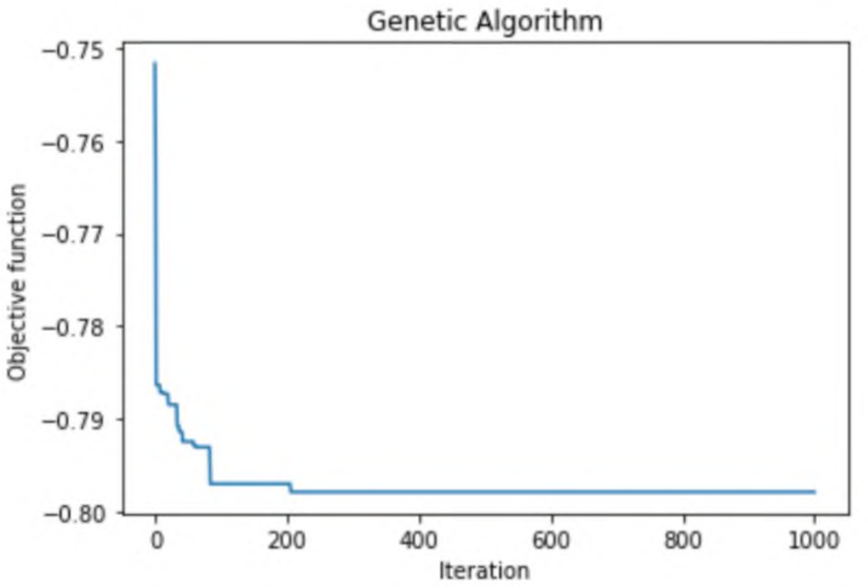


Figure 1. Parameter optimization process with genetic algorithm

In the proposed XGBoost-GA model, the F-measure is found to be 79.8% for the training data. The parameters of the XGBoost model are alfa=0.9857, gamma=1, max depth=10, min child weight =3, max delta step=1, subsample=0.8869 and the number of trees=95.

The performance analysis of the classification models for training data are given in Table 7 and Fig. 2.

Table 7. Performance analysis of the classification models for training data

Model	F-measure	Accuracy	Precision	Recall
XGBoost-GA	0.798	0.962	0.817	0.780
XGBoost	0.752	0.948	0.826	0.691
Logistics Regression	0.090	0.907	0.658	0.048
Naive Bayes	0.393	0.916	0.628	0.286
k-NN	0.718	0.949	0.772	0.672

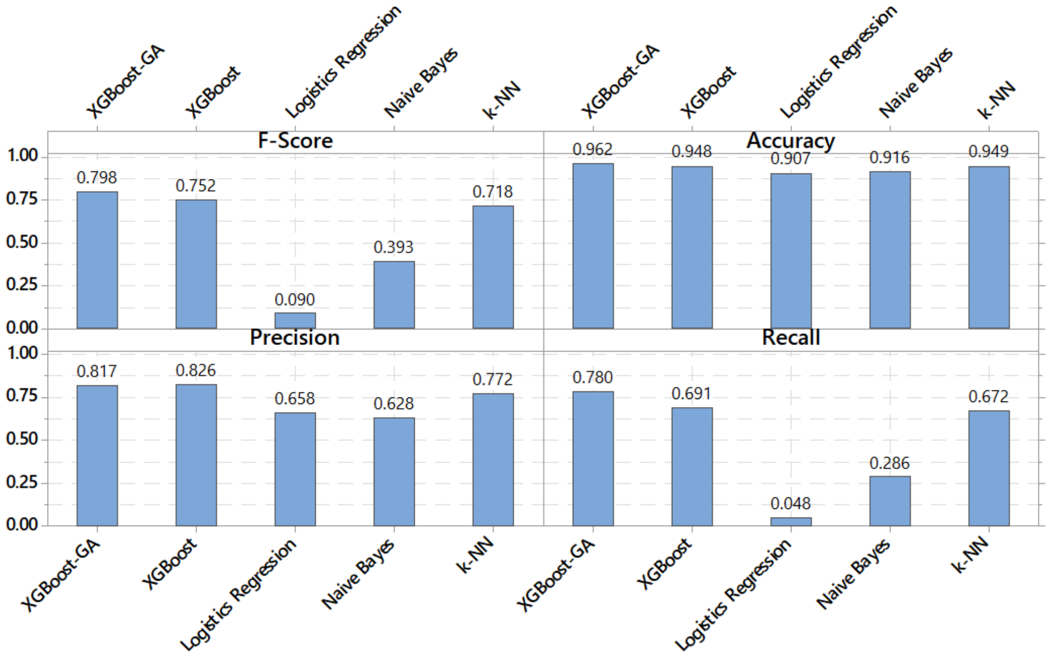


Figure 2. Performance metrics of the classification models for the testing data set

The performance analysis of the classification models for testing data are presented in Table 8 and Fig. 3.

Table 8. Performance analysis of the classification models for testing data

Model	F-measure	Accuracy	Precision	Recall
XGBoost-GA	0.733	0.959	0.718	0.749
XGBoost	0.712	0.951	0.780	0.655
Logistics Regression	0.050	0.901	0.389	0.027
Naive Bayes	0.381	0.912	0.608	0.278
k-NN	0.644	0.938	0.727	0.578

According to the results in Table 7, the proposed hybrid XGBoost-GA model has the maximum F-measure value of 79.8% whereas LR model has the lowest F-Measure of 5%. The numerical results also indicate that parameter optimization improves the classification accuracy of XGBoost model.

As can be seen from Table 8 and Fig. 3, the proposed XGBoost-GA model has a F-measure of 73.3% for validation data. The proposed hybrid XGBoost-GA model produces higher accuracy, recall, precision and F1-score than other classification algorithms such as XGBoost, logistics regression, naïve bayes, k-NN.

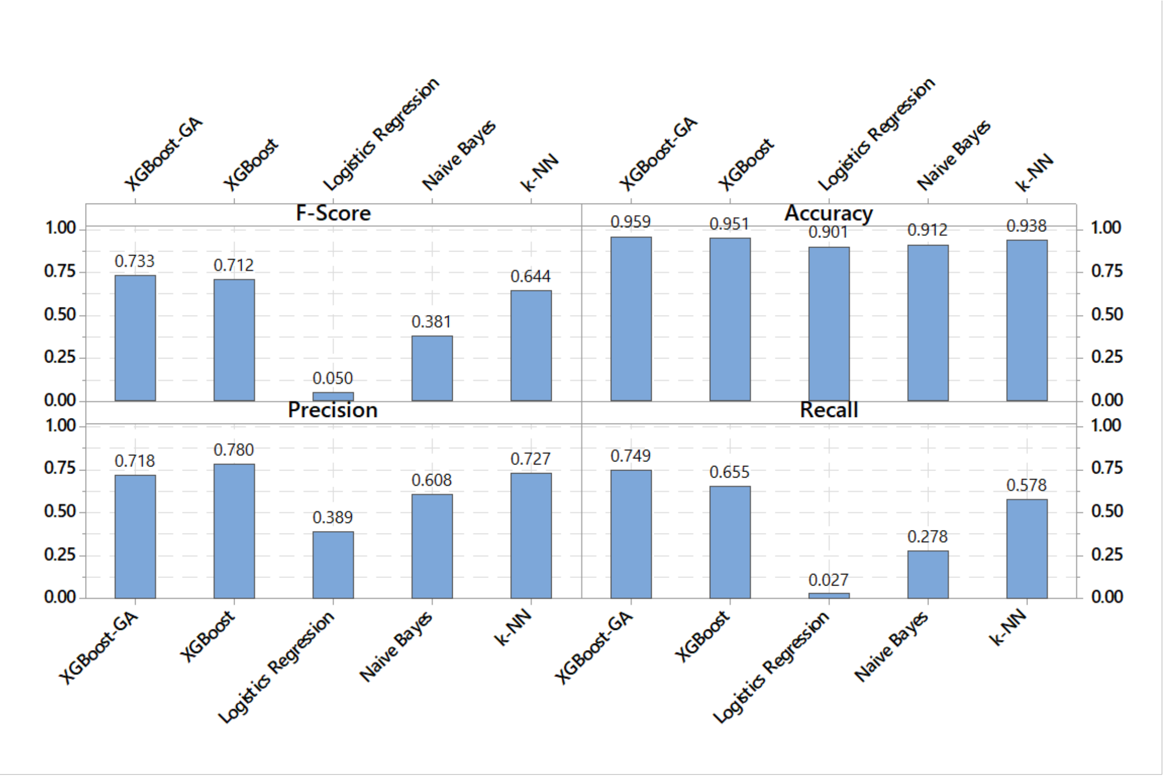


Figure 3. Performance measures of the classification models for the testing data set

5. CONCLUSIONS

With the digitalizing world, efforts to prevent and detect the increase in fraudulent activities have accelerated. While fraud detection performs digitally in sectors such as finance, banking and insurance, it is observed that applications in the retail sector have only just begun and analytical approaches are used to a limited extent in solving the problem. At the same time, while there are many studies in the literature on fraud detection for the financial sectors, there are very few studies for the retail sector.

Failure to detect and prevent fraudulent activities causes, businesses to experience significant financial losses. Gee and Button (2019) stated that financial losses resulting from fraud cases around the world are more than 80% of the UK's Gross Domestic Product. In another study published by ACFE (Association of Certified Fraud Examiners), 91 cases of fraud in the retail sector were examined and the median value of financial losses resulting from these cases was determined to be $85000.

One of the sources of fraud in the retail industry is store personnel. Accordingly, in this study, the cashier fraud detection problem is addressed with real-life data taken from a retail chain. To solve this problem, a hybrid approach is developed using extreme gradient boosting (XGBoost) and genetic algorithm (GA). In this approach, genetic algorithm is used to optimize the parameters of the XGBoost algorithm. The performance of the developed approach is compared with basic classification algorithms such as default XGBoost, logistic regression, naive bayes and k-nearest neighbor. Numerical results showed that the proposed approach has better performance than other classification algorithms for training and testing data. Also, with its high accuracy rate and F-measure, the proposed approach offers an effective solution for detecting cashier fraud in retail.

In future studies, the performance can be increased by adding new attributes to the proposed model. Additionally, approaches such as SMOTE can be used to solve the class imbalance problem. Different metaheuristic algorithms or Bayesian optimization can be used to solve the parameter optimization problem.

Peer Review: Externally peer-reviewed.

Author Contributions: Conception/Design of Study- A.D., P.V., Y.M., T.Ö.; Data Acquisition- A.D., T.Ö.; Data Analysis/Interpretation- A.D., T.Ö.; Drafting Manuscript- A.D., P.V., Y.M., T.Ö.; Critical Revision of Manuscript- A.D., P.V., Y.M., T.Ö.; Final Approval and Accountability- A.D., P.V., Y.M., T.Ö.; Supervision- T.Ö.

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

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

ORCID IDs of the authors

Aytek Demirdelen	0000-0002-6005-4604
Pelin Vardarlıer	0000-0002-5101-6841
Yurdagül Meral	0000-0001-9244-1994
Tuncay Özcan	0000-0002-9520-2494

REFERENCES

Askari, S. M. S., & Hussain, M. A. (2020). IFDTC4. 5: Intuitionistic fuzzy logic based decision tree for E-transactional fraud detection. *Journal of Information Security and Applications, 52*, 102469.

Chen, T., & Guestrin, C. (2016, August). Xgboost: A scalable tree boosting system. In *Proceedings of the 22nd acm sigkdd international conference on knowledge discovery and data mining* (pp. 785-794).

Chen, Z., Jiang, F., Cheng, Y., Gu, X., Liu, W., & Peng, J. (2018, January). XGBoost classifier for DDoS attack detection and analysis in SDN-based cloud. In *2018 IEEE international conference on big data and smart computing (bigcomp)* (pp. 251-256). IEEE.

Erol, S. (2016). *Hile denetiminde proaktif yaklaşımlar* (Master's thesis, Sosyal Bilimler Enstitüsü).

ESMERAY, A. (2018). BİLİŞİM TEKNOLOJİSİNDEKİ GELİŞMELERİN MUHASEBE DENETİMİNE KATKISI. *Muhasebe Bilim Dünyası Dergisi, 20*, 294-309.

Gee, J., & Button, M. (2019). The financial cost of fraud 2019: The latest data from around the world.

Hanagandi, V., Dhar, A., & Buescher, K. (1996, March). Density-based clustering and radial basis function modeling to generate credit card fraud scores. In *IEEE/IAFE 1996 Conference on Computational Intelligence for Financial Engineering (CIFEr)* (pp. 247-251). IEEE.

Holland, J. H. (1975). Adaptation in natural and artificial systems: An introductory analysis with applications to biology, control, and artificial intelligence.

Huang, Z., Zheng, H., Li, C., & Che, C. (2024). Application of Machine Learning-Based K-Means Clustering for Financial Fraud Detection. *Academic Journal of Science and Technology, 10*(1), 33-39.

Mahmoudi, N., & Duman, E. (2015). Detecting credit card fraud by modified Fisher discriminant analysis. *Expert Systems with Applications, 42*(5), 2510-2516.

Nadim, A. H., Sayem, I. M., Mutsuddy, A., & Chowdhury, M. S. (2019, December). Analysis of machine learning techniques for credit card fraud detection. In *2019 International Conference on Machine Learning and Data Engineering (iCMLDE)* (pp. 42-47). IEEE.

Niu, X., Wang, L., & Yang, X. (2019). A comparison study of credit card fraud detection: Supervised versus unsupervised. *arXiv preprint arXiv:1904.10604*.

Parmar, J., Patel, A., & Savsani, M. (2020). Credit card fraud detection framework-a machine learning perspective. *International Journal of Scientific Research in Science and Technology, 7*(6), 431-435.

Pehlivanli, D., Eken, S., & AYAN, E. B. (2019). Detection of fraud risks in retailing sector using MLP and SVM techniques. *Turkish Journal of Electrical Engineering and Computer Sciences, 27*(5), 3633-3647.

Perols, J. (2011). Financial statement fraud detection: An analysis of statistical and machine learning algorithms. *Auditing: A Journal of Practice & Theory, 30*(2), 19-50.

Renjith, S. (2018). Detection of fraudulent sellers in online marketplaces using support vector machine approach. *arXiv preprint arXiv:1805.00464*.

Roseline, J. F., Naidu, G. B. S. R., Pandi, V. S., alias Rajasree, S. A., & Mageswari, N. (2022). Autonomous credit card fraud detection using machine learning approach. *Computers and Electrical Engineering, 102*, 108132.

Sahin, Y., & Duman, E. (2011, March). Detecting credit card fraud by decision trees and support vector machines. In *Proceedings of the International MultiConference of Engineers and Computer Scientists* (Vol. 1, pp. 1-6).

Seyedhossein, L., & Hashemi, M. R. (2010, December). Mining information from credit card time series for timelier fraud detection. In *2010 5th International Symposium on Telecommunications* (pp. 619-624). IEEE.

Shen, A., Tong, R., & Deng, Y. (2007, June). Application of classification models on credit card fraud detection. In *2007 International conference on service systems and service management* (pp. 1-4). IEEE.

Shukur, H. A., & Kurnaz, S. (2019). Credit card fraud detection using machine learning methodology. *International Journal of Computer Science and Mobile Computing, 8*(3), 257-260.

Van Vlasselaer, V., Bravo, C., Caelen, O., Eliassi-Rad, T., Akoglu, L., Snoeck, M., & Baesens, B. (2015). APATE: A novel approach for automated credit card transaction fraud detection using network-based extensions. *Decision support systems, 75*, 38-48.

Varmedja, D., Karanovic, M., Sladojevic, S., Arsenovic, M., & Anderla, A. (2019, March). Credit card fraud detection-machine learning methods. In *2019 18th International Symposium INFOTEH-JAHORINA (INFOTEH)* (pp. 1-5). IEEE.

Walke, A. (2019). Comparison of supervised and unsupervised fraud detection. In *Advances in Data Science, Cyber Security and IT Applications: First International Conference on Computing, ICC 2019, Riyadh, Saudi Arabia, December 10–12, 2019, Proceedings, Part I 1* (pp. 8-14). Springer International Publishing.

Yi, Z., Cao, X., Pu, X., Wu, Y., Chen, Z., Khan, A. T., ... & Li, S. (2023). Fraud detection in capital markets: A novel machine learning approach. *Expert Systems with Applications, 231*, 120760.

How cite this article

Demirdelen, A., Vardarlıer, P., Meral, Y., & Özcan, T. (2024). Diagnosis of internal frauds using extreme gradient boosting model optimized with genetic algorithm in retailing. *Acta Infologica, 8*(1), 60-70. https://doi.org/10.26650/acin.1475658

İSTANBUL UNIVERSITY PRESS

Acta Infologica
acin 2024, 8 (1), 71–87
DOI: 10.26650/acin.1389571

Review Article / Derleme Makalesi

An Overview of Paradigm Shift Dynamics in Transportation: Use of Artificial Intelligence in Intelligent Transportation Systems in Türkiye

Esma Dilek[1], Özgür Talih[2], Türksel Kaya Bensghir[3]

[1]Gazi University, Graduate School of Natural and Applied Sciences, Department of Information Security Engineering, Ankara, Türkiye
[2]Bandırma Onyedi Eylül University, Graduate School of Applied Sciences, Intelligent Transportation Systems and Technologies Master's Thesis Program, Balıkesir, Türkiye
[3]Hacı Bayram Veli University, Faculty of Economics and Administrative Sciences, Department of Business Administration, Ankara, Türkiye

Corresponding author : Esma Dilek
E-mail : esma.dilek@gazi.edu.tr

ABSTRACT

Currently, technology-based methods are widely used in the solutions developed for smart cities and sustainable transportation. Owing to the rapid advances in technology, the traditional structure of transportation networks is undergoing a paradigm shift. Artificial Intelligence (AI), that has started to have disruptive effects in many sectors in the future, is expected to be one of the most influential factors in the paradigm shift in transportation. In this paper, the dynamics of the paradigmatic shift in transportation to promote sustainable transportation in Türkiye are evaluated through conducting a review of the existing literature. Scenarios exhibiting the use of disruptive and innovative technologies in transport systems, specifically, AI applications in intelligent transportation systems (ITS), are examined. Additionally, the economic, environmental, and social impacts of AI applications are discussed by emphasizing the need to identify priority areas for the effective use of AI in the field of intelligent transport. Thus, this paper, by summarizing the use of AI-based technologies for intelligent transport in Türkiye, contributes to the literature by providing an overview of the existing knowledge.

Keywords: Paradigm shift in transportation, artificial intelligence, intelligent transportation systems, sustainable transportation, Türkiye

Submitted	**:** 11.11.2023
Revision Requested	**:** 19.02.2024
Last Revision Received	**:** 06.03.2024
Accepted	**:** 14.03.2024
Published Online	**:** 24.04.2024

1. INTRODUCTION

Transportation systems are characterized by an ever-increasing claim on public resources and many challenges need to be overcome for the proper functioning of these systems. Some of these challenges include the efficient management of the large number of components involved in the systems, involvement of multiple stakeholders seeking to achieve several goals that may conflict with each other, and their multidisciplinary structure.

Considering the increasing demand for transportation triggered by economic development and the growing desire of individuals to travel with more comfort, there is a need to strengthen traditional transportation systems with innovative approaches to overcome the difficulties presently experienced by passengers. In this regard, it becomes necessary to adopt a management strategy that focuses on the needs of stakeholders in the field of transportation and allows the integration of different transportation modes and stakeholders in an efficient and coordinated manner. While determining these strategies, a paradigm shift is observed that focuses on the digitalization of transportation processes and ensuring optimal interoperability between different actors and transportation modes. In such a paradigmatic shift, it is observed that principles such as sustainability, accessibility, connectivity, security, and safety constitute the fundamentals of transportation and mobility design. Intelligent transportation systems (ITS), that is one of the main approaches employed in this paradigmatic shift, and which is in line with these principles, is increasingly attracting researchers' attention due to its potential to revolutionize the way passengers and freight are transported. Disruptive technologies such as artificial intelligence (AI), Big Data, Internet of Things (IoT), and distributed ledgers are used in ITS to effectively address and prevent accidents, recognize unusual traffic conditions, optimize routes, and maintain roads (Oladimeji et al., 2023). In addition to AI and its technologies, several innovative technologies such as Connected, Cooperative and Automated Mobility (CCAM), cloud computing, open data, blockchain technologies, drones, air taxis, immersive interfaces, digital twins, Mobility as a Service (MaaS), smart roads, and hyperloop are used in this paradigmatic shift to address transportation problems. These technologies offer a diverse range of benefits and impacts for individuals, countries, and governments as compared to that by traditional approaches. This digital transportation infrastructure also brings with it a dynamic governance approach that involves continuous change and improvement.

This paradigm shift in the field of transportation, that is associated with dynamic governance and factors (hereinafter referred to as "paradigm shift dynamics"), is dominated by disruptive and innovative technologies that provide solutions to the existing problems, benefits from several AI applications, and reveals a data-based, human and environment-oriented perspective. It is open to constant development, and this vivid and developing nature of the technologies in the paradigm shift in transportation is effective in achieving sustainable, fair, environmentally friendly, and efficient transportation by focusing on data, which is considered a valuable mine in this century. Thus, more effective and intelligent decision-making is possible in such data-driven transportation management frameworks. Within the context of this shift, the variety and volume of mobility data collected is increasing every day due to the widespread use of different technologies such as surveillance cameras, LIDARs (Laser Imaging Detection and Ranging), and detectors for effective traffic management. The data collected from sensors are obtained from monitoring (i) physical elements such as roads, vehicles, and pedestrians, and (ii) digital components for ensuring the reliability and security of the communication network. AI systems employing these data are increasingly penetrating into our lives every day, triggering radical changes in transport systems. Machine Learning (ML) approaches, subfield of AI, are among the currently used innovative methods to analyze the collected transport data (Schneider, Kutila, & Hoess, 2021).

The main topics focused in this paper that aim to shed light on the factors currently driving the paradigm shift in the transportation sector, are as follows:

- Dynamics of the paradigmatic shift in transportation systems
- Disruptive and innovative technologies that play a significant role in the paradigm shift in transport
- AI-based ITS applications that are used to promote sustainable transportation in Türkiye
- The use of AI applications in ITS and their economic, environmental, and social impacts
- Academic studies in Türkiye that employed AI algorithms for various aspects of transportation management
- Evaluations and suggestions for the stakeholders in Türkiye to enable them to benefit from the paradigmatic shift in the field of transport, including legal infrastructure

The remainder of this study is organized as follows: in Section 2, the method followed by the authors of this paper for conducting a review of past literature is summarized. In Section 3, an overview of the dynamics of the paradigmatic shift in transportation systems is presented with a focus on disruptive and innovative technologies that play a role in the transformation of transport systems; additionally, AI applications in ITS along with AI's historical evolution and impacts are discussed. In Section 4, the dynamic paradigmatic shift in transportation and AI applications used in ITS regarding the promotion of sustainable transportation in Türkiye are examined. Finally, Section 5 presents the discussion and conclusions.

2. METHOD

This paper reviews recent studies, national strategy and policy documents, and evaluations to shed light on the factors currently shaping the paradigm shift in the transportation sector. It aims to highlight the dynamics of transformation affecting the future of sustainable transportation. Considering past studies in the literature, the paper focuses on the disruptive and innovative technologies used in transport and AI technologies that help achieve sustainability goals in the transportation sector. In this regard, these topics are evaluated from a global perspective, and impacts of AI, especially in the ITS sector, are discussed. A current situation analysis is conducted to reveal the dynamic paradigmatic shift and the view of sustainable transportation in Türkiye.

3. DYNAMICS OF THE PARADIGM SHIFT IN TRANSPORTATION

In addition to changes in the demand for transport services, the attributes of the emerging technologies also shape the transport sector. Especially in urban transport, various types of mobility systems are emerging; due to the radical change facilitated by these technologies, they are often termed “disruptive and innovative technologies.” The concept of intelligent transportation has gained prominence in smart cities due to the integration of elements of the fourth industrial revolution into transportation systems, including IoT, sensor technologies, autonomous vehicles (AVs), cloud computing, Big Data, AI, digital twins, and blockchain (Önder & Akdemir, 2020), thus taking its place among the dynamics leading the paradigm shift in transportation.

Urban mobility, through which human needs can be observed intensively, is one of the most dynamic areas in the paradigmatic shift experienced in transportation. Consequently, many researchers are currently working on the development of new mobility technologies in line with smart city and urban transport concepts that have the potential to change many aspects of urban transport, from the type of fuel to the style of driving. Some of these developments are as follows: the ability of vehicles to cooperate and communicate with each other while gradually becoming automated, the emergence of radical changes in travel patterns in the long term with the concept of MaaS, etc. (Medina-Tapia & Robusté, 2018). Table 1 presents the paradigm shift in transportation. This has been adapted from (Litman, 2013)’s work and updated in accordance with the current trends in transport sector.

Table 1. Paradigm Shift in Transportation

Criteria for Comparison	Old Paradigm	New Paradigm
Definition of transport	Mobility (physical traveling)	• Accessibility (the ability of people to reach services and activities)
Planning objectives	Reduced congestion, journey times, accidents, emissions, and costs	• Provision of efficient and fair transport services for all road users, including disadvantaged groups
Modes considered	Preference for individual car use	• Multimodal approach (walking, public transport, cycling, etc.)
Common effects	Travel speed and congestion duration, cost of vehicle operations, accident and emission rates	• Several economic, social, environmental, and mental impacts, including indirect impacts • Sustainable transportation
Performance indicators	Vehicle traffic speed, road level of service (LoS), distance-based accident and emission rates	• Quality of transport options • Multimodal LoS • Compatibility of land use with accessibility
Consideration of transport demand management (TDM)	Deeming it necessary to reduce individual vehicle use in general and acceptance of TDM as the last alternative	• Support for TDM when it is cost effective
Transport improvement strategies	Increased road capacity and parking areas	• Improvement of transport options • Use of TDM • More accessible urban space planning
Impact on sustainability	Reduced rates of traffic accidents and pollution emissions per kilometer	• Reduced accident and emission rates per capita • Improved physical activity and basic access conditions
Technology utilization level	Use of traditional transport modes and transport systems	• Intensive use of ITS that combines traditional and innovative technologies • Integrating CCAM with the existing technologies in use • Widespread use of electric vehicles and technologies • Increasing use of disruptive and innovative technologies

In the following sections, disruptive and innovative technologies that constitute the paradigmatic shift dynamics in transportation systems and AI applications used in ITS that lead this paradigmatic shift are examined.

3.1. Disruptive and Innovative Technologies in Transport

Currently, traditional methods used for solving critical transport and infrastructure problems are increasingly exposed to disruptive effects due to technological advances. This has led to providing people with new opportunities for development. The innovative modes of transportation facilitated by these opportunities are transforming the way we travel; they are also making transport solutions increasingly dependent on the use of digital technologies (European Bank, 2019). The disruptive and innovative technologies that are triggering a paradigmatic shift in transport systems are depicted in Fig. 1. They are categorized in terms of (i) Information and Communication Technologies (ICT), (ii) Transportation, and (iii) ICT and Transportation.

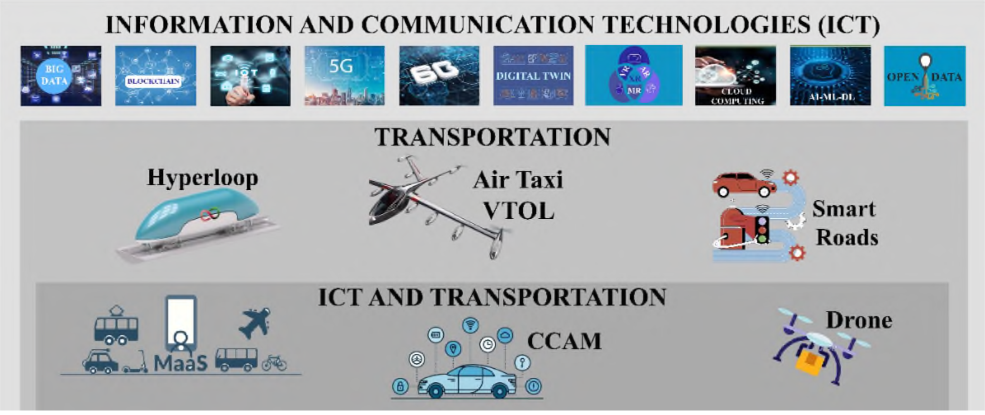


Figure 1. Disruptive and Innovative Technologies in Transport (Source: Authors)

3.1.1. Disruptive Technologies in ICT

IoT technology offers solutions that introduce radical changes in generally accepted approaches in several areas of transport such as traffic management, public transport, electric vehicles and charging infrastructures, railway systems, logistics and supply chain, fleet tracking, CCAM technologies, smart contracts, and last mile transport. IoT technology-based transport solutions aim to provide more flexible, efficient, safe, and secure mobility options. Traffic safety and security is one of the main issues of urban mobility and IoT solutions play an active role in detecting user errors, preventing traffic accidents, and facilitating effective traffic management (Derawi, Dalveren, & Cheikh, 2020). Further, Connected and Automated Vehicles (CAVs) can act in a more informed and coordinated manner owing to their ability to interact with everything using IoT devices (Nikitas, Michalakopoulou, Njoya, & Karampatzakis, 2020).

Big Data has been increasingly attracting the attention of researchers worldwide in terms of its application for the transportation sector. This technology provides data for transportation from several sources such as social media, data centers, sensors, vehicles, and open data platforms (Torre-Bastida et al., 2018). Regarding ITS, the application of Big Data can help reduce supply chain wastes, consolidate shipments, optimize logistics activities, contribute to improving the end-to-end user experience, and reduce negative environmental impacts of transportation (Hashem et al., 2016).

Within the framework of cloud computing technologies, researchers have proposed ITS-Cloud, that comprises traditional-static and temporary-dynamic cloud submodels for ITS to improve the efficiency of transport applications and services (Bitam & Mellouk, 2012).

Open transport data platforms provide information to the public and private sectors to create efficient transport solutions. They also provide opportunities for road users to efficiently manage their travels (T.C. Ulaştırma ve Altyapı Bakanlığı, 2022b).

Blockchain technology can be used in various areas related to transport and mobility such as in data sharing, payment systems, MaaS, supply chain and logistics services, and CCAM technologies (T.C. Ulaştırma ve Altyapı Bakanlığı, 2020).

The use of digital twins facilitates the accurate simulation of the real transport and road network. This technology can also be applied to predictive analytics-based approaches to make decisions regarding transport enhancements, including making decisions based on the analysis of traffic congestion (Rudskoy, Ilin, & Prokhorov, 2021).

Immersive technologies offer cost-effective solutions that address transportation requirements such as improving user experiences, testing areas for AVs, platforms for urban design, and training and education of pilots, drivers, and passengers (Li, Trappey, Lee, & Li, 2022). These technologies contribute to the transformation of the automotive and transport sector by virtual (VR), augmented (AR), and mixed (MR) reality solutions. These immersive systems and scenarios are often used to understand transport-related situations such as the operation of AVs, operation of railway systems, maintenance for motorways and vehicles, air traffic management, complexity of the real dynamic traffic environment, and cost uncertainties.

3.1.2. Disruptive Technologies in Transportation

Roads no longer constitute a physical entity or solid ground alone, but also include many supporting elements that are introduced by disruptive and innovative technologies. Currently, they are equipped with ICT, provide renewable energy, weigh the load of the moving vehicles, automatically charge electric vehicles, instantly detect traffic violations, communicate with road assets, have smart intersections and lights, and include fast emergency rescue features (Toh, Sanguesa, Cano, & Martinez, 2020). Smart roads not only use devices such as speed and noise detectors, CCTV (Closed Circuit Television) cameras, smart traffic lights and street lights, road and weather monitoring systems, digital signage, parking systems to improve transport safety, but also interact with CAVs, making the driving experience safer by enabling route tracking (Koptelov, 2022).

Hyperloop, described as the “fifth mode of transport,” is a transport and mobility alternative that is operated with energy obtained from renewable energy resources. It consumes less energy as compared that of an aircraft and can travel at a speed of approximately 1,200 km/h. It is resistant to earthquakes and weather conditions, and is not exposed to traffic problems (T.C. Cumhurbaşkanlığı Dijital Dönüşüm Ofisi, 2020).

The features and benefits of an air taxi that is one of the disruptive and innovative technologies in transportation, are summarized in (Pan & Alouini, 2021) as (i) supporting environmentally friendly transport, (ii) enabling transport without traffic congestion, (iii) providing flexible and fast door-to-door transport, (iv) requiring less ground support infrastructure, (v) allowing more space for road users, and (vi) lower maintenance and construction costs.

3.1.3. Disruptive Technologies in ICT and Transportation

In (ERTRAC, 2022), that discusses the effects of AI in the framework of CCAM, the applications of AI and its effects in the context of autonomous driving are comprehensively examined. AV is a broad and dynamic concept under development, that plays an important role in the paradigm shift of urban transport (Medina-Tapia & Robusté, 2018). AV technologies use a combination of hardware and software components to effectively employ driverless features. Owing to the sensitive communication AVs establish with each other, they can reduce traffic congestion and simultaneously perform intelligent fleet management. Additionally, current AVs are equipped with a 360° vision, that can significantly help to reduce accidents (Oladimeji et al., 2023). AVs, that use hardware such as radar, LIDAR, and cameras to perceive the environment, can benefit from ML models for planning and decision-making by interpreting data collected from sensors (Schwarting, Alonso-Mora, & Rus, 2018). CAVs can transform mobility needs, transportation networks, and road infrastructure by transferring vehicle control and driving responsibility from humans to machines with immense AI and wireless connectivity capabilities (Nikitas et al., 2020).

Drones, especially when used in logistics activities, have the potential to have a positive impact on the environment by significantly reducing the number of vehicles on the road. The use of drones is bringing about radical changes to asset management, infrastructure inspections, and field survey tasks. It also supports logistics activities within the framework of consumer services and product deliveries in several developed and developing global markets (European Bank, 2019). In future, a drone (i) can be an alternative vehicle for the delivery of emergency aid, time-sensitive products, medical supplies, etc. without facing the difficulties of waiting in traffic or location access, (ii) can be a more feasible solution for urban air mobility, (iii) can record high-resolution images and videos, (iv) can be employed in emergency response and disaster relief cases, (v) can be used in the monitoring and management of parking spaces in urban areas owing to its advanced sensors, real-time data transfer ability, and effective traffic management (Vega, 2023). Thus, Unmanned Aerial Vehicles (UAV), also known as drones, are among the paradigmatic shift tools that use AI and wireless technologies for air transportation (Nikitas et al., 2020).

MaaS is a system that offers digital packages of personalized multimodal mobility services through an intelligent online platform that can provide integrated journey planning, booking, smart ticketing, and real-time information services (Nikitas et al., 2020). Thus, MaaS, as one of the technologies facilitating the paradigmatic shift in transportation, has the potential of reducing individual car use, and helps passengers to choose the most efficient mobility service based on their specifications. In this approach, optional access to various types of transportation is possible in line with user demands, individual travel options are offered for each user, and users can choose the best offers from a digital platform (Talih & Tektaş, 2023).

3.2. AI in ITS and its Impacts

This section discusses the following: the historical development of AI, and the use of AI in ITS and mobility solutions, and AI's economic, social, and environmental impacts.

3.2.1. Historical Evolution of AI

Although AI technology is currently widely used in numerous sectors, including transportation, its emergence dates back to the 1950s. Alan Turing's article titled "Computing Machinery and Intelligence" laid out the framework on how to build and test AI in 1950. In 1949, Norbert Wiener, who is considered the modern-day father of the field of cybernetics, stated that machines could accomplish everything that was done in an understandable and straightforward manner. Wiener revealed the possible disruptive effects of AI on the economy by emphasizing that the value of regular factory workers will gradually decrease with the development of the skills of machines (Markoff, 2013). Between 1958 and 1959, Ord. Dr. Cahit Arf gave talks on the thinking potential of machines as part of the Atatürk University Public Conferences (Sarı, 2021).

However, subsequently, the development of research in the field of AI stagnated due to financial difficulties and the inadequacy of the equipment to conduct practical experiments based on the theory put forward in the 1950s. In 1966, a report published by the Automatic Language Processing Advisory Committee argued that no improvement was achieved by AI-based systems after evaluating these systems in terms of quality, speed, and cost, without considering their scientific value. Later, in the Lighthill report published by the British government, it was claimed that both the public and scientific community found AI to be insufficient in terms of industrial applications. The report anticipated that traditional engineering approaches would produce better results. In this environment, the funds allocated to AI research decreased and developments were disrupted. After this period, AI evolved into a new paradigm, Expert Systems, that focused on transferring the experience of humans in their field of expertise to machines and running it on personal computers. Digital Equipment Corporation announced that the expert system called XCON (eXpert CONfigurer) saved them 40 million dollars a year between 1980–1986. The US invested approximately over 1 billion dollars in Expert Systems in 1985. While the UK launched the Alvey program worth 350 million pounds, Japan announced an investment of 850 million dollars within the scope of the 5th generation computer project (Delipetrev, Tsinaraki, & Kostic, 2020).

In the years between 1990 and 2010, the concept of ML, a subbranch of AI, dominated the research landscape (Balasubramanian, Libarikian, & McElhaney, 2021). With the developing hardware and increasing amount of data, the problem-solving capabilities of algorithms began to be evident in many areas. However, it was noted that when there is an abundance of data, standard ML algorithms encounter challenges, whereas Artificial Neural Network (ANN) based models provide better results. In 2009, in the ImageNet (Hou et al., 2022) competition, an event in which advanced image classification models compete, 3.2 million labeled images were made available for 5,247 classes. In the performance evaluations during this competition, the deep learning model known as AlexNet achieved significant success as compared to other competitors; this attracted the attention of researchers toward this direction (Krizhevsky, Sutskever & Hinton, 2012). Thus, deep learning (DL) models, that are a subset of ML methods, started to emerge. While there are separate models in ML algorithms that extract summary information from images and perform classification and recognition using this information, both tasks can be performed within a single ANN in DL models (Dilek & Dener, 2023). This pioneered a paradigmatic shift in AI technologies. Currently, with the introduction of Big Data and powerful hardware, AI models and algorithms have become a formidable actor that forces radical changes in traditional methods used in ITS. Specifically, deep ANN-based models developed after 2010 have begun to be used in many sectors, including transportation. As research in the field of AI accelerates, models with new capabilities will emerge. The historical evolution of AI technologies is depicted in Fig. 2.

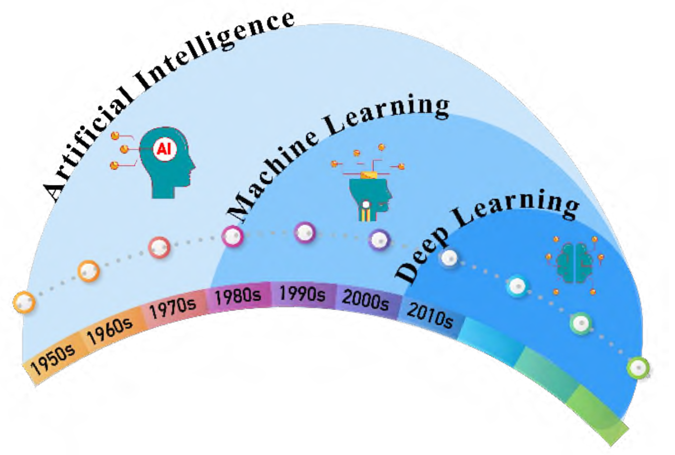


Figure 2. Historical Evolution of AI Technologies (Source: Authors)

3.2.2. AI in ITS

ITS, that can be defined as the transition of traditional transportation systems to a data-driven mechanism, is an ecosystem that includes many subtopics—ranging from innovative transportation systems to logistics, and from communication to security of its infrastructure. AI, which is one of the pioneers of the paradigmatic shift in transportation, has a wide range of application areas within the framework of ITS, as shown in Fig. 3.

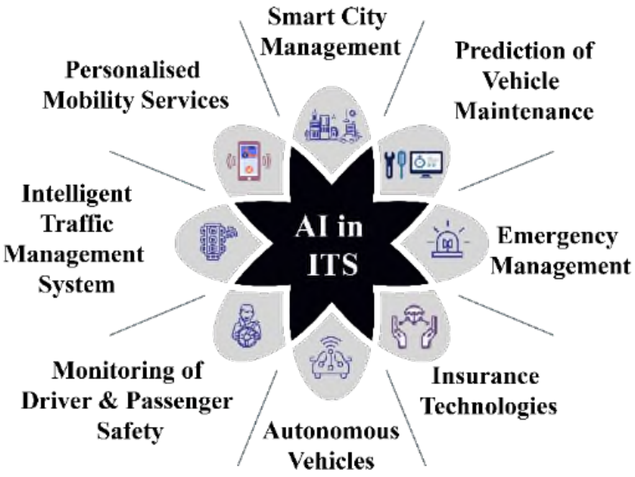

Smart City Management
Prediction of Vehicle Maintenance
Personalised Mobility Services
Intelligent Traffic Management System
AI in ITS
Emergency Management
Monitoring of Driver & Passenger Safety
Insurance Technologies
Autonomous Vehicles

Figure 3. AI Application Areas Within the Context of ITS (Source: Authors)

AI is a powerful tool that, when used consciously, has the potential to foster a shift to a paradigm that uses resources more efficiently, facilitating sustainable living. AI technologies are used in various areas of ITS. Some of these areas include AVs, traffic monitoring, traffic sign recognition, pedestrian detection, traffic flow analysis, computer vision (CV) assisted parking management, road condition monitoring, automatic detection of traffic incidents, automatic number plate recognition, driver tracking systems, and infrastructure health inspection. AI models are also widely used in the prediction and detection of traffic situations such as road conditions, and traffic volume, density, and incidents. Owing to AI applications, it is possible to improve decision-making mechanisms, increase the efficiency of public transportation systems to promote sustainability, and reduce accidents in the road network (Abduljabbar, Dia, Liyanage, & Bagloee, 2019).

A detailed analysis of the use of ML algorithms used in ITS was presented by (Yuan et al., 2022). In their study, the use of ML algorithms in the field of ITS has been systematically addressed under three categories, mainly (i) detection/recognition, (ii) prediction, and (iii) management. The AI tasks and application scenarios utilized in the field of ITS are summarized in Table 2.

Table 2. AI Tasks and Application Scenarios in the Field of ITS (Adapted by the authors)

AI Tasks	Application Scenarios
Detection/Recognition	· Detection and classification of road assets (vehicles, pedestrians, traffic lights, signs, road lines, etc.); detection of obstacles, transportation infrastructure damages, road status, etc. · Detection and classification of vehicles according to their types · Detection of pedestrians and behavior analysis of drivers · Grouping data packets in order of importance in traffic communication networks, detection of network intrusions, and security vulnerabilities
Prediction	· Traffic density prediction · Travel time estimation · Prediction of vehicle and pedestrian behaviors · Route planning · Accident prediction and warning systems
Management	· Traffic signal management: Organizing traffic lights and active speed limits according to road conditions · Resource management: Optimizing communication systems and balancing storage/computing loads · Demand management: Providing route optimization for service providers such as those providing rental car, taxi, bicycle services, etc.

ANNs, Genetic Algorithms, Simulated Annealing, Artificial Immune System, Bee Colony Optimization, Ant Colony Optimiser, and Fuzzy Logic Model are among the AI methods that are used to increase the intelligence level of ITS (Abduljabbar et al., 2019).

3.2.3. Impacts of AI

It is anticipated that the use of AI technologies in ITS can provide time and financial savings, reveal potential assets that generally remain hidden in business development processes, and provide substantial gains in data governance in many areas, ranging from individual mobility of people to management of traffic in the logistics sector. Some economic, environmental, and social impacts of AI technologies are examined in the following sections.

- Economic Impacts

As expressed by (Howarth, 2023), 83% of companies state that using AI in their business strategies is their top priority. In 2035, the base gross revenue in the transport, logistics, and warehousing sector is estimated to be USD 2,131 billion, where the contribution of AI is estimated to be approximately USD 744 billion (nearly 35%). AI is also projected to cause severe fluctuations in employment, as automating physical and cognitive tasks is likely to lead to massive job losses for low-skilled workers in the transportation industry (Manyika James and Sneader Kevin, 2018). This negative gross impact is estimated to be approximately 10% by 2030 (Bughin Jacques, Seong Jeongmin, Manyika James, Chui Michael, & Joshi Raoul, 2018). While low-skilled workers are at risk of being replaced by technology and machines, demand for highly skilled workers in areas such as data analytics, engineering, cybersecurity, and vehicle monitoring, that enable the development of AI-powered mobility solutions, is expected to increase.

It is projected that AI models will gradually add approximately 13 trillion euros to the global economy by 2030 that is equivalent to approximately 14% (Rao, Verweij, & Cameron, 2020). Moreover, AI is estimated to contribute to an average annual productivity growth of approximately 1%–2% for the same period (Bughin Jacques et al., 2018).

Daily losses in automated parking systems and city traffic management can be prevented by employing AI-based techniques. Improving traffic congestion saves time and money by minimizing unnecessary fuel consumption. In the European Union (EU), congestion often occurs in and around urban areas. This congestion costs approximately 100 billion euros annually, which is roughly 1% of the EU’s Gross Domestic Product (GDP). AI has the potential to reduce travel times for the benefit of both individuals and the industry. AI applications can be used to adaptively design road infrastructure and signaling systems to better evenly distribute traffic by forecasting future demands. Efficient AI-based traffic management is expected to reduce waiting time at traffic lights by up to 47% (Batura et al., 2021).

Another contribution of AI is reducing road maintenance costs. It is reported that expenses will be minimized with fully automated transportation systems and will create a contribution of 38 billion euros across Europe (Batura et al., 2021). With the spread of AVs, a significant decrease in accident rates is expected, leading to reducing damage to public property and health costs due to injuries.

The interaction between ITS and AI also has an economic impact on energy savings. Some studies show that for various deployment scenarios, significant energy savings will be achieved from the use of smart cars (both electric and hybrid) (Chase, Maples, & Schipper, 2018). Specifically, the management of public transport and taxi mobility and shared passenger strategies can avoid a significant waste of public resources.

Fleet tracking systems are an additional energy saving mechanism. Truck platooning can reduce logistics costs and it is indicated that a fuel saving of 4% can be achieved in such a driving protocol as wind resistance will be reduced. This system's ability to keep the convoy model continually mobile is another economic benefit (Pham, 2018). By changing the driver of the tired leader vehicle, the convoy can be actively maintained.

Another dimension related to the economic impacts of AI in the field of ITS is related to insurance companies. Individual policies can be created based on the data collected from personalized vehicles; thus, more advantageous insurance service packages can be created. However, it is also predicted that the insurance industry's revenues will tend to decrease in the long term, especially with the introduction of fully automated vehicles. Additionally, the commercial taxi sector will also be affected as AI-assisted innovative transport trends will replace traditional approaches. In addition to these expectations, it is estimated that AI will trigger new sectors and lead to the creation of new business areas. Business areas pioneered by CAV technologies are expected to create 25.000 new jobs by 2035 (Batura et al., 2021).

- *Environmental Impacts*

The use of AI within the framework of ITS will also have positive effects on the environmental conditions. AI will have such positive effects through ensuring energy efficiency of vehicles, AV applications that can make decisions according to traffic conditions, and improving road conditions. With the support of ITS powered with AI, it is predicted that traffic management systems will be enhanced and the traffic congestion problem will be mitigated (Batura et al., 2021). AI-assisted management of traffic flow will enable the reduction of congestion and engine idling by ensuring steady traffic flow at optimum speeds. AI will also make it feasible to select the fastest and most energy-efficient route using traffic prediction. Thus, greenhouse gas emissions, fuel consumption, air pollution, and noise pollution can be prevented. As there will be less braking in a steady traffic flow, nonexhaust emissions will also be reduced. Thus, owing to the increased operational efficiency using AI-assisted systems, especially in the logistics sector, savings of approximately 500 billion dollars will be achieved, while the creation of 280 megatons of waste will be prevented worldwide in terms of reduction in carbon dioxide (CO_2) emissions (Transforming Transport, 2017).

The use of AI-supported applications in the field of ITS may have negative impacts as well. For example, since the training of advanced AI models has long computation times, model training processes have a significant carbon footprint. Hence, this needs to be considered in terms of environmental impacts of AI.

- *Social Impacts*

The European Environment Agency states that transport has a significant impact on the quality of life, especially in cities, and that reducing traffic-related pollution (air, noise, greenhouse gas emissions) significantly improves the quality of life (European Environment Agency, 2016). With the autonomy facilitated by AI applications, it is expected that there will be significant changes in people's lives from social aspects. Psychological problems caused by traffic congestion will decrease and the quality of life will increase. Since AVs are expected to obey traffic rules and avoid over speeding, they will help mitigate the emission of harmful gases and reduce the number of accidents caused by drivers due to making mistakes or intake of alcohol. It is also estimated that, due to increased road safety, fatal accident rates will decrease.

AI applications have the potential to enhance transportation services in both urban and rural regions, resulting in an improved travel experience (Batura et al., 2021). Owing to the increased accessibility to various destinations, significant improvements can be achieved with AI technologies for less mobile population groups such as the elderly and disabled. Thus, the disadvantaged groups will have increased opportunities of socialization.

However, with the penetration of AI and autonomous systems into our lives, security-related issues will emerge. Detection of incidents in public transportation or in public areas will be possible with CV technologies using cameras. Identification of vehicles or people sought by law enforcement and the routes followed by criminals can be easily accomplished by processing millions of data by AI. However, the use of AI brings the issue of personal rights to the fore. With the development of facial recognition systems, concerns arise that governments may implement more pressured policies on some individuals (Dilek & Dener, 2023).

4. PARADIGMATIC SHIFT DYNAMICS AND AI APPLICATIONS IN TRANSPORTATION SYSTEMS OF TÜRKİYE

Paradigmatic shift dynamics in transportation and AI applications used in ITS to promote sustainable transportation in Türkiye are examined in the following sections.

4.1. Paradigmatic Shift Dynamics in Transportation

It is observed that the disruptive and innovative technologies that focus on intelligent transportation and data are the triggers of the new transportation paradigm. AI strategies for Türkiye are discussed (T.C. Cumhurbaşkanlığı Dijital Dönüşüm Ofisi, 2021) in "2023 Industry and Technology Strategy", which states that disruptive technologies such as AI transforms the national policies of countries where there is a significant investment made in automated mobility and UAVs (T.C. Sanayi ve Teknoloji Bakanlığı, 2019). Additionally, "Mobility Vehicles and Technologies Roadmap" reveals a governance model that includes innovative solutions and technologies in line with the global context and trends, strategic goals, action plans, and critical projects within the framework of mobility (T.C. Sanayi ve Teknoloji Bakanlığı, 2022). Moreover, there are short-, medium-, and long-term goals and actions that support innovation in transportation in the strategy document and action plan for the advancement of ITS in Türkiye (T.C. Ulaştırma ve Altyapı Bakanlığı, 2020).

In the "11th Transportation and Communication Council" held in 2013, where developments in transportation were discussed, intelligent transportation targets were determined as follows: (i) Deployment of ITS in cities and ensuring the integration of national ITS applications, (ii) improving traffic safety and travel comfort, (iii) ensuring uninterrupted traffic flow conditions, (iv) developing intelligent transportation infrastructures and systems to ensure the integration of the railway network with other transportation systems, (v) establishment of smart motorways and state roads by 2035, and (vi) effective use of ITS in all cities with clean-fueled and highly energy-efficient vehicle technologies. While the focus was on the mission, vision, and targets set in the "National ITS Strategy Document and the 2020–2023 Action Plan" (T.C. Ulaştırma ve Altyapı Bakanlığı, 2020) along with sustainable transportation and increasing the deployment of ITS in the "12th Transport and Communication Council" report published in 2021, it is observed that the targets set in the "Green Deal Action Plan" (T.C. Ticaret Bakanlığı, 2021) are in line with the targets of the "European Green Deal" that states that Europe aims to be the first climate-neutral continent in the world in 2050. In the last two council meetings, it is noticed that the issues of digitalization, decarbonization, and autonomous and universal access to transportation have been brought to the agenda of transportation (T.C. Ulaştırma ve Altyapı Bakanlığı, 2023d).

In the "12th Development Plan", AI is regarded as one of the pioneering technologies for the green and digital transformation of sectors and it includes several measures to strengthen the cooperation of the public, academia, and private sectors in the field of AI (T.C. Cumhurbaşkanlığı Strateji ve Bütçe Başkanlığı, 2023).

In addition to the issues covered in the abovementioned policy and strategy documents, some initiatives among the paradigmatic shift dynamics of transportation aimed at ensuring sustainable transportation in Türkiye are summarized as follows:

- The concept of IoT is considered as an important paradigmatic shift dynamic for developing innovative solutions in transport, industry, research, and service projects for many sectors in Türkiye. In several Turkish cities, such as İstanbul, Kayseri, Konya, Ankara, and İzmir, intelligent transportation solutions and applications designed for urban traffic management using data collected from IoT assets such as traffic measurement sensors, surveillance cameras, and Global Positioning System (GPS) sensors are employed (T.C. Ulaştırma ve Altyapı Bakanlığı, 2020; T.C. Çevre Şehircilik ve İklim Değişikliği Bakanlığı, 2022).
- Safe, secure, and smart roads that can communicate with their environment constitutes one of the main elements of the paradigm shift in transportation. Some motorways in Türkiye that include basic smart road components are the Ankara-Niğde Motorway[1] and Northern Marmara Motorway[2].
- Long-term targets were included in the "National ITS Strategy Document and the 2020–2023 Action Plan" in line with the need for legislative regulations for vehicles such as air taxis, drones, etc. (T.C. Ulaştırma ve Altyapı Bakanlığı, 2020).
- The "National ITS Strategy Document and the 2020–2023 Action Plan" is a supportive and guiding roadmap to promote CCAM technologies in Türkiye, and it is observed that the use of disruptive and innovative technologies such as AI in transportation is a part of the strategy (T.C. Ulaştırma ve Altyapı Bakanlığı, 2020).

[1] https://www.ankaranigdeotoyolu.com

[2] https://www.kuzeymarmaraotoyolu.com

- "Driving Architecture for Autonomous Vehicles and Determination of Connected Vehicle Traffic Test Scenarios Project" and "Determination of Technical Characteristics of In-Vehicle Information and Communication System Project" (HGM, 2022b) have the potential to be effective in the development, deployment, and raising awareness regarding CAV technologies in Türkiye. Additionally, with the implementation of "Deployment of C-ITS (Cooperative ITS) Test and Application Corridor" action, a test zone will be built in Türkiye where several C-ITS use cases will be tested and new solutions can be developed (T.C. Ulaştırma ve Altyapı Bakanlığı, 2020).
- The "Mobility as a Service Project" (HGM, 2022a), that is under development in line with the long-term targets set in the "National ITS Strategy Document and the 2020–2023 Action Plan" (T.C. Ulaştırma ve Altyapı Bakanlığı, 2020), is among the personalized sustainable mobility solutions being developed in Türkiye.
- The "Digital City Twins–Creating a 3D Model" study, initiated by the Ministry of Environment, Urbanization and Climate Change in 2016–2017, continues to cover all Turkish provinces and districts and coastal regions (T.C. Çevre Şehircilik ve İklim Değişikliği Bakanlığı, 2022).
- Hyperloop is a promising and innovative transportation solution of the future for long-distance freight and passenger mobility; however, as it a future solution, there is no associated regulation in place in Türkiye (T.C. Cumhurbaşkanlığı Dijital Dönüşüm Ofisi, 2020).
- Several datasets are shared in open data portals regarding transportation and mobility categories. In order to promote the sharing of collected data, studies are generally conducted by countries within the framework of the existing legal infrastructure. In both the "11th Development Plan" and "12th Development Plan", a measure to establish a "National Open Data Portal", where public data will be shared, including transportation and mobility data in Türkiye, has been outlined. These plans also outline the process to determine data anonymization standards (T.C. Cumhurbaşkanlığı Strateji ve Bütçe Başkanlığı, 2019), (T.C. Cumhurbaşkanlığı Strateji ve Bütçe Başkanlığı, 2023). Some of the municipalities in Türkiye that have open data portals are presented in Table 3 (AVTED, 2023).

Table 3. Some of the Municipalities That Have Open Data Portals (Listed by the authors)

Category	Residential Location	Open Data Portal Address
CITY	Ankara	https://seffaf.ankara.bel.tr
	Antalya	https://acikveri.antalya.bel.tr
	Balıkesir	https://acikveri.balikesir.bel.tr
	Bursa	https://acikyesil.bursa.bel.tr
	Gaziantep	https://acikveri.gaziantep.bel.tr
	İstanbul	https://data.ibb.gov.tr
	İzmir	https://acikveri.bizizmir.com
	Kayseri	https://acikveri.kayseri.bel.tr
	Kocaeli	https://veri.kocaeli.bel.tr
	Konya	https://acikveri.konya.bel.tr
	Ordu	https://acikveri.ordu.bel.tr
	Sakarya	https://acikveri.sakarya.bel.tr
DISTRICT	Beyoğlu	https://acikveri.beyoglu.bel.tr
	Eyüpsultan	https://acikveri.eyupsultan.bel.tr
	Kadıköy	https://acikveri.kadikoy.bel.tr
	Küçükçekmece	https://acikveri.kucukcekmece.bel.tr
	Tuzla	https://veri.tuzla.bel.tr

4.2. AI-Supported ITS Applications

The new generation vehicles in Türkiye include systems such as pedestrian detection, automatic emergency braking, lane keeping assistant, smart speed control, fatigue detection, parking assistant, and adaptive cruise control and navigation. These AI-supported systems generally use components such as sensors and camera systems to help ensure safety and secure driving. Furthermore, AI-supported applications can quickly identify circumstances such as failure to wear seat belts and cell phone usage while driving, as well as driver fatigue symptoms such as exhaustion, distraction, and sleeplessness (Ulaşım Yönetim Merkezi, 2022).

For the first time in Türkiye, an AI-supported incident detection system was designed with domestic resources and deployed for the effective traffic management of the Ankara-Niğde Motorway. With the integration of the domestic AI-supported incident detection system into the cameras on the motorway, operators and drivers are warned if anomalous situations occur on the motorway (AUSPOSTASI, 2020).

Within the scope of the intelligent motorway project signed with the Scientific and Technological Research Council of Türkiye (TÜBİTAK) and the Northern Marmara Motorway Administration, detection of anomalies on roads will be made possible by equipping the roads with fiber optic infrastructure by utilizing advanced signal processing and AI techniques; further, imaging will be collected by sending autonomous drones to the relevant regions to collect further information (AUS Türkiye, 2022).

Fog, icing, and accident risks are automatically detected, and necessary precautions can be taken at the control center on a 24/7 basis due to the integration of AI systems into the cameras located on the 1915 Çanakkale Bridge (T.C. Ulaştırma ve Altyapı Bakanlığı, 2022a).

Trains are monitored in real time using AI technologies in the Train Monitoring and Coordination Center established by the General Directorate of Turkish State Railways Transportation Inc. In trains, CV technology detects sleepiness, absent-mindedness, and fatigue of the driver, and abnormal situations such as the train slipping off the track. Thus, drivers are warned when required so that possible accidents caused by human errors can be prevented (T.C. Ulaştırma ve Altyapı Bakanlığı, 2023b).

The disruptive impact of AI is predicted to be in the logistics sector. Within the framework of the agreements signed with AVL in Türkiye, a significant investment has been made in the truck platooning-automated convoy systems. It is emphasized that especially in 2025, heavy commercial vehicle classes will switch to smart truck convoys, and that owing to this new convoy technology, companies expect to obtain financial gains between 8–15% in convoys with three trucks (Ford Otosan, 2019).

An AI-based Video Analysis System (Eyeminer) developed by HAVELSAN; autonomous buses that have Level 4 autonomy developed by several technology and automotive manufacturers such as ADASTEC, Karsan, Anadolu Isuzu and Otokar; and, trucks that have Level 4 autonomy manufactured by Ford Otosan are among the other sustainable transportation applications in Türkiye that benefit from AI technologies. Efforts are also being made to increase the safety and security of transportation systems with AI technologies; some of these efforts include the Horizon 2020 InSecTT Project and Horizon Europe BRIGHTER Project, of which Marmara University is a participant. The InSecTT Project utilizes AI for two core tasks: (i) AI-supported embedded processing for industrial tasks and (ii) AI enhanced wireless transmission. Within the scope of the BRIGHTER Project, images obtained from the cameras located at smart intersections are processed by AI on edge computing devices and subsequently transmitted to the vehicles via Road Side Unit (RSU) devices. Thus, it aims to provide useful services for traffic efficiency and safety (T.C. Ulaştırma ve Altyapı Bakanlığı, 2023a, 2023c).

AI systems can identify structural flaws in the transportation infrastructure, including potholes, flooding, and ice on the road. While several studies regarding this topic have been conducted worldwide, Pendik Municipality in İstanbul has also taken steps in this direction (Pendik Belediyesi, 2015).

Some AI-supported solutions within the framework of ITS in Türkiye can be summarized as follows:

- AI-based incident detection and traffic prediction (INTETRA, 2023).
- Dynamic intersection control, vehicle counting, automatic tunnel incident detection, and the electronic enforcement system used within the scope of traffic control and management (ISSD, 2022).
- AI-supported cameras for vehicle detection, classification, and counting (Asya Trafik, 2022).
- Attention and fatigue measurement of drivers with image analysis, object detection, meta learning, and video segmentation; development of autonomous systems that provide environmental perception and decision support; adaptation to radar systems to improve remote sensing targets; and, increasing the capabilities of existing systems (ASELSAN, 2022).
- Autonomous vehicle technologies (Leo Drive, 2023).

- Intelligent intersection management and signaling system, vehicle recognition, speed detection, traffic anomaly detection, and traffic management and control (MIA Technology, 2024).
- Traffic data collection, real-time video processing, real-time vehicle counting, and video processing (Neovision, 2021).
- Smart intersection management with AI-supported cameras (MOSAŞ, 2024).
- License plate recognition; recognition of the vehicle's model, color and brand; detection of the vehicle speed and vehicle tracking, and transfer of black and white list number plates to security units via the Police Information System (POLNET) (Divit, 2023).

The academia in Türkiye has also focused on utilizing AI in solving problems related to transport and traffic. Academic studies including scientific projects, theses, articles, and conference papers are published by several departments of universities. For conducting this review, we searched the Dergipark Academic Information System, National Thesis Centre, and Google Scholar for obtaining academic studies using Turkish keywords equivalent to "*artificial intelligence, transport, transportation, traffic, intersection, adaptive, vehicle, driver, signal, road, strategy, autonomous*". Some academic studies on the use of AI in intelligent transport solutions in Türkiye, obtained as a result of the search, are provided in Table 4.

Table 4. Some of Academic Studies Related to AI-Supported ITS

Study	Title
(Tektaş et al., 2002)	A Review on The Use of Artificial Intelligence Techniques in Traffic (Yapay Zekâ Tekniklerinin Trafik Kontrolünde Kullanılması Üzerine Bir İnceleme)
(Doğan, 2007)	Regression Analysis and Artificial Intelligence Approach with Traffic Accident Prediction Models for Türkiye and Some Chosen Big Cities
(Çevik, 2010)	Vehicle License Plate Recognition System with Artificial Intelligence Methods
(Erdal, 2018)	Use of Artificial Intelligence Techniques and Expert Systems in the Control of Terrestrial Intelligent Transportation Systems (Yapay Zekâ Teknikleri ve Uzman Sistemlerin Karasal Akıllı Ulaşım Sistemlerinin Denetiminde Kullanımı)
(Gülsün & Gonca, 2019)	Adaptive Traffic Management Systems
(Pazar et al., 2020)	Development of Artificial Intelligence-Based Vehicle Detection System
(Başkaya et al., 2020)	A Model Proposal on The Preparation of Effective Transportation Plans Using Artificial Intelligence Techniques Within the Context of Smart Cities
(Kadiroğulları et al., 2020)	Determination of Vehicle Number and Vehicle Transit Time for a Sample Intersection in Isparta Province Using ARIMA Artificial Intelligence
(Palandız et al., 2021)	Classification of Traffic Signs with Artificial Intelligence: A Sample Application for Denizli City Center
(Özmen et al., 2022)	Detection of Foreign Material Under Vehicle by Artificial Intelligence Methods and Automatic Passing System
(Demir, 2023)	Model Proposal for the Creation of Transportation Stories in The Context of Smart Transportation Systems and Artificial Intelligence
(Toğaç, 2023)	Gaziantep Province Artificial Intelligence-based Intelligent Transportation Systems, Adaptive Signalization Control and Simulation
(Ulu, 2023)	An Artificial Intelligence-Based Optimization Model and Application in Traffic Incident Management
(Narbay & Kirazlı, 2023)	Artificial Intelligence in Autonomous Vehicles, Processing of Personal Data and its Results
(Şafak, 2024)	Application of Artificial Intelligence Methods for Driver Assistance in Vehicles

AI-supported ITS studies in Türkiye focus on traffic management techniques, incident management and prediction methods, traffic control methods and measures, optimization models, development of vehicle number plate recognition and detection systems, classification of traffic signs, driver support systems, adaptive signaling control, and model proposals for the preparation of effective transportation plans with AVs, as shown in Table 4.

Based on our review, we observed that academia and public and private sectors in Türkiye employ AI-based methods in ITS such as AVs, pedestrian detection, management of traffic signalization systems and detection of traffic signs, travel time estimation, monitoring of road conditions, traffic incident detection, automatic number plate recognition, monitoring of driver behaviors, and vehicle detection systems. In this context, the main areas where AI and intelligent transport issues are addressed together in Türkiye are summarized below. It should be noted that the issues are not limited to these alone.

- Traffic analysis, control, and optimization methods, including traffic monitoring and management
- Decision support for ITS
- Detection of driver behaviors
- Travel time prediction
- Preparation of effective transport plans including the use of infrastructure
- AVs

5. DISCUSSION AND CONCLUSION

Issues such as increase in population and vehicle ownership rates and rapid urbanization in developing countries as well as in developed countries worldwide have made it necessary to investigate innovative approaches for efficient transport management. Dynamic factors such as the increase in daily journeys, transformation in mobility trends, digitalization, and autonomy in vehicles make it necessary to plan, develop, and deliver transport services beyond the traditional boundaries. To ensure improvement and sustainability in the field of transport, a paradigm shift involving disruptive and innovative technologies is observed that has become a necessity in the age of globalization.

It is noticed that transportation and mobility solutions can be significantly improved with the introduction of ITS, CCAM technologies, and innovative technologies that leverage AI, Big Data, and IoT. AI technologies stand out among the elements responsible for the paradigm shift dynamics in transportation since they enable efficient use of resources, that accelerates the development of countries with strategic steps and lower budgets. Consequently, it is essential for Türkiye to focus on AI technologies in the field of transportation by referring to a variety of use cases in the world, investing in this field using domestic resources, and adopting it in national policies. By following the worldwide trends in the public sector, academia, and private sector, determining prioritized scenarios in the field of ITS, conducting studies in line with the visionaries of the academia together with the workforce of the private sector will contribute to the development of sustainable transportation systems in Türkiye.

In this study, the dynamics of a paradigmatic shift in the field of transportation are examined and the role and impacts of AI-supported ITS applications for the establishment of sustainable transportation systems are analyzed. Moreover, AI applications used in the ITS sector and academic studies to promote sustainable transportation in Türkiye are presented. Considering the current situation in Türkiye, in order for the paradigmatic shift dynamics in transportation to successfully transform the transport sector, it is necessary to determine the legal framework, facilitate interoperability and data sharing, and create a supportive framework for CCAM technologies. Additionally, while deploying AI-supported systems in transport services, ensuring data security, identifying potential risks, addressing liability concerns, determining the strategy to increase the dissemination and acceptance rate of innovative applications, and building capacity through training and awareness activities is of critical importance.

Peer Review: Externally peer-reviewed.

Author Contributions: Conception/Design of Study- E.D., Ö.T., T.K.B.; Drafting Manuscript- E.D., Ö.T., T.K.B.; Critical Revision of Manuscript- E.D., Ö.T., T.K.B.; Final Approval and Accountability- E.D., Ö.T., T.K.B.

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

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

ORCID IDs of the authors

Esma Dilek	0000-0002-7994-0294
Özgür Talih	0000-0002-5899-2511
Türksel Kaya Bensghir	0000-0002-2313-5325

REFERENCES

Abduljabbar, R., Dia, H., Liyanage, S., & Bagloee, S. A. (2019). Applications of Artificial Intelligence in Transport: An Overview. *Sustainability, 11*(1). https://doi.org/10.3390/su11010189

ASELSAN. (2022). *Yapay Zekâ Teknolojileri*. https://www.aselsan.com/tr/arge#temelTeknolojiler

Asya Trafik. (2022). *Focus Loop Trafik Kamerası*. https://www.asyatrafik.com/focus-loop-trafik-kamerasi/

AUS TÜRKİYE. (2022). Kuzey Marmara Otoyolu (KMO) ve TÜBİTAK'tan Fiber Optik Tabanlı Akıllı Ulaşım Sistemi. *Bülten*, Vol. 7. Retrieved from https://austurkiye.org.tr/uploads/blog/file_8-07-numarali-bulten-474.pdf

AUSPOSTASI. (2020). *Sektörden haberler* (Vol. 4). Vol. 4. AUS TÜRKİYE. Retrieved from https://www.austurkiye.org.tr/uploads/blog/file_5-04-numarali-bulten-952.pdf

AVTED. (2023). *Yerel Yönetimler Açık Veri Endeksi*. Açık Veri ve Teknoloji Derneği (AVTED). Retrieved from www.acikveriendeksi.org

Balasubramanian, Ramrath., Libarikian, Ari., & McElhaney, Doug. (2021, March 12). Insurance 2030—The impact of AI on the future of insurance. Retrieved from https://www.mckinsey.com/industries/financial-services/our-insights/insurance-2030-the-impact-of-ai-on-the-future-of-insurance

Başkaya, O., Ağaçsapan, B., & Çabuk, A. (2020). Akıllı Şehirler Kapsamında Yapay Zekâ Teknikleri Kullanarak Etkin Ulaşım Planlarının Oluşturulması Üzerine Bir Model Önerisi. *GSI Journals Serie C: Advancements in Information Sciences and Technologies, 3*(1), 1–21.

Batura, O., Regeczi, D., Vassilev, A., Yagafarova, A., Bani, E., Bonneau, V., . . . DE STREEL, A. (2021). *Artificial intelligence in road transport: annex to cost of non-Europe report*. Retrieved from http://www.europarl.europa.eu/RegData/etudes/STUD/2021/654212/EPRS_STU(2021)654212(ANN1)_EN.pdf

Bitam, S., & Mellouk, A. (2012). ITS-cloud: Cloud computing for Intelligent transportation system. *2012 IEEE Global Communications Conference (GLOBECOM)*, 2054–2059. https://doi.org/10.1109/GLOCOM.2012.6503418

Bughin Jacques, Seong Jeongmin, Manyika James, Chui Michael, & Joshi Raoul. (2018, September 4). Notes from the AI frontier: Modeling the impact of AI on the world economy. Retrieved June 9, 2023, from McKinsey Global Institute website: https://www.mckinsey.com/featured-insights/artificial-intelligence/notes-from-the-ai-frontier-modeling-the-impact-of-ai-on-the-world-economy

Chase, N., Maples, J., & Schipper, M. (2018). Autonomous vehicles: Uncertainties and energy implications. *2018 EIA Energy Conference*.

Çevik, K. K. (2010). *Yapay Zekâ Yöntemleri ile Araç Plaka Tanıma Sistemi* (Master Thesis, Selçuk Üniversitesi Fen Bilimleri Enstitüsü). Selçuk Üniversitesi Fen Bilimleri Enstitüsü. Retrieved from https://acikbilim.yok.gov.tr/bitstream/handle/20.500.12812/463051/yokAcikBilim_381549.pdf?sequence=-1&isAllowed=y

Delipetrev, B., Tsinaraki, C., & Kostic, U. (2020). *Historical evolution of artificial intelligence*.

Demir, K. (2023). *Akıllı ulaşım sistemleri ve yapay zekâ bağlamında ulaşım hikayelerinin oluşturulması için model önerisi* (Yüksek Lisans Tezi). Erciyes Üniversitesi-Fen Bilimleri Enstitüsü.

Derawi, M., Dalveren, Y., & Cheikh, F. A. (2020). Internet-of-things-based smart transportation systems for safer roads. *2020 IEEE 6th World Forum on Internet of Things* (WF-IoT), 1–4. IEEE.

Dilek, E., & Dener, M. (2023). Computer vision applications in intelligent transportation systems: a survey. *Sensors, 23*(6), 2938.

Divit. (2023). Divit Plaka Tanıma Sistemleri. Retrieved June 9, 2023, from http://www.divit.com.tr/#anchorpts website: http://www.divit.com.tr/#anchorpts

Doğan, E. (2007). *Regresyon Analizi ve Yapay Zekâ Yaklaşımı ile Türkiye ve Seçilen Bazı Büyük İller için Trafik Kaza Tahmin Modelleri* (Yüksek Lisans Tezi). Kırıkkale Üniversitesi Fen Bilimleri Enstitüsü.

Erdal, H. (2018). Yapay Zekâ Teknikleri ve Uzman Sistemlerin Karasal Akıllı Ulaşım Sistemlerinin Denetiminde Kullanımı. *Akıllı Ulaşım Sistemleri ve Uygulamaları Dergisi, 1*(1), 32–39.

ERTRAC. (2022). *Connected, Cooperative and Automated Mobility Roadmap*. Retrieved from www.ertrac.org

European Bank. (2019). *Disruptive technology and innovation in transport Policy paper on sustainable infrastructure*.

European Environment Agency. (2016, June 3). Front-running cities changing transport, improving quality of life. Retrieved June 13, 2023, from European Environment Agency website: https://www.eea.europa.eu/media/newsreleases/front-running-cities-changing-transport

Ford Otosan. (2019). *Platooning: Otonom Konvoy*. Retrieved June 9, 2023, from https://blog.ford.com.tr/platooning-otonom-konvoy

Gülsün, B., & Gonca, C. K. (2019). Adaptif Trafik Yönetim Sistemleri. *OHS ACADEMY, 2*(1), 32–40. Retrieved from https://dergipark.org.tr/tr/pub/ohsacademy/issue/44841/516737

Hashem, I. A. T., Chang, V., Anuar, N. B., Adewole, K., Yaqoob, I., Gani, A., . . . Chiroma, H. (2016). The role of big data in smart city. *International Journal of Information Management, 36*(5), 748–758. https://doi.org/10.1016/J.IJINFOMGT.2016.05.002

HGM. (2022a, September 27). *Bir Hizmet Olarak Hareketlilik Projesi Başlangıç Toplantısı Yapıldı*. T.C. Ulaştırma ve Altyapı Bakanlığı. Retrieved June 9, 2023, from https://hgm.uab.gov.tr/haberler/bir-hizmet-olarak-hareketlilik-projesi-baslangic-toplantisi-yapildi

HGM. (2022b, September 27). *Projelerimiz*. T.C. Ulaştırma ve Altyapı Bakanlığı. Retrieved June 9, 2023, from https://hgm.uab.gov.tr/projelerimiz

Hou, Y., Shi, H., Chen, N., Liu, Z., Wei, H., & Han, Q. (2022). Vision image monitoring on transportation infrastructures: a lightweight transfer learning approach. *IEEE Transactions on Intelligent Transportation Systems*.

Howarth, J. (2023, February 8). 57+ Amazing Artificial Intelligence Statistics (2023). Retrieved June 9, 2023, from https://explodingtopics.com/blog/ai-statistics website: https://explodingtopics.com/blog/ai-statistics

INTETRA. (2023). *Artificial Intelligence Digital Solution*. https://intetra.com.tr/en/artificial-intelligence/

ISSD. (2022). *Product Catalogs*. https://www.issd.com.tr/en/36484/

Kadiroğulları, G., Aksoy, B., Sayın, H., & Melek, Ö. (2020). Arıma Yapay Zekâ Yöntemi Kullanılarak Isparta İlindeki Örnek Bir Kavşak İçin Araç Sayısı ve Araç Geçiş Sürelerinin Tespiti. *Mühendislik Bilimleri ve Tasarım Dergisi, 8*(5), 11–24.

Krizhevsky, A., Sutskever, I., & Hinton, G. E. (2012). Imagenet classification with deep convolutional neural networks. Advances in neural information processing systems, 25.

Koptelov, A. (2022). *Smart Roads: How AI in Transportation Keeps Drivers Safe*. Retrieved from https://towardsdatascience.com/smart-roads-how-ai-in-transportation-keeps-drivers-safe-98e4dfd4a7e8

Leo Drive. (2023). *Uçtan uca otonom teknoloji oluşturma*. https://www.leodrive.ai/

Li, F., Trappey, A. J. C., Lee, C. H., & Li, L. (2022). Immersive technology-enabled digital transformation in transportation fields: A literature overview. *Expert Systems with Applications, 202*, 117459. https://doi.org/10.1016/J.ESWA.2022.117459

Litman, T. (2013). Transportation and Public Health. *Annual Review of Public Health, 34*. https://doi.org/10.1146/annurev-publhealth-031912-114502

Manyika James and Sneader Kevin. (2018, June 1). AI, automation, and the future of work: Ten things to solve for. Retrieved June 9, 2023, from https://www.mckinsey.com/featured-insights/future-of-work/ai-automation-and-the-future-of-work-ten-things-to-solve-for

Markoff, J. (2013, May 20). M.I.T. Scholar's 1949 Essay on Machine Age Is Found. Retrieved June 8, 2023, from https://www.nytimes.com/2013/05/21/science/mit-scholars-1949-essay-on-machine-age-is-found.html

Medina-Tapia, M., & Robusté, F. (2018). Exploring paradigm shift impacts in urban mobility: Autonomous Vehicles and Smart Cities. *Transportation Research Procedia, 33*, 203–210. https://doi.org/10.1016/J.TRPRO.2018.10.093

MIA Teknoloji. (2024). *Çözümler*. https://www.miateknoloji.com/cozumler/

MOSAŞ. (2024). *Kameralı Akıllı Kavşak Yönetim Sistemi*. https://www.mosas.com.tr/sinyalizasyon/cozumler/cyclops/kamera/

Narbay, Ş., & Kirazlı, Ş. N. (2023). Otonom Araçlarda Yapay Zekâ, Kişisel Verilerin İşlenmesi ve Sonuçları. *Sakarya Üniversitesi Hukuk Fakültesi Dergisi, 11*(1), 49–66.

Neovision. (2021). *Yeni Nesil Veri Toplama ve Analitiği*. https://neovisiontr.com/

Nikitas, A., Michalakopoulou, K., Njoya, E. T., & Karampatzakis, D. (2020). Artificial Intelligence, Transport and the Smart City: Definitions and Dimensions of a New Mobility Era. *Sustainability, 12*(7). https://doi.org/10.3390/su12072789

Oladimeji, D., Gupta, K., Kose, N. A., Gundogan, K., Ge, L., & Liang, F. (2023). Smart Transportation: An Overview of Technologies and Applications. *Sensors, 23*(8). https://doi.org/10.3390/s23083880

Önder, H., & Akdemir, F. (2020). Kentsel Ulaşımın Dijital Boyutu: Ulaşım 4.0. Journal, 3(2), 202–215.

Özmen, M. M., Eylence, M., Şenol, R., & Aksoy, B. (2022). Yapay Zekâ Yöntemleriyle Araç Altı Yabancı Madde Tespit Edilmesi ve Otomatik Geçiş Sistemi. *El-Cezeri, 9*(4), 1495–1505.

Palandız, T., Bayrakçı, H. C., & Özkahraman, M. (2021). Yapay Zekâ Kullanılarak Trafik İşaret Levhalarının Sınıflandırılması: Denizli İl Merkezi İçin Örnek Bir Uygulama. *International Journal of 3D Printing Technologies and Digital Industry, 5*(3), 645–653.

Pan, G., & Alouini, M.-S. (2021). Flying Car Transportation System: Advances, Techniques, and Challenges. *IEEE Access, 9*, 24586–24603. https://doi.org/10.1109/ACCESS.2021.3056798

Pazar, Ş., Bulut, M., & Uysal, C. (2020). Yapay Zekâ Tabanlı Araç Algılama Sistemi Geliştirilmesi. *Journal of Science, Technology and Engineering Research, 1*(1), 31–37.

Pendik Belediyesi. (2015, March 16). Bozuk yola 'cep'ten uyarı. Retrieved June 9, 2023, from https://www.pendik.bel.tr/haber/detay/bozuk-yola-cepten-uyari website: https://www.pendik.bel.tr/haber/detay/bozuk-yola-cepten-uyari

Pham, Q. (2018). *Autonomous vehicles and their impact on road transportations*.

Rao, A., Verweij, G., & Cameron, E. (2020). *Sizing the prize What's the real value of AI for your business and how can you capitalise?* Retrieved from https://www.pwc.com/gx/en/issues/analytics/assets/pwc-ai-analysis-sizing-the-prize-report.pdf

Rudskoy, A., Ilin, I., & Prokhorov, A. (2021). Digital Twins in the Intelligent Transport Systems. *Transportation Research Procedia, 54*, 927–935. https://doi.org/10.1016/J.TRPRO.2021.02.152

Sarı, F. (2021). Cahit Arf'in "Makine Düşünebilir mi ve Nasıl Düşünebilir?" Adlı Makalesi Üzerine Bir Çalışma. *TRT Akademi, 6*(13), 812–833.

Schneider, M., Kutila, M., & Hoess, A. (2021). *Applications of AI in Transportation Industry*. https://doi.org/10.1201/9781003337232-29

Schwarting, W., Alonso-Mora, J., & Rus, D. (2018). Planning and Decision-Making for Autonomous Vehicles. *Annual Review of Control, Robotics, and Autonomous Systems, 1*(1), 187–210. https://doi.org/10.1146/annurev-control-060117-105157

Şafak, C. B. (2024). *Araçlarda Sürücü Destek Amaçlı Yapay Zeka Yöntemlerinin Uygulanması* (Elektrik ve Bilgisayar Mühendisliği Tezli Yüksek Lisans Programı, KTO Karatay Lisansüstü Eğitim Enstitüsü). KTO Karatay Lisansüstü Eğitim Enstitüsü, Konya. Retrieved from https://acikerisim.karatay.edu.tr/handle/20.500.12498/6098

Talih, Ö., & Tektaş, N. (2023). Bir Hizmet Olarak Hareketlilik-MaaS Perspektifi ve Türkiye Analizi. *Journal, 7*(2), 431–463.

T.C. Çevre Şehircilik ve İklim Değişikliği Bakanlığı. (2022). *Şehirlerin Dijital İkizleri*. Ankara. Retrieved from https://cbs.csb.gov.tr/sehirlerin-dijital-ikizlerini-guncelliyoruz-haber-278818

T.C. Cumhurbaşkanlığı Dijital Dönüşüm Ofisi. (2020). *Yeni Nesil Ulaşım Teknolojisi Hyperloop*. Retrieved from https://cbddo.gov.tr/SharedFolderServer/Genel/3.Aras%CC%A7t%C4%B1rma-Raporu-Yeni-Nesil-Ulas%CC%A7%C4%B1m-Teknolojisi-Hyperloop.pdf

T.C. Cumhurbaşkanlığı Dijital Dönüşüm Ofisi. (2021). *Ulusal Yapay Zeka Stratejisi 2021-2025*. Retrieved from https://cbddo.gov.tr/SharedFolderServer/Genel/File/TR-UlusalYZStratejisi2021-2025.pdf

T.C. Sanayi ve Teknoloji Bakanlığı. (2019). *2023 Sanayi ve Teknoloji Stratejisi*. Ankara. Retrieved from https://www.sanayi.gov.tr/assets/pdf/SanayiStratejiBelgesi2023.pdf

T.C. Sanayi ve Teknoloji Bakanlığı. (2022). *Mobilite Araç ve Teknolojileri Yol Haritası*. Retrieved from https://www.sanayi.gov.tr/assets/pdf/plan-program/MobiliteAracveTeknolojileriYolHaritasi.pdf

T.C. Ticaret Bakanlığı. (2021). *Yeşil Mutabakat Eylem Planı*. Retrieved from https://ticaret.gov.tr/data/60f1200013b876eb28421b23/MUTABAKAT%20YE%C5%9E%C4%B0L.pdf

T.C. Ulaştırma ve Altyapı Bakanlığı. (2020). *Ulusal AUS Strateji Belgesi ve 2020-2023 Eylem Planı*. Retrieved from https://hgm.uab.gov.tr/uploads/pages/akilli-ulasim-sistemler-aus/ulusal-akilli-ulasim-sistemleri-strateji-belgesi-ve-2020-2023-eylem-plani.pdf

T.C. Ulaştırma ve Altyapı Bakanlığı. (2022a). Enlerin, İlklerin ve Rekorların Projesi 1915 Çanakkale Köprüsü Yarın Açılıyor. Retrieved March 4, 2024, from https://www.uab.gov.tr/haberler/ulastirma-ve-altyapi-bakani-karaismailoglu-enlerin-ilklerin-ve-rekorlarin-projesi-1915-canakkale-koprusu-yarin-aciliyor

T.C. Ulaştırma ve Altyapı Bakanlığı. (2022b). *T.C. Ulaştırma ve Altyapı Bakanlığı Stratejik Plan 2019-2023*. Retrieved from https://www.uab.gov.tr/uploads/pages/stratejik-yonetim/uab-stratejik-plani-guncellenmis-versiyon-16-09-2021.pdf

T.C. Ulaştırma ve Altyapı Bakanlığı. (2023a). *Araç İçi Bilgi ve Haberleşme Sisteminin (ABHS) Mevcut Durum Analiz Raporu*. Ankara. Retrieved from https://hgm.uab.gov.tr//uploads/pages/akilli-ulasim-sistemleri-aus/abhs-mevcut-durum-analizi-raporu.pdf

T.C. Ulaştırma ve Altyapı Bakanlığı. (2023b). Demiryollarında Yapay Zekâ Sistemi. Retrieved March 4, 2024, from https://www.uab.gov.tr/haberler/demiryollarinda-yapay-zeka-sistemi

T.C. Ulaştırma ve Altyapı Bakanlığı. (2023c). *K-AUS Türkiye Mevcut Durum Analiz Raporu*. Ankara. Retrieved from https://hgm.uab.gov.tr/uploads/pages/akilli-ulasim-sistemleri-aus/k-aus-turkiye-mevcut-durum-analizi.pdf

T.C. Ulaştırma ve Altyapı Bakanlığı. (2023d). *Ulaştırma ve Haberleşme Şura Raporları*. Retrieved from https://sgb.uab.gov.tr/suralar

Tektaş, M., Akbaş, A., & Topuz, V. (2002). Yapay zekâ tekniklerinin trafik kontrolünde kullanılması üzerine bir inceleme. *Uluslararası Trafik ve Yol Güvenliği Kongresi, Gazi Üniversitesi*, Ankara.

Toğaç, M. G. (2023). *Gaziantep province artificial intelligence basedintelligent transportation systems, adaptive signalizationcontrol and simulation* (Yüksek Lisans). Gaziantep İslam Bilim ve Teknoloji Üniversitesi-Lisansüstü Eğitim Enstitüsü.

Toh, C. K., Sanguesa, J. A., Cano, J. C., & Martinez, F. J. (2020). Advances in smart roads for future smart cities. *Proceedings of the Royal Society A: Mathematical, Physical and Engineering Sciences, 476*(2233), 20190439. https://doi.org/10.1098/rspa.2019.0439

Torre-Bastida, A. I., Del Ser, J., Laña, I., Ilardia, M., Bilbao, M. N., & Campos-Cordobés, S. (2018). Big Data for transportation and mobility: recent advances, trends and challenges. *IET Intelligent Transport Systems, 12*(8), 742–755. https://doi.org/https://doi.org/10.1049/iet-its.2018.5188

Transforming Transport. (2017). *MOBILITY MEETS BIG DATA*. Retrieved from https://transformingtransport.eu/sites/default/files/2017-07/TT_BROCHURE_WEB.pdf

Ulaşım Yönetim Merkezi. (2022, December 27). Ulaşımda 'Yapay Zekâ' Dönemi. Retrieved March 5, 2024, from İstanbul Büyükşehir Belediyesi website: https://uym.ibb.gov.tr/kurumsal/haberler-ve-duyurular/ula%C5%9F%C4%B1mda-yapay-zek%C3%A2-d%C3%B6nemi

Ulu, M. (2023). *Gaziantep ili yapay zekâ tabanlı akıllı ulaşım sistemleri ile adaptif sinyalizasyon kontrolü ve simülasyonu* (Doktora Tezi). İstanbul Üniversitesi-Cerrahpaşa Lisansüstü Eğitim Enstitüsü.

Vega, E. D. La. (2023). *Drones And the Future of Transportation*. Retrieved from https://www.futureelectronics.com/blog/article/drones-and-the-future-of-transportation/

Yuan, T., da Rocha Neto, W., Rothenberg, C. E., Obraczka, K., Barakat, C., & Turletti, T. (2022). Machine learning for next-generation intelligent transportation systems: A survey. *Transactions on Emerging Telecommunications Technologies, 33*(4), e4427.

How cite this article

Dilek, E., Talih, Ö., & Kaya Bensghir, T. (2024). An overview of paradigm shift dynamics in transportation: Use of artificial intelligence in intelligent transportation systems in Türkiye. *Acta Infologica, 8*(1), 71-87. https://doi.org/10.26650/acin.1389571

DESCRIPTION

Acta Infologica (ACIN) is the publication of Informatics Department of the Istanbul University. It is an open access, scholarly, peerreviewed journal published biannually in June and December. The journal was founded in 2017.

AIM AND SCOPE

ACIN aims to contribute to the scientific community interested in the field of informatics and aims to provide a platform for researchers exploring issues based on the concepts of data-information-knowledge, information and communication technologies and applications. The journal welcomes multidisciplinary studies regarding the field as well.
The areas of study covered in the scope of ACIN are in below;
Intelligent Systems
Information Security and Law
Knowledge Management
Computer Networks
Computer Architecture
Information Systems
Bioinformatics
Geographic Information Systems
E-Applications
Internet Technologies
Decision Support Systems and Business Intelligence
Microcontroller and Applications
Mobile Systems
Modeling and Optimization
Project Management
Social and Digital Media
Data Mining
Database Systems
Artificial Intelligence and Machine Learning
Software Engineering

EDITORIAL POLICIES AND PEER REVIEW PROCESS

Publication Policy

The subjects covered in the manuscripts submitted to the Journal for publication must be in accordance with the aim and scope of the journal. The journal gives priority to original research papers submitted for publication.

General Principles

Only those manuscripts approved by its every individual author and that were not published before in or sent to another journal, are accepted for evaluation.

Submitted manuscripts that pass preliminary control are scanned for plagiarism using iThenticate software. After plagiarism check, the eligible ones are evaluated by editor-in-chief for their originality, methodology, the importance of the subject covered and compliance with the journal scope.

Short presentations that took place in scientific meetings can be referred if indicated in the article. The editor hands over the papers matching the formal rules to at least two national/international referees for evaluation and gives green light for publication upon modification by the authors in accordance with the referees' claims. Changing the name of an author (omission, addition or order) in papers submitted to the Journal requires written permission of all declared authors. Refused manuscripts and graphics are not returned to the author.

Open Access Statement

The journal is an open access journal and all content is freely available without charge to the user or his/her institution. Except for commercial purposes, users are allowed to read, download, copy, print, search, or link to the full texts of the articles in this journal without asking prior permission from the publisher or the author. This is in accordance with the BOAI definition of open access.

The open access articles in the journal are licensed under the terms of the Creative Commons Attribution-NonCommercial 4.0 International (CC BY-NC 4.0) license.

Article Processing Charge

All expenses of the journal are covered by the Istanbul University. Processing and publication are free of charge with the journal. There is no article processing charges or submission fees for any submitted or accepted articles.

Peer Review Process

Only those manuscripts approved by its every individual author and that were not published before in or sent to another journal, are accepted for evaluation.

Submitted manuscripts that pass preliminary control are scanned for plagiarism using iThenticate software. After plagiarism check, the eligible ones are evaluated by Editor-in-Chief for their originality, methodology, the importance of the subject covered and compliance with the journal scope. Editor-in-Chief evaluates manuscripts for their scientific content without regard to ethnic origin, gender, sexual orientation, citizenship, religious belief or political philosophy of the authors and ensures a fair double-blind peer review of the selected manuscripts.

The selected manuscripts are sent to at least two national/international external referees for evaluation and publication decision is given by Editorin- Chief upon modification by the authors in accordance with the referees' claims.

Editor-in-Chief does not allow any conflicts of interest between the authors, editors and reviewers and is responsible for final decision for publication of the manuscripts in the Journal.

Reviewers' judgments must be objective. Reviewers' comments on the following aspects are expected while conducting the review.

- Does the manuscript contain new and significant information?
- Does the abstract clearly and accurately describe the content of the manuscript?
- Is the problem significant and concisely stated?
- Are the methods described comprehensively?
- Are the interpretations and consclusions justified by the results?
- Is adequate references made to other Works in the field?
- Is the language acceptable?

Reviewers must ensure that all the information related to submitted manuscripts is kept as confidential and must report to the editor if they are aware of copyright infringement and plagiarism on the author's side.

Areviewer who feels unqualified to review the topic of a manuscript or knows that its prompt review will be impossible should notify the editor and excuse himself from the review process.

The editor informs the reviewers that the manuscripts are confidential information and that this is a privileged interaction. The reviewers and editorial board cannot discuss the manuscripts with other persons. The anonymity of the referees is important.

COPYRIGHT NOTICE

Authors publishing with the journal retain the copyright to their work licensed under the Creative Commons Attribution-NonCommercial 4.0 International license (CC BY-NC 4.0) (https://creativecommons.org/licenses/by-nc/4.0/) and grant the Publisher non-exclusive commercial right to publish the work. CC BY-NC 4.0 license permits unrestricted, non-commercial use, distribution, and reproduction in any medium, provided the original work is properly cited.

PUBLICATION ETHICS AND PUBLICATION MALPRACTICE STATEMENT

Acta Infologica (ACIN) is committed to upholding the highest standards of publication ethics and pays regard to Principles of Transparency and Best Practice in Scholarly Publishing published by the Committee on Publication Ethics (COPE), the Directory of Open Access Journals (DOAJ), to access the Open Access Scholarly Publishers Association (OASPA), and the World Association of Medical Editors (WAME) on https://publicationethics.org/resources/guidelines-new/principles-transparency-and-best-practice-scholarly-publishing All parties involved in the publishing process (Editors, Reviewers, Authors and Publishers) are expected to agree on the following ethical principles.

All submissions must be original, unpublished (including as full text in conference proceedings), and not under the review of any other publication synchronously. Each manuscript is reviewed by one of the editors and at least two referees under double-blind peer review process. Plagiarism, duplication, fraud authorship/denied authorship, research/data fabrication, salami slicing/salami publication, breaching of copyrights, prevailing conflict of interest are unnethical behaviors.

All manuscripts not in accordance with the accepted ethical standards will be removed from the publication. This also contains any possible malpractice discovered after the publication. In accordance with the code of conduct we will report any cases of suspected plagiarism or duplicate publishing.

RESEARCH ETHICS

Acta Infologica (ACIN) adheres to the highest standards in research ethics and follows the principles of international research ethics as defined below. The authors are responsible for the compliance of the manuscripts with the ethical rules.

- Principles of integrity, quality and transparency should be sustained in designing the research, reviewing the design and conducting the research.
- The research team and participants should be fully informed about the aim, methods, possible uses and requirements of the research and risks of participation in research.
- The confidentiality of the information provided by the research participants and the confidentiality of the respondents should be ensured. The research should be designed to protect the autonomy and dignity of the participants.
- Research participants should participate in the research voluntarily, not under any coercion.
- Any possible harm to participants must be avoided. The research should be planned in such a way that the participants are not at risk.
- The independence of research must be clear; and any conflict of interest or must be disclosed.
- In experimental studies with human subjects, written informed consent of the participants who decide to participate in the research must be obtained. In the case of children and those under wardship or with confirmed insanity, legal custodian’s assent must be obtained.
- If the study is to be carried out in any institution or organization, approval must be obtained from this institution or organization.
- In studies with human subject, it must be noted in the method’s section of the manuscript that the informed consent of the participants and ethics committee approval from the institution where the study has been conducted have been obtained.

AUTHOR RESPONSIBILITIES

It is authors’ responsibility to ensure that the article is in accordance with scientific and ethical standards and rules. And authors must ensure that submitted work is original. They must certify that the manuscript has not previously been published elsewhere or is not currently being considered for publication elsewhere, in any language. Applicable copyright laws and conventions must be followed. Copyright material (e.g. tables, figures or extensive quotations) must be reproduced only with appropriate permission and acknowledgement. Any work or words of other authors, contributors, or sources must be appropriately credited and referenced.

All the authors of a submitted manuscript must have direct scientific and academic contribution to the manuscript. The author(s) of the original research articles is defined as a person who is significantly involved in "conceptualization and design of the study", "collecting the data", "analyzing the data", "writing the manuscript", "reviewing the manuscript with a critical perspective" and "planning/conducting the study of the manuscript and/or revising it". Fund raising, data collection or supervision of the research group are not sufficient roles to be accepted as an author. The author(s) must meet all these criteria described above. The order of names in the author list of an article must be a co-decision and it must be indicated in the Copyright Agreement Form. The individuals who do not meet the authorship criteria but contributed to the study must take place in the acknowledgement section. Individuals providing technical support, assisting writing, providing a general support, providing material or financial support are examples to be indicated in acknowledgement section.

All authors must disclose all issues concerning financial relationship, conflict of interest, and competing interest that may potentially influence the results of the research or scientific judgment.

When an author discovers a significant error or inaccuracy in his/her own published paper, it is the author's obligation to promptly cooperate with the Editor to provide retractions or corrections of mistakes.

RESPONSIBILITY FOR THE EDITOR AND REVIEWERS

Editor-in-Chief evaluates manuscripts for their scientific content without regard to ethnic origin, gender, sexual orientation, citizenship, religious belief or political philosophy of the authors. He/She provides a fair double-blind peer review of the submitted articles for publication and ensures that all the information related to submitted manuscripts is kept as confidential before publishing.

Editor-in-Chief is responsible for the contents and overall quality of the publication. He/She must publish errata pages or make corrections when needed.

Editor-in-Chief does not allow any conflicts of interest between the authors, editors and reviewers. Only he has the full authority to assign a reviewer and is responsible for final decision for publication of the manuscripts in the Journal.

Reviewers must have no conflict of interest with respect to the research, the authors and/or the research funders. Their judgments must be objective.

Reviewers must ensure that all the information related to submitted manuscripts is kept as confidential and must report to the editor if they are aware of copyright infringement and plagiarism on the author's side.

A reviewer who feels unqualified to review the topic of a manuscript or knows that its prompt review will be impossible should notify the editor and excuse himself from the review process.

The editor informs the reviewers that the manuscripts are confidential information and that this is a privileged interaction. The reviewers and editorial board cannot discuss the manuscripts with other persons. The anonymity of the referees must be ensured. In particular situations, the editor may share the review of one reviewer with other reviewers to clarify a particular point.

MANUSCRIPT ORGANIZATION

LANGUAGE

The publication language of the journal is English.

Manuscript Organization and Submission

All correspondence will be sent to the first-named author unless otherwise specified. Manuscpript is to be submitted online via dergipark. org.tr/login that can be accessed at http://acin.istanbul.edu.tr and it must be accompanied by a title page specifying the article category (i.e. research article, review etc.) and including information about the manuscript (see the Submission Checklist) and cover letter to the editor. Manuscripts should be prepared in Microsoft Word 2003 and upper versions. In addition, Copyright Agreement Form that has to be signed by all authors must be submitted.

1. Use ACIN article document as a template if you are using Microsoft Word 6.0 or upper versions. Otherwise, use this document as an instruction set.

2. The first letters of words in the article title should be written in uppercase; the entire title should not be capitalized. Avoid writing formulas in the title. Do not write "(Invited)" or similar expressions in the title.
3. The abstract must be between 150–250 words and written as one paragraph. It should not contain displayed mathematical equations or tabular material. The abstract should include three to five different keywords or phrases, as this will help readers to find it. It is important to avoid over-repetition of such phrases as this can result in a page being rejected by search engines. Ensure that your abstract reads well and is grammatically correct.
4. Underneath the abstracts, a minimum of 3 and a maximum of 5 keywords that inform the reader about the content of the study should be specified. Keywords must be defined by taking into consideration authorities like "TR Dizin Anahtar Terimler Listesi", "Medical Subject Headings", "CAB Theasarus", "JISCT, "ERIC", etc.
5. The manuscripts should contain mainly these components: title, abstract and keywords; sections, references, tables and figures.
6. A title page including author information must be submitted together with the manuscript. The title page is to include fully descriptive title of the manuscript and, affiliation, title, e-mail address, ORCID, postal address, phone, mobile phone and fax number of the author(s) (see The Submission Checklist).
7. References should be prepared as APA 6th edition.

REFERENCES

Reference Style and Format

Acta Infologica (ACIN) complies with APA (American Psychological Association) style 6th Edition for referencing and quoting. For more information:

- American Psychological Association. (2010). Publication manual of the American Psychological Association (6th ed.). Washington, DC: APA.
- http://www.apastyle.org

Accuracy of citation is the author's responsibility. All references should be cited in text. Reference list must be in alphabetical order. Type references in the style shown below.

Citations in the Text

Citations must be indicated with the author surname and publication year within the parenthesis.
If more than one citation is made within the same paranthesis, separate them with (;).

Samples:

More than one citation;
(Esin, et al., 2002; Karasar, 1995)
Citation with one author;
(Akyolcu, 2007)
Citation with two authors;
(Sayıner & Demirci, 2007)
Citation with three, four, five authors;
First citation in the text: (Ailen, Ciambrune, & Welch, 2000) Subsequent citations in the text: (Ailen, et al., 2000)
Citations with more than six authors;
(Çavdar, et al., 2003)

Citations in the Reference

All the citations done in the text should be listed in the References section in alphabetical order of author surname without numbering. Below given examples should be considered in citing the references.

Basic Reference Types

Book

a) Turkish Book
Karasar, N. (1995). *Araştırmalarda rapor hazırlama* (8th ed.) [Preparing research reports]. Ankara, Turkey: 3A Eğitim Danışmanlık Ltd.
b) Book Translated into Turkish
Mucchielli, A. (1991). *Zihniyetler* [Mindsets] (A. Kotil, Trans.). İstanbul, Turkey: İletişim Yayınları.
c) Edited Book
Ören, T., Üney, T., & Çölkesen, R. (Eds.). (2006). *Türkiye bilişim ansiklopedisi* [Turkish Encyclopedia of Informatics]. İstanbul, Turkey: Papatya Yayıncılık.
d) Turkish Book with Multiple Authors
Tonta, Y., Bitirim, Y., & Sever, H. (2002). *Türkçe arama motorlarında performans değerlendirme* [Performance evaluation in Turkish search engines]. Ankara, Turkey: Total Bilişim.
e) Book in English
Kamien R., & Kamien A. (2014). *Music: An appreciation*. New York, NY: McGraw-Hill Education.
f) Chapter in an Edited Book
Bassett, C. (2006). Cultural studies and new media. In G. Hall & C. Birchall (Eds.), *New cultural studies: Adventures in theory* (pp. 220–237). Edinburgh, UK: Edinburgh University Press.
g) Chapter in an Edited Book in Turkish
Erkmen, T. (2012). Örgüt kültürü: Fonksiyonları, öğeleri, işletme yönetimi ve liderlikteki önemi [Organization culture: Its functions, elements and importance in leadership and business management]. In M. Zencirkıran (Ed.), *Örgüt sosyolojisi* [Organization sociology] (pp. 233–263). Bursa, Turkey: Dora Basım Yayın.
h) Book with the same organization as author and publisher
American Psychological Association. (2009). *Publication manual of the American psychological association* (6th ed.). Washington, DC: Author.

Article

a) Turkish Article
Mutlu, B., & Savaşer, S. (2007). Çocuğu ameliyat sonrası yoğun bakımda olan ebeveynlerde stres nedenleri ve azaltma girişimleri [Source and intervention reduction of stress for parents whose children are in intensive care unit after surgery]. *Istanbul University Florence Nightingale Journal of Nursing, 15*(60), 179–182.
b) English Article
de Cillia, R., Reisigl, M., & Wodak, R. (1999). The discursive construction of national identity. *Discourse and Society, 10*(2), 149–173. http://dx.doi.org/10.1177/0957926599010002002
c) Journal Article with DOI and More Than Seven Authors Lal, H., Cunningham, A. L., Godeaux, O., Chlibek, R., Diez-Domingo, J., Hwang, S.-J. ... Heineman, T. C. (2015). Efficacy of an adjuvanted herpes zoster subunit vaccine in older adults. *New England Journal of Medicine, 372*, 2087–2096. http://dx.doi. org/10.1056/NEJMoa1501184
d) Journal Article from Web, without DOI
Sidani, S. (2003). Enhancing the evaluation of nursing care effectiveness. *Canadian Journal of Nursing Research, 35*(3), 26–38. Retrieved from http://cjnr.mcgill.ca
e) Journal Article wih DOI
Turner, S. J. (2010). Website statistics 2.0: Using Google Analytics to measure library website effectiveness. *Technical Services Quarterly, 27*, 261–278. http://dx.doi.org/10.1080/07317131003765910
f) Advance Online Publication
Smith, J. A. (2010). Citing advance online publication: A review. *Journal of Psychology.* Advance online publication. http://dx.doi. org/10.1037/a45d7867
g) Article in a Magazine Henry, W. A., III. (1990, April 9). Making the grade in today's schools. *Time, 135*, 28–31.

Doctoral Dissertation, Master's Thesis, Presentation, Proceeding

a) Dissertation/Thesis from a Commercial Database Van Brunt, D. (1997). *Networked consumer health information systems* (Doctoral dissertation). Available from ProQuest Dissertations and Theses database. (UMI No. 9943436)

b) Dissertation/Thesis from an Institutional Database
Yaylalı-Yıldız, B. (2014). *University campuses as places of potential publicness: Exploring the politicals, social and cultural practices in Ege University* (Doctoral dissertation). Retrieved from Retrieved from: http://library.iyte.edu.tr/tr/hizli-erisim/iyte-tez-portali
c) Dissertation/Thesis from Web
Tonta, Y. A. (1992). *An analysis of search failures in online library catalogs* (Doctoral dissertation, University of California, Berkeley). Retrieved from http://yunus.hacettepe.edu.tr/ tonta/yayinlar /phd/ickapak.html
d) Dissertation/Thesis abstracted in Dissertations Abstracts International
Appelbaum, L. G. (2005). Three studies of human information processing: Texture amplification, motion representation, and figureground segregation. *Dissertation Abstracts International: Section B. Sciences and Engineering, 65*(10), 5428.
e) Symposium Contribution
Krinsky-McHale, S. J., Zigman, W. B., & Silverman, W. (2012, August). Are neuropsychiatric symptoms markers of prodromal Alzheimer's disease in adults with Down syndrome? In W. B. Zigman (Chair), *Predictors of mild cognitive impairment, dementia, and mortality in adults with Down syndrome*. Symposium conducted at the meeting of the American Psychological Association, Orlando, FL.
f) Conference Paper Abstract Retrieved Online
Liu, S. (2005, May). *Defending against business crises with the help of intelligent agent based early warning solutions*. Paper presented at the Seventh International Conference on Enterprise Information Systems, Miami, FL. Abstract retrieved from http://www.iceis. org/iceis2005/abstracts_2005.htm
g) Conference Paper - In Regularly Published Proceedings and Retrieved Online
Herculano-Houzel, S., Collins, C. E., Wong, P., Kaas, J. H., & Lent, R. (2008). The basic nonuniformity of the cerebral cortex. *Proceedings of the National Academy of Sciences, 105*, 12593–12598. http://dx.doi.org/10.1073/pnas.0805417105
h) Proceeding in Book Form
Parsons, O. A., Pryzwansky, W. B., Weinstein, D. J., & Wiens, A. N. (1995). Taxonomy for psychology. In J. N. Reich, H. Sands, & A. N. Wiens (Eds.), *Education and training beyond the doctoral degree: Proceedings of the American Psychological Association National Conference on Postdoctoral Education and Training in Psychology* (pp. 45–50). Washington, DC: American Psychological Association.
i) Paper Presentation Nguyen, C. A. (2012, August). *Humor and deception in advertising: When laughter may not be the best medicine.* Paper presented at the meeting of the American Psychological Association, Orlando, FL.

Other Sources
a) Newspaper Article
Browne, R. (2010, March 21). This brainless patient is no dummy. *Sydney Morning Herald*, 45.
b) Newspaper Article with no Author
New drug appears to sharply cut risk of death from heart failure. (1993, July 15). *The Washington Post*, p. A12.
c) Web Page/Blog Post Bordwell, D. (2013, June 18). David Koepp: Making the world movie-sized [Web log post]. Retrieved from http://www.davidbordwell. net/blog/page/27/ *d) Online Encyclopedia/Dictionary*
Ignition. (1989). In *Oxford English online dictionary* (2nd ed.). Retrieved from http://dictionary.oed.com Marcoux, A. (2008). Business ethics. In E. N. Zalta (Ed.). The Stanford encyclopedia of philosophy. Retrieved from http://plato.stanford.edu/entries/ethics-business/
e) Podcast
Dunning, B. (Producer). (2011, January 12). *inFact: Conspiracy theories* [Video podcast]. Retrieved from http://itunes.apple.com/
f) Single Episode in a Television Series
Egan, D. (Writer), & Alexander, J. (Director). (2005). Failure to communicate. [Television series episode]. In D. Shore (Executive producer), *House*; New York, NY: Fox Broadcasting.
g) Music
Fuchs, G. (2004). Light the menorah. *On Eight nights of Hanukkah* [CD]. Brick, NJ: Kid Kosher.

SUBMISSION CHECKLIST

Ensure that the following items are present:

- Cover letter to the editor
  - The category of the manuscript
  - Confirming that “the paper is not under consideration for publication in another journal”.
  - Including disclosure of any commercial or financial involvement.
  - Confirming that last control for fluent English was done.
  - Confirming that journal policies detailed in Information for Authors have been reviewed.
  - Confirming that the references cited in the text and listed in the references section are in
  - line with APA 6.
- Copyright Agreement Form
- Permission of previous published material if used in the present manuscript
- Title page
  - The category of the manuscript
  - The title of the manuscript
  - All authors’ names and affiliations (institution, faculty/department, city, country),
  - e-mail addresses
  - Corresponding author’s email address, full postal address, telephone and fax number
  - ORCIDs of all authors.
- Main Manuscript Document
- The title of the manuscript
- Abstract (150-250 words)
- Key words: 3-5 words
- Grant support (if exists)
- Conflict of interest (if exists)
- Acknowledgement (if exists)
- References
- All tables, illustrations (figures) (including title, explanation, captions)

TELİF HAKKI ANLAŞMASI FORMU / COPYRIGHT AGREEMENT FORM

İstanbul Üniversitesi
Istanbul University

Dergi Adı: Acta INFOLOGICA (ACIN
Journal name: Acta INFOLOGICA (ACIN)

Telif Hakkı Anlaşması Formu
Copyright Agreement Form

Sorumlu Yazar *Responsible/Corresponding Author*	
Makalenin Başlığı *Title of Manuscript*	
Kabul Tarihi *Acceptance Date*	
Yazarların Listesi *List of Authors*	

Sıra *No*	**Adı-Soyadı** *Name - Surname*	**E-Posta** *E-Mail*	**İmza** *Signature*	**Tarih** *Date*
1				
2				
3				
4				
5				

Makalenin türü (Araştırma makalesi, Derleme, v.b.)
Manuscript Type (Research Article, Review, etc.)

Sorumlu Yazar:
Responsible/Corresponding Author:

Çalıştığı kurum	*University/company/institution*	
Posta adresi	*Address*	
E-posta	*E-mail*	
Telefon no; GSM no	*Phone; mobile phone*	

Yazar(lar) aşağıdaki hususları kabul eder:
Sunulan makalenin yazar(lar)ın orijinal çalışması olduğunu ve intihal yapmadıklarını,
Tüm yazarların bu çalışmaya asli olarak katılmış olduklarını ve bu çalışma için her türlü sorumluluğu aldıklarını,
Tüm yazarların sunulan makalenin son halini gördüklerini ve onayladıklarını,
Makalenin başka bir yerde basılmadığını veya basılmak için sunulmadığını,
Makalede bulunan metnin, şekillerin ve dokümanların diğer şahıslara ait olan Telif Haklarını ihlal etmediğini kabul ve taahhüt ederler.
İSTANBUL ÜNİVERSİTESİ'nin bu fikri eseri, Creative Commons Atıf-GayrıTicari 4.0 Uluslararası (CC BY-NC 4.0) lisansı ile yayınlamasına izin verirler. Creative Commons Atıf-GayrıTicari 4.0 Uluslararası (CC BY-NC 4.0) lisansı, eserin ticari kullanım dışında her boyut ve formatta paylaşılmasına, kopyalanmasına, çoğaltılmasına ve orijinal esere uygun şekilde atıfta bulunmak kaydıyla yeniden düzenleme, dönüştürme ve eserin üzerine inşa etme dâhil adapte edilmesine izin verir.
Yazar(lar)ın veya varsa yazar(lar)ın işvereninin telif dâhil patent hakları, fikri mülkiyet hakları saklıdır.
Ben/Biz, telif hakkı ihlali nedeniyle üçüncü şahıslarca vuku bulacak hak talebi veya açılacak davalarda İSTANBUL ÜNİVERSİTESİ ve Dergi Editörlerinin hiçbir sorumluluğunun olmadığını, tüm sorumluluğun yazarlara ait olduğunu taahhüt ederim/ederiz.
Ayrıca Ben/Biz makalede hiçbir suç unsuru veya kanuna aykırı ifade bulunmadığını, araştırma yapılırken kanuna aykırı herhangi bir malzeme ve yöntem kullanılmadığını taahhüt ederim/ederiz.
Bu Telif Hakkı Anlaşması Formu tüm yazarlar tarafından imzalanmalıdır/onaylanmalıdır. Form farklı kurumlarda bulunan yazarlar tarafından ayrı kopyalar halinde doldurularak sunulabilir. Ancak, tüm imzaların orijinal veya kanıtlanabilir şekilde onaylı olması gerekir.

The author(s) agrees that:
The manuscript submitted is his/her/their own original work and has not been plagiarized from any prior work,
all authors participated in the work in a substantive way and are prepared to take public responsibility for the work,
all authors have seen and approved the manuscript as submitted,
the manuscript has not been published and is not being submitted or considered for publication elsewhere,
the text, illustrations, and any other materials included in the manuscript do not infringe upon any existing copyright or other rights of anyone.
ISTANBUL UNIVERSITY will publish the content under Creative Commons Attribution-NonCommercial 4.0 International (CC BY-NC 4.0) license that gives permission to copy and redistribute the material in any medium or format other than commercial purposes as well as remix, transform and build upon the material by providing appropriate credit to the original work.
The Contributor(s) or, if applicable the Contributor's Employer, retain(s) all proprietary rights in addition to copyright, patent rights.
I/We indemnify ISTANBUL UNIVERSITY and the Editors of the Journals, and hold them harmless from any loss, expense or damage occasioned by a claim or suit by a third party for copyright infringement, or any suit arising out of any breach of the foregoing warranties as a result of publication of my/our article. I/We also warrant that the article contains no libelous or unlawful statements and does not contain material or instructions that might cause harm or injury.
This Copyright Agreement Form must be signed/ratified by all authors. Separate copies of the form (completed in full) may be submitted by authors located at different institutions; however, all signatures must be original and authenticated.

Sorumlu Yazar; *Responsible/Corresponding Author*	**İmza / Signature**	**Tarih** / *Date*
		……../……../……………