



95. yıl

ICI-CS2021

ULUSLARARASI BİLİŞİM VE BİLGİSAYAR BİLİMLERİ KONFERANSI



2022 1st International Conference On Informatics And Computer Science (ICI-CS2021)

October 9-11, 2022

Ankara, Turkey

ISBN: 978-975-507-301-9



Preface

2021 International Conference On Informatics and Computer Science (ICI-CS202) is a comprehensive conference which focuses on both informatics and computer science. The solution of complex global agendas such as ensuring sustainable development, accelerating digitalization in all areas, evaluating opportunities created by intelligent systems, improving industry practices, applying new technologies to real life, and overcoming challenges in technological developments does not seem possible a single field perspective. International Conference on Informatics and Computer Science (ICI-CS2021) held for the first time in December 2021 with participants from Turkey and abroad and hosted by Gazi University Institute of Informatics, aims to pave the way for innovative approaches for these agendas from an interdisciplinary perspective. We invited researchers, academics and students, and industry practitioners of Informatics and computer science and interdisciplinary fields to contribute to our conference.

ICI-CS2021 provides an opportunity for the scientists, engineers, industrialists, scholars and other professionals from all over the world to interact and exchange their new ideas and research outcomes in related fields and develop possible chances for future collaboration. We would like to express our grateful thanks to the conference program chairs and committee members, and all the reviewers for their great professionalism and efforts in reviewing the submitted papers. And thank all the participants and sponsors for their valuable contributions and support to ICI-CS2021.



Conference Chair

Prof. Dr. Aslıhan TÜFEKÇİ - Gazi University, Informatics Institute

Conference Committees

Honorary President

Prof. Dr. Musa YILDIZ - Gazi University Rector

Chairman of the Management Committee

Prof. Dr. Aslihan TÜFEKÇİ - Gazi University

Member of the Management Committee

Prof. Dr. Mahir DURSUN - Gazi University

Prof. Dr. Nursal ARICI - Gazi University

Prof. Dr. Mustafa ALKAN - Gazi University

Prof. Dr. M. Ali AKCAYOL - Gazi University

Prof. Dr. Hasan Şakir BİLGE - Gazi University

Prof. Dr. Serdar KULA - Gazi University

Assoc. Prof. Dr. Recep BENZER - Gazi University

Asst. Prof. Öner BARUT - Gazi University

Asst. Prof. Sinan TOKLU - Gazi University

Res. Asst. M. Ali KOŞAN - Gazi University

Res. Asst. Mehmet SEVRİ - Gazi University

Sedat HAKYEMEZ - Gazi University

Dr. Nadircan DAĞLI - Gazi University

Conference Secretary

Res. Asst. Tuana İRKEY - Gazi University

Res. Asst. Merve ORAKCI - Gazi University

Conference Steering Committee

Prof. Dr. Kemal LEBLEBİCİOĞLU - ODTÜ

Prof. Dr. Ferat ŞAHİN - Rochester Institute of Technology

Member of the Technical Committee

- Assoc. Prof. Hacer KARACAN - Gazi University
- Assoc. Prof. Mehmet ATAK - Gazi University
- Assoc. Prof. İbrahim Alper DOĞRU - Gazi University
- Asst. Prof. Sinan TOKLU - Gazi University
- Res. Ass. Uğurhan KUTBAY - Gazi University
- Prof. Dr. Fırat HARDALAÇ - Gazi University
- Prof. Dr. Ayhan ERDEM - Gazi University
- Assoc. Prof. Oktay YILDIZ - Gazi University
- Assoc. Prof Erkan BOSTANCI - Ankara University
- Asst. Prof. Ayetül GELEN - Bursa Technical University
- Asst. Prof. Nurettin Gökhan ADAR - Bursa Technical University
- Asst. Prof. Serdar ENGİNOĞLU - Çanakkale Onsekiz Mart University
- Assoc. Prof Zerrin IŞIK - Dokuz Eylül University
- Assoc. Prof. Abdullah Talha KABAKUŞ - Düzce University
- Assoc. Prof Orhan DAĞDEVİREN - Ege University
- Assoc. Prof Bilal BABAYİĞİT - Erciyes University
- Prof. Dr. Ahmet YAZICI - Eskişehir Osman Gazi University
- Prof. Dr. Nezih GEREK - Eskişehir Technical University
- Assoc. Prof. Ahmet Mete VURAL - Gaziantep University
- Asst. Prof. İpek ABASIKELEŞ-TURGUT - Iskenderun Technical University
- Asst. Prof. Uğur ERKAN - Karamanoglu Mehmetbey University
- Asst. Prof. RECEP SİNAN ARSLAN - Kayseri University
- Assoc. Prof Kazım YILDIZ - Marmara University
- Asst. Prof. Kemal UÇAK - Muğla Sıtkı Koçman University
- Asst. Prof. Anıl UTKU - Munzur University
- Assoc. Prof. Utku KÖSE - Süleyman Demirel University
- Assoc. Prof. Gökçe Nur YILMAZ - TED University
- Assoc. Prof Korhan KARABULUT - Yasar University
- Prof. Dr. Matthew SADIKU - Prairie View A&M University, USA

Assoc. Prof. Gülüstan DOĞAN - University of North Carolina Wilmington, USA
Asst. Prof. Abul K.M. AZAD - Northern Illinois University, USA
Assoc. Prof David HICKS - Texas A&M University - Kingsville, USA
Dr. Sait Ahmet BINSELAM - FM Global, Center for Property Risk Solutions, USA
Assoc. Prof Daniel MORRIS - Michigan State University, USA
Prof. Dr. Mehmet CELENK - Ohio University, USA
Assoc. Prof Priyanka SURENDRAN - AMA International University, Bahrain
Assoc. Prof Moharram CHALLENGER - University of Antwerp, Belgium
Assoc. Prof Jauhar ALI - Abu Dhabi University, UAE
Assoc. Prof Mostefa BELARBI - Ibn Khldoun University, Algeria
Assoc. Prof Mingqiang YANG - Shandong University, China
Asst. Prof. Sigit ADINUGROHO - Brawijaya University, Indonesia
Prof. Dr. Youness TABII - University of Abdelmalek Essaâdi, Tetouan, Morocco
Assoc. Prof Ankur Singh BIST - KIET Ghaziabad, India
Asst. Prof. Yazen SHEET - Mosul University, Iraq
Asst. Prof. Ghassan ALADOOL - University of Mosul, Iraq
Asst. Prof. Ann ALKAZZAZ - University of Mosul, Iraq
Prof. Dr. Amine BERMAK - HBKU, Qatar Foundation, Qatar
Prof. Dr. Young-Jin KIM - Kyung Hee University, Korea
Prof. Dr. A Nwar ALYATAMA - Kuwait University, Kuwait
Assoc. Prof Prof. Luay TAHAT - Gulf University for Science and Technology, Kuwait
Asst. Prof. Nada ALMASRI - Gulf University for Science and Technology, Kuwait
Asst. Prof. Nur Shamsiah ABDUL RAHMAN - Universiti Malaysia Pahang, Malaysia
Asst. Prof. Nor Saradatul Akmar ZULKIFLI - Universiti Malaysia Pahang, Malaysia
Asst. Prof. Anis Farihan Mat RAFFEI - Univeristi Malaysia Pahang, Malaysia
Asst. Prof. Syifak Izhar HĪSHAM - Universiti Malaysia Pahang, Malaysia
Asst. Prof. Mohd Fadzil ABDUL KADIR - University Sultan Zainal Abidin, Malaysia
Assoc. Prof Siti Zaleha Zainal ABĪDĪN - Universiti Teknologi Mara, Malaysia
Prof. Dr. Ahmed R. ABDELAZĪZ - Alexandria University, Mısır
Prof. Dr. Subarna SHAKYA - Tribhuvan University, Nepal
Assoc. Prof Faisal KHAN - COMSATS University Islamabad, Pakistan

Assoc. Prof Zeeshan BHATTI - University of Sindh, Pakistan

Prof. Dr. Adnan RASHDI - National University of Sciences and Technology (NUST), Islamabad, Pakistan

Assoc. Prof Ayesha SAADIA - Muslim Youth University (MYU), Islamabad, Pakistan

Assoc. Prof Cristian Paul CHIONCEL - Eftimie Murgu University, Romania

Asst. Prof. Mohammed Hayyan ALSIBAI - International University for Science and Technology, Syria

Assoc. Prof Talal M. AL-KHAROUBI - King Fahd University for Petroleum and Minerals, Saudi Arabia

Asst. Prof. Hani. K. ALMOHAIR - King Fahd University of Petroleum and Minerals, Saudi Arabia

Prof. Dr. Rehan Ullah KHAN - Qassim University, Saudi Arabia

Assoc. Prof Marlen ASANOV - V.I. Vernadsky Crimean Federal University, Ukraine

Assoc. Prof Maksym ZALISKYĬ - National Aviation University, Ukraine

Assoc. Prof Naoufel KRAIEM - Sultan Qaboos University, Oman

Assoc. Prof Ashraf SALEEM - Sultan Qaboos University, Oman

Assoc. Prof Serein AL-RATROUT - Al-Zahra College for Women, Oman

Assoc. Prof Ala KHALIFEH - German Jordanian University, Jordan

Prof. Dr. Ahmad HIASAT - Princess Sumaya University for Technology, Jordan

Assoc. Prof Wael TOGHUJ - Al-Ahliyya Amman University, Jordan

Assoc. Prof Dang Ngoc Hoang THANH - Hue College of Industry Hue, Vietnam

2022 1st International Conference On Informatics And Computer Science (ICI-CS2021) Table of Contents

Preface	ii
Conference Committees	iv
Virtual Reality Based Tower Crane Operator Training Simulator..... Mahir DURSUN, Kadir YILMAZ	1
Machine Learning Prediction Models of COVID-19 Prevalence in Turkey..... Hatice Nur CANPOLAT, Ertan GÜNER	8
Impact of Feature Configurations on Speech Recognition Using Listen, Attend and Spell Network..... Zekeriya TÜFEKÇİ, Gökay DİŞKEN	15
Performance of YOLO on a Custom Dataset of Infrared Laser Images..... Emre Ege EŞER , Fırat HARDALAÇ	19
ISO/IEC 27001 Information Security Management System and Cybersecurity Maturity Model Certification-CMMC Applications..... İsmail Yiğit COŞAR, Aslıhan TÜFEKÇİ	24
Natural Language Processing Based Recommendation System to Improve Patent Preliminary Search..... Yavuz YILMAZ, Tahsin ÇETİNYOKUŞ	31
Hair Removal and Lesion Segmentation with Fully Convolutional Neural Network and Image Processing in Images of Skin Cancer..... Cihan AKYEL, Nursal ARICI	35
The Importance and Challenges of Twenty-First Century Skills for the Information and Communication Sector..... Isra' Ala'Addin DWIEK, Hasan ÇAKIR	40
A Decision Support System Proposal for the Assignment of a Lawyer-Case in Law Offices. Nadide Gizem GÜRSON DOLAR, Başak GÖK, Hadi GÖKÇEN	48
Prediction of Life Quality Index Value Rankings of Countries After the COVID-19 Pandemic with Artificial Neural Networks..... Ali UÇUCU1, Başak GÖK, Hadi GÖKÇEN	57

An Information System Proposal for Determining of Material Picking Route and Picking Box Quantity in Warehouse.....	64
Volkan ÇETİN, Başak GÖK, Hadi GÖKÇEN	
Decision Support Systems in Engineering: A Bibliometric Analysis and Systematic Review..	72
Musab Talha Akpınar, Mehmet Atak, Uzeyir Fidan	
Deep Learning Based Forecasting Of Cancellation, Delay And Orientation On Flights.....	79
Anıl AYAYDIN, M. Ali AKCAYOL	
Comparison of AutoML Libraries in order to Use in Anomaly Detection Processes for Conveyor Line Used in Autonomous Production System.....	85
Cihan BAYRAKTAR, Hadi GÖKÇEN	
Sentiment, Stance, and Emotion Analysis on Twitter for COVID-19 Vaccination: A Brief Survey.....	91
Doğan Küçük, NursalArıcı	
Accessibility Evaluation of e-Nabız Web Site.....	96
Mehmet Fevzi Akdeniz, Pınar Onay Durdu	
Mask Based Image Encryption Using Chaotic Logistic Map.....	102
Hasan Hüseyin Sayan, Ali UZUN	
Deep Learning-Based Traffic Accidents Severity Forecast.....	109
Mustafa CAYMAZ, İ. Alper DOĞRU, Sinan TOKLU	
Systematic Literature Review of Detecting Topics and Communities in Social Networks.....	114
Ömer Ayberk ŞENCAN, İsmail ATACAK, İbrahim Alper DOĞRU	

Virtual Reality Based Tower Crane Operator Training Simulator

Mahir DURSUN¹, Kadir YILMAZ²

¹Department of Electrical and Electronics Engineering, Gazi University, Ankara, Turkey

² Education of Computer, Gazi University, Ankara, Turkey

mdursun@gazi.edu.tr, kadiryilmaza@gmail.com

Abstract— The tower crane is a very powerful lifting machine that is often used in the construction field. Its use is increasing day by day in line with the needs of the growing industry. Accordingly, the need for operators is also increasing. Due to its high structure and lifting heavy loads, the tower crane is in the category of very risky construction equipment. In terms of occupational health and safety, operators must be well-trained and experienced. The difficulties encountered during the practical implementation of these trainings in real environments reduce the quality of training. Virtual reality technologies allow the creation of a computer-generated three-dimensional environment that users can direct and interact with. With this study, it is aimed to develop a virtual reality training simulator that is compatible with real environments to make crane operators gain experience without endangering human life before practical training.

Keywords— virtual reality, VR, tower crane, operator, training simulator

1. INTRODUCTION

Construction machines are used a lot in many sectors. These machines are designed to move large blocks easily, and therefore, construction machines are very effective in the successful realization of construction projects [1]. In order for the developed machines to be used safely, trained and experienced operators are needed. Negative results can occur when heavy construction machines are used by operators who are not well trained or lack experience. These results cause great financial losses as well as health problems that cannot be compensated. In the evaluation made by the Construction Machinery Engineering Association in Turkey, it was stated that the majority of the operators were primary school graduates, but 15% were high school graduates [2]. This shows that operator training should be done more effectively. Demonstration method in educational sciences is a method of first showing and explaining the application of a process, the operation of a tool, and then teaching them by having the student drill and practice [3]. When the literature is examined, there are studies showing that the demonstration method is more effective in terms of student achievement than the traditional method [4-8]. Demonstration method is a widely used method in operator training. Trainings for the use of many machines are carried out by using the machine. Factors such as the high cost of these machines and the security risks in educational environments are the factors that prevent the quality of education from being realized at the desired level. Applications developed with virtual reality technologies appear as a powerful method to complete the deficiencies encountered at this point.

Virtual reality is defined as the use of a computer-generated three-dimensional environment that a user can direct and interact with [9]. In another study, it was defined as “computer simulation that displays a computer-generated virtual environment in which a person can interact with objects” [10]. The virtual environment is three-dimensional and often tries to replicate the real world view. It simulates the user's physical presence in an artificially created world and allows to interact with virtual objects. This feature of virtual reality technologies shows that it is a suitable environment for learning the use of a device or machine. In this way, these technologies can be used to meet the qualified personnel needed in different business areas. It is seen in the literature that the use of virtual reality in education has become widespread. For example, in parallel with the rapidly increasing need for renewable energy, the required installation personnel training will be provided in a shorter time and in a safer method, at the same standards with this technology [11].

One of the areas where heavy machinery is used intensively is the construction industry. In addition to being an important employment area in many parts of the world, the construction sector stands out with high rates of work accidents and occupational diseases [12]. Tower cranes, which we frequently encounter especially in high-rise buildings, are important in the rapid progress of the construction. Recent developments in the construction industry show that the use of tower cranes will increase rapidly. A study on tower cranes stated that “as construction projects grow and become more modular, the use of lifting equipment such as cranes and hoists is

expected to increase” [13]. Accordingly, the need for well-trained and experienced operators will continue to increase.

The common use of mobile devices in the 2000s contributed to the reach of virtual reality technologies to many people. Due to the high costs and low portability of virtual reality technologies, virtual reality applications could not go further than project-based studies until today. However, the spread of mobile devices has enabled the production of virtual reality devices and applications that can appeal to all segments of society. Virtual reality experience via mobile applications on smartphones with gyroscope, virtual reality glasses with their own integrated lenses are important steps in the field of virtual reality [14]. Today, the number of applications compatible with virtual reality glasses is increasing day by day. Many virtual reality applications are produced with virtual reality glasses such as Oculus Rift, Sony Playstation VR, Samsung Gear VR, Sony Morpheus, HTC vive. The application in this study was carried out with the Samsung Gear VR device. The biggest factor in choosing this device is that it is more affordable than its competitors in terms of price-performance. Developing the application with the most widely produced VR device in the market will ensure that the device can be obtained easily. As of 2007, Samsung Gear VR ranks first with 782.000 sales.

Practical trainings using heavy machinery have an important place in operator training. However, there are some limitations to these trainings. Tower cranes are the most expensive resources on construction sites [15]. In a study on the spatial and temporal planning of tower cranes in construction projects, it was stated that “tower cranes are important pieces of equipment in the construction industry and, although very expensive, their use is widespread” [16]. This high cost is a situation that hinders applied education. In addition, since the use of construction equipment is very dangerous, it is among the "very dangerous" work branches in terms of work and worker health. Tower crane height is high. Tower heights are between 20-60 meters. Crane arm lengths are between 6 meters and 30 meters, and lifting capacities are between 0.3 tons and 10 tons [17]. Occupational accidents related to tower cranes result in death to a large extent. The majority of work accidents are due to operator mistakes and lack of experience. However, experience can be gained on the job. In order to gain this experience, it is necessary to increase the number of applied trainings. However, reasons such as the high cost of the heavy machine, the high risk in work security in the environment where training is carried out with the heavy machine, the height of the tower, the bad weather conditions when training takes place outside, and the operator's cabin being wide enough for only one person decreases the training quality.

In the light of all this information, in this study, it is aimed to design a virtual reality-based training simulator that can be used as a supportive training tool for crane operators before practical training in real environments. The aim is

to provide a virtual training environment for trainees to gain experience without being exposed to life-threatening risks. In this study, a construction site and tower crane model is developed in a virtual environment. With this application, a safe classroom environment and the guidance of the trainer, the use of construction equipment, risk and accident scenarios will be taught to the trainees by experience.

In the second part of this study, information is given on the use of tower crane machines, occupational accidents encountered in work areas and the way of operator training.

In the third chapter, the development stages of a tower crane training simulator with virtual reality technologies are presented.

In the last part, the limitations of the tower crane operator training simulator and future studies are presented. It has been determined that this software can be used as a tool to increase operator development in training activities.

2. TOWER CRANE

2.1. Tower Crane Working Principle

Tower crane consists of many parts coming together. Tower cranes are named after their towering body [18]. The body consisting of steel structure, on which tower cranes are built, is called mast. The arm on which the load is carried at the top of the mast is called the jib. The carriage that allows the hook to move back and forth on the jib is called a trolley. A hook attached to the trolley with steel ropes provides up and down movement. The tower crane basically makes 3 different movements. These; turn movement left and right thanks to the jib, back and forth movement with the trolley, up and down movement with the hook.

Tower cranes are exposed to heavy forces due to their structure and the weight of the load they carry. Additional weights are used to balance these forces. In order to balance the weight on the load arm and prevent the mast from tipping over, the counter arm on the mast top, and balance stones on this arm are used [19]. If not used correctly, the size of the work machine itself and the weight of the load it carries cause serious work accidents. An inexperienced operator is faced with unexpected forces during use. The knowledge of what the operator should do in the face of these situations is the gain that can only be learned by experience.

As seen in Figure 1., the report published by EU-OSHA (European Union Occupational Health and Safety agency) in 2010 shows occupational accidents resulting in death [20]. EU-OSHA shows that 38% of electric shocks from machinery used in construction are from cranes. Occupational accidents that occur as a result of works carried out near electric lines constitute the majority of

these accidents[21]. The best way to reduce the rate of these accidents, which are caused by the operator's failure to predict the swing of the tower crane rope, is through better training of the tower crane operator, who is primarily responsible for the tower crane.

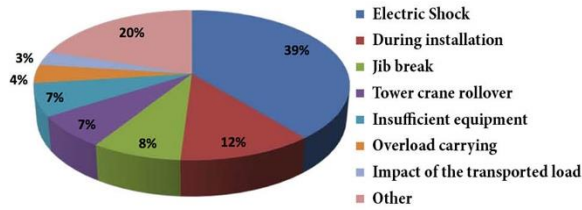


Figure 1. Analysis of work accident resulting in crane-related death [20].

According to data published by Towercranesupport.com (2021), 1125 accidents occurred worldwide between 2000 and 2010, and 780 of these accidents were recorded as fatal accidents [22]. Tower crane usage is provided by control sticks inside the cabin. The right and left movement of the right stick provides the jib to turn to the right and left. The forward and backward movement of the same control unit provides the forward and backward movement of the trolley. With the forward and backward movement of the stick on the left side, the hook moves up and down. All these movements should be started from slow to fast, and when the target point is approached, the movement should be terminated gradually, from fast to slow. The majority of work accidents result in the crashing, knocking or falling of the load against the surrounding elements in the construction site as a result of the operator not taking into account the moment caused by the weight of the load during these movements.

2.2. Tower Crane Operator Training Program

Tower crane operator training is given by private institutions affiliated to the Ministry of National Education in Turkey. The training programs given in these courses are based on the Lifting-Elevating (Tower Crane) Operator Training Program prepared by the General Directorate of Private Education Institutions and approved by the Board of Education. In this course program, there are group practices with the actual use of heavy machine, and personal practices.

The reasons such as climbing a machine at this height by a person without tower crane experience, the excitement caused by this process being done for the first time, and adverse weather conditions reduce the effectiveness of the information to be given to the trainees in the trainings

where the machine is used. The biggest disadvantage is that it is very difficult and risky to imitate the situations that the trainee will encounter in the field of work in a real environment. Virtual reality simulation will enable the operator to learn by experiencing scenarios prepared in accordance with real environments.

3. DESIGN AND IMPLEMENTATION

3.1. Tools and Software

In this study, 3DSMax program was used to produce the required 3D models. The virtual reality application was developed using the Unity3D game engine. C# language was used as the programming language. Samsung Gear VR was used for the VR imaging system. Oculus Integration was used to integrate the Unity3D game engine and virtual glasses.

3DSMAX, a three-dimensional visualization and animation program, was used to model the construction area in the screenshot shown in Figure 2. In order to virtualize a construction site, a tower crane consisting of three parts is modeled first. This model consists of the mast and jib that performs the right-left movement, the trolley that makes the forward-backward movement, and the hook that will move up and down. In addition, a cabin is modeled on the tower crane, where the operator's working environment will be located and which will form the main viewing angle. In addition, the loads and target buildings to be used in the load carrying process are modeled. Since the horizon line ends in the same way at every point of view, it will cause a decrease in reality and make it difficult to feel the movement, so the surrounding of the construction area is equipped with models such as buildings and roads.

Unity3D is a game engine developed by Unity3D Technologies, which is used to develop games and simulations that can be run on computers, game consoles and mobile devices. Unity3D, the first version of which was released in 2005, offers the opportunity to develop applications for 27 platforms. Unity3D is a very practical and easy to use game development engine [23]. It allows to use many easy-to-use plugins. The platform's material library has rich materials. It is easy to debug and optimize the system using Unity3D development. Unity3D game engine is a suitable platform for developing applications that will enable the user and objects to interact in a three-dimensional space. Libraries that will enable it to work in harmony with existing virtual glasses can be integrated into this software. In this study, integration between the Samsung Gear VR glasses and the development environment was achieved using Oculus Integration.

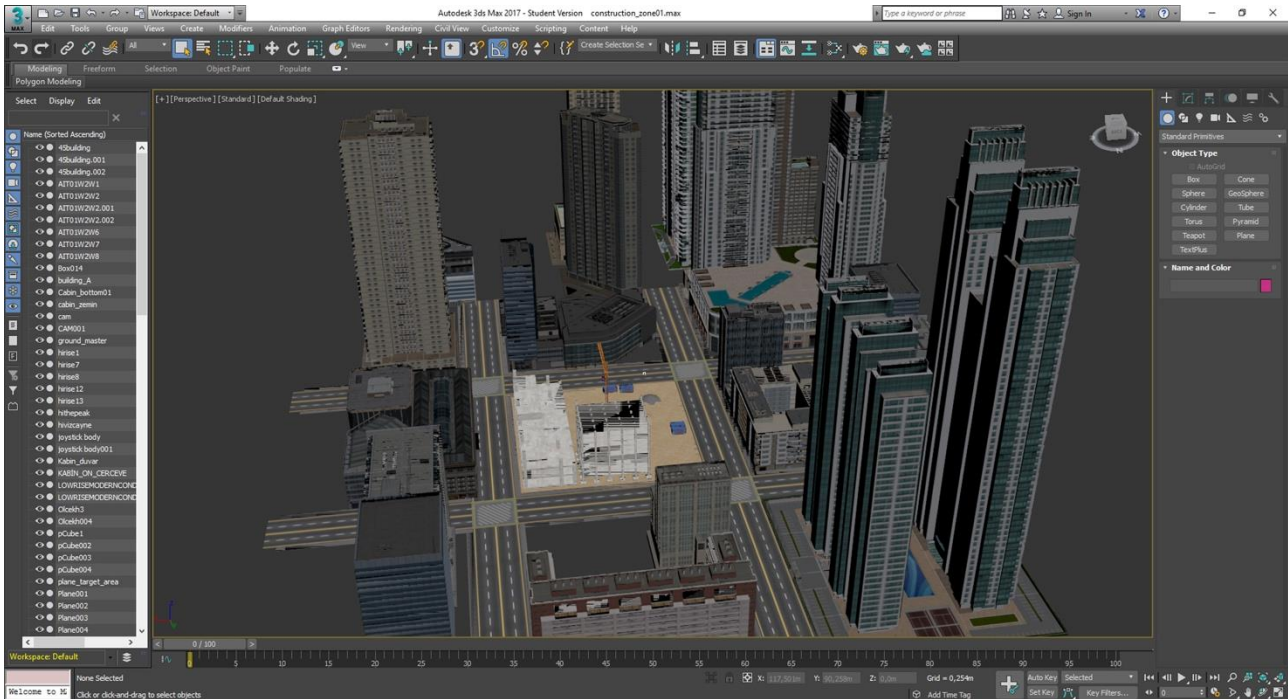


Figure 2. Modeling of the construction area

The control bar that comes with the Samsung Gear VR is an input tool that will allow the user to interact with the objects in the scene. The button and touchpad shown in Figure 3 are programmed according to need. The movements of the control tool in the real environment and the movement of the control stick created in the virtual environment work synchronously. When the control stick in the virtual environment is hovered over the objects on the stage that have a function assigned to them, selection can be made with the trigger button. In this way, it is ensured that the tower crane is operated and the left and right control sticks are activated. Depending on the selection, the movements of the tower crane are provided as a result of finger scrolling on the touchpad.

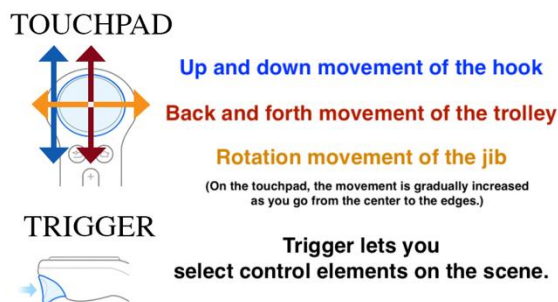


Figure 3. Control stick function

3.2. System Design

Before the virtual reality application, the requirements of the system were determined through interviews with experienced trainers who provide operator training. As a

result of these interviews, it was decided that the application would consist of three parts.

Scenario 1: The phase of operating the vehicle and moving the crane 3 axes, using the control hardware inside the tower crane. In this section, the trainee is expected to make a visual inspection of the work site with a tower crane. The task is completed successfully by observing the dangerous situations in the construction site and notifying the course trainer. This section is also evaluated as a section for learning the use of virtual reality headset by the trainee who is introduced to virtual reality headset for the first time.

Scenario 2: In this scenario, the trainee starts the crane by pressing the safety button. S/he tests each movement of the crane with no load. S/he moves the hook to the red areas appearing at different points on the application screen and successfully completes the scenario.

Scenario 3 Carrying the load: In this scenario, the trainee moves the hook closer to the center of gravity of the load that appears on the screen. S/he lifts the load connected to the hook without making a cross pulling motion. By pulling up on the hook, S/he allows the load to be raised to a safe height. S/he makes the back and forth movement of the trolley in a way that S/he can reach the target and not pose a danger. S/he enables the jib to approach the target by making a wheel movement and successfully completes the mission by safely leaving the load to the red area appearing at the target point. In all these stages, the touch of the load and the crane to the environmental elements in the working area causes the mission to fail.

In the interviews, the trainers drew attention to the fact that the rope should oscillate if the tower crane movement is

done quickly. They stated that many accidents were caused by the operator not taking this into account. Another remarkable situation is that the end of the wheel movement of the jib is not instantaneous due to the weight of the vehicle. It was stated that the operator should gradually reduce the movement before reaching the target. It has been stated that if these two situations are not taken into consideration, accidents such as hitting the environmental elements, knocking over, knocking down or falling of the load from the hook are experienced.

3.3. Scene Configuration with Game Engine

Figure 4 shows the hierarchy of game objects in Unity3D scene. Directional Light is used for outdoor lighting in the stage. In this way, it was ensured that the shadows fall on all objects in the same direction. The event system has been added to the stage to provide virtual reality headset control. The tower crane game object has also been added so that it can make movements within the application. The cabin and jib are placed under the crane game object so that it can move according to the movement of the hull. The trolley is formed under it so that it can move relative to the jib. In addition, OVRCameraRig Game object has been added to the cabinet in order to create the Vr image. In this way, it is possible to change the direction of the virtual camera according to the point of view of the virtual reality headset.

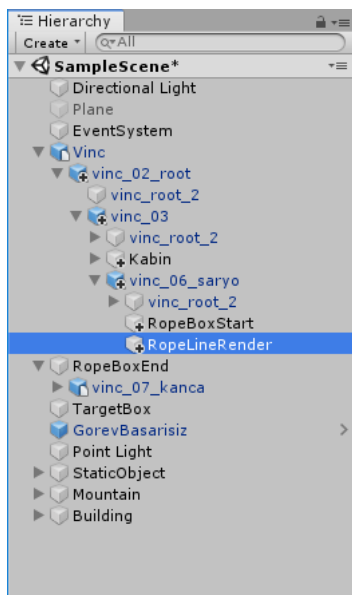


Figure 4. Hierarchy of objects created on stage

The operator's perspective displayed in Figure 5 is positioned in a place that can fully view the work area. In order to feel the crane movements more accurately, buildings as environmental elements, mountains as landforms and a skybox for the sky, was created.

Unity helps you simulate physics in your Project to ensure that the objects correctly accelerate and respond to collisions, gravity, and various other forces. Unity provides different physics engine implementations which

you can use according to your Project needs [24]. Physics module included in Unity3D was used to oscillate the Rope of the tower crane in up-down and right-left turns and to control collisions. Spring joint is used together with the physics module in order to make a swing of the hook and the load at the end of the rope in accordance with physics. The Spring Joint joins two Rigidbodies together but allows the distance between them to change as though they were connected by a spring [25]. The spring acts like a piece of elastic that tries to pull the two anchor points together to the same position. The strength of the pull is proportional to the current distance between the points with the force per unit of distance set by the Spring property. A Damping value can be given that reduces the spring force proportionally to the relative velocity between the two objects. The higher the value, the faster the oscillation decreases. By using this feature, the oscillation is changed according to the weight of the load.

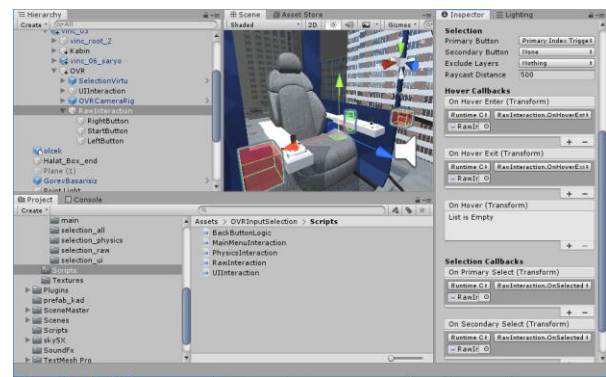


Figure 5. Creating models to interact

In order to control collisions depending on the movement of the tower crane in the virtual environment, the box collider game object working with the physics module is used. Warning messages are given to the user interface with the onTrigger method triggered by the collisions.

The tower crane movement is provided by applying the change in x and y values obtained from the touchpad to the objects in the virtual environment with the code block in Figure 6.

4. CONCLUSION AND RECOMMENDATIONS

Within the scope of this study, a virtual reality based tower crane operator training simulator has been developed. With this simulator, it is aimed that the trainees get to know and experience the construction equipment more closely before the practical trainings, thus reducing the rates of work accidents caused by operator mistakes.

Due to the structure of the tower crane, it is risky and difficult to carry out a training activity on it. The point that makes the developed training simulator important is that it can be ensured that the deficiencies that may occur in training can be gained in the classroom environment, as if on a tower crane. In addition, it has been observed that

these cranes are installed on a flat area in order to minimize safety risks. On the real construction site, it is likely that there will be obstacles such as buildings, electricity poles near the tower cranes. In operator training, trainees cannot experience this and many other situations in real environments. There is a chance to learn this in the virtual environment by making any number of mistakes.

```

//...
//...
//INPUT DATA START
bool hasChangedRope = false;
if (startButton)
{
    input = OVRInput.Get(OVRInput.Axis2D.PrimaryTouchpad);
    if (leftButton == true)
    { moveHook = input.y; moveTrolley = 0; rotateJib = 0; }
    if (rightButton == true)
    {
        moveHook = 0;
        moveTrolley = (input.y > 0.3f || input.y < -0.3f) ? input.y : 0;
        rotateJib = (input.x > 0.3f || input.x < -0.3f) ? input.x : 0;
    }
}
else
{ moveHook = 0; moveTrolley = 0; rotateJib = 0; }
//END

// HOOK MOVE START
if (moveHook > 0.5 && ropeLength < maxRopeLength)
{
    ropeLength += winchSpeed * Time.deltaTime;
    hasChangedRope = true;
}
if (moveHook < -0.5 && ropeLength > minRopeLength)
{
    ropeLength -= winchSpeed * Time.deltaTime;
    hasChangedRope = true;
}
//END

//JIB ROTATE
if (rotateJib > BOM_SPEED)
{
    BOM_SPEED += BOM_SPEED < rotateJib ? 0.005f : 0;
}
else if (rotateJib < -BOM_SPEED)
{
    BOM_SPEED -= BOM_SPEED > rotateJib ? 0.005f : 0;
}
currentEulerAngles = new Vector3(0, BOM_SPEED / 4, 0);
Bom.transform.localEulerAngles = currentEulerAngles;
//END

//TROLLEY MOVE
if (limit >= 750 || limit <= -750)
{ limit = 749; moveTrolley = 0; }
else
{ limit += moveTrolley; }
trolley.transform.Translate(Vector3.right * Time.deltaTime + moveTrolley * TROLLEY_SPEED);
//END
//...
//...

```

Figure 6. Code block provide movement of the tower crane

In future studies, based on the design drawn in this study, bigger, more powerful, and more realistic simulation software can be developed. Such software can be internet-based, record trainees' work in the application, detect successful and unsuccessful situations of trainees, and allow trainees to access remotely.

One of the limitations is that virtual glasses have effects on health (eye pain, dizziness, nausea, contagious disease etc.) that vary from person to person. It is important that the scenarios to be used with virtual glasses are made in a way that does not exceed a certain period of time. In addition, the course instructor should be careful in minimizing such effects during training.

As a result, Virtual Reality Environments are a very powerful tool that can be used in the training of many heavy machines. Virtual reality-based software to be used in operator training for highly dangerous heavy machines will be an effective method in reducing loss of life and property.

REFERENCES

- [1] I. Naskoudakis, K. Petrousatou, "A thematic review of main researches on construction equipment over the recent years", *Procedia Engineering*, 164, 206–213, 2016.
- [2] Internet: İş Makinaları Mühendisleri Birliği [Association of Construction Machinery Engineers], İMMB Meslek Kursu [Vocational Course], 2015, <http://www.ismakinalari.org.tr/tr/page.asp?id=69>, 01.09.2021.
- [3] G. Aksu, N. Doğan, "Scaling of teaching methods and techniques according to student opinions by pairwise comparison method", *Journal of Measurement and Evaluation in Education and Psychology*, 6(2), 194–206, 2015.
- [4] Z. Aydın, **The Effect of Active Learning Based Activities Used in Primary 6th Grade Mathematics Lesson on Students' Attitudes Towards Mathematics Lesson, Academic Achievement and Creative Thinking Levels**, Master's Thesis, Gaziantep University, Institute of Social Sciences, Gaziantep, 2011.
- [5] C. Eker, **The Effects of Video and Demonstration Teaching Methods on Student Success**, Master's Thesis, Sakarya University, Institute of Social Sciences, Sakarya, 2021.
- [6] W.C. Deese, L.L. Ramsey, J. Walczyk and D. Eddy, "Using demonstration assessments to improve learning", *Journal of Chemical Education*, 1511–1516, 2000.
- [7] N.E. Daluba, "Effect of demonstration method of teaching on students achievement in agricultural Science", *World Journal of Education*, 5–7, 2013.
- [8] O.U. Koğ, **The Effect of Teaching Mathematics with Visualization Approach on Cognitive and Affective Development of Students**, Doctoral Dissertation, Dokuz Eylül University, Institute of Educational Sciences, İzmir, 2012.
- [9] İ.E. Emre, M. Selçuk, V.Ö. Budak, M. Bütün, İ. Şimşek, "Immersion Investigation of Devices Used in Educational Virtual Reality Applications", *Journal of Information Technologies*, 12(2), 119–127, 2019.
- [10] G.C. Burdea, P. Coiffet, **Virtual reality technology**, New Jersey: John Wiley & Sons, New York, U.S.A., 2013.
- [11] M. Dursun, "An augmented reality based modular platform for solar energy", **ICCTA: International Conference on Computer and Technology Applications**, Antalya, Turkey, 149-153, 14 April 2020.
- [12] International Labour Organization, **Facts on safety at work**, Geneva: International Labour Office, (2005).
- [13] P. Hsu, P. Angeloudis, M. Aurisicchio, "Optimal logistics planning for modular construction using two-stage stochastic programming", *Automation in Construction*, 94, 47–61, 2018.
- [14] L. Wang, F. Hampel, J.A. Gladysz, "Gyroskop- Giganten: dipolare Cl-Rh-CO-Rotatoren. umgeben von Statoren aus drei Speichen 25-bis 27-gliedriger Makrocyclen", *Angewandte Chemie*, 118(26), 4479-4482, 2006.
- [15] M. Al-Hussein, M.A. Niaz, H. Yu, H. Kim, "Integrating 3D visualization and simulation for tower crane operations on construction sites", *Automation in Construction*, 15, 554-562, 2006.

- [16] K. Wu, B. G. Soto, F. Zhang, "Spatio-temporal planning for tower cranes in construction projects with simulated annealing", *Automation in Construction*, 111, 103060, 2020.
- [17] İ. Kökçü, **Tower Crane Design and Analysis**, Master's Thesis, Balıkesir University, Institute of Science and Technology, Balıkesir, 2015.
- [18] H. Urul, **Occupational Health and Safety Precautions to be Taken in Works with Cranes Used in Construction Works**, ÇSGB El Kitabı, İstanbul, 2013.
- [19] S.İ. Sayer, **Kule Vinçlerin Hesaplarının Ve Tasarımlarının Standartlar Açısından İncelenmesi**, Master's Thesis, Akdeniz University, Institute of Natural and Applied Sciences, Antalya, 2018.
- [20] Occupational Safe and Health Branch Labour Department, **Code of Practice for Safe Use of Tower Cranes**, 2011.
- [21] A. Nazhoğlu, **Detection of the Risks Encountered in the Works with Tower Cranes Used in the Construction Industry and Ways of Protection**, Occupational Health and Safety Specialization Thesis / Research, The Ministry of Labor and Social Security General Directorate of Occupational Health And Safety, Ankara, 2014.
- [22] Internet: Tower Crane Operator's Cab, http://www.towercrane.com/K10000_tower_crane_07_00.htm, 25.05.2021.
- [23] G. Liu, J. Wu, "Design and Implementation of Virtual Campus Roaming System Based on Unity3D", **2019 International Conference on Machine Learning**, Big Data and Business Intelligence, Taiyuan, China, 147-150, 8-10 Nov 2019.
- [24] Internet: Unity Documentation Physics, <https://docs.unity3d.com/Manual/PhysicsSection.html>, 01.12.2021.
- [25] Internet: Unity Documentation Spring Joint, <https://docs.unity3d.com/Manual/class-SpringJoint.html>, 01.12.2021.

Machine Learning Prediction Models of COVID-19 Prevalence in Turkey

Hatice Nur CANPOLAT, Ertan GÜNER

Industrial Engineering Department, Gazi University, Ankara, Turkey
haticenurcanpolat@gmail.com , erguner@gazi.edu.tr

Abstract— In December 2019, cases of pneumonia were detected in Wuhan, China. As a result of the examinations, it was reported that the disease was caused by a new coronavirus called SARS-COV-2. The disease called COVID-19 turned into an epidemic that affected the whole world in a short time. The COVID-19 pandemic caused a total of 197.905,518 cases and 4.218,403 deaths worldwide by August 1, 2021. The rapid spread of the COVID-19 epidemic has brought many serious problems with it. While countries were affected socially and economically, the biggest negative impact of the epidemic was on health systems. For this reason, it is vital to estimate the burden that will occur in a country's health system in such global epidemic situations. In this study, the prevalence of COVID-19 in Turkey was examined and the number of future cases and deaths were estimated with Linear Regression (LR), Random Forest Regression (RFR) and Bayesian Ridge Regression (BRR). The proposed estimation models were compared in terms of Coefficient of Determination (R^2), Mean Absolute Error (MAE), Mean Square Error (MSE) and Mean Squared Deviation (RMSE). The R^2 value was found to be 1 for RFR 0.97 for BRR and 0.97 for LR. The smallest MAE, MSE, RMSE values are respectively 955,17; 1527893,16; 1236,08 belongs to the RFR model. The results showed that the RFR prediction model is more effective than other methods in estimating the spread of the epidemic in Turkey.

Keywords— Turkey, COVID-19, prediction, linear regression, random forest regression, bayesian ridge regression

1. INTRODUCTION

Date of December 31, 2019, in Wuhan, China's Hubei community, cases of pneumonia of unknown cause were established, and on January 5, 2020, the World Health Organization created a new attack that had already been created [1]. Coronaviruses (CoV) are viruses that are in the group of RNA viruses to which SARS-CoV and MERS-CoV viruses belong and cause serious diseases by showing symptoms similar to the common cold in humans in general [2]. Untreatable pneumonia-like symptoms have been observed in people who have been in contact with this new coronavirus, called SARS-CoV-2. Initially expressed as 2019-nCoV, this disease was later named Covid-19 and affected the whole world in a short time like three months after it appeared in China [3]. COVID-19, which was declared a global epidemic by the World Health Organization on March 12, 2020, continues to threaten all humanity. While the number of cases caused by COVID-19 worldwide was 197 million 905 thousand 518 people until August 1, 2021, 4 million 218 thousand 403 people died. In Turkey, the total number of cases reached 5 million 567 thousand 142 from the first COVID-19 case on March 11, 2020, until August 1, 2021, and 51 thousand 332 people died due to COVID-19 [4].

When we examine the pattern of epidemics from the past to the present, we can observe that millions of people have died as a result of them. Some epidemic diseases that are thought to have disappeared in the world re-emerge in different regions by renewing themselves and continue to threaten human life [5]. The COVID-19 epidemic has seriously affected countries in terms of health infrastructure, economic system, military structure, and social life, including in developed countries. From this point of view, the critical point in combating epidemics is to prevent the spread of the epidemic by estimating the spread potential of the disease at the initial stage, by taking the necessary precautions [6].

2. LITERATURE REVIEW

As with many epidemics in the past, researchers are currently working on modeling the spread of the COVID-19 outbreak and estimating epidemiological curves [7]. Within the scope of these studies, warning systems that predict the future spread of the COVID-19 epidemic [8], developing various practices for the authorities to determine strategies to combat the situations that will occur as a result of the spread of the epidemic [9] topics such as the estimation of the impact of COVID-19 on different sectors such as education, transportation, economy were

covered. Some researchers prefer traditional modeling methods to predict the spread of COVID-19, while others prefer methods such as dynamic mathematical modeling and calculating the basic reproduction number [10]. Some studies in the literature show that artificial intelligence-based COVID-19 spread models attract great attention from the global scientific community. An electronic literature search was conducted in various databases to comprehensively examine the spread patterns of AI-based COVID-19. Based on the reviewed articles, it was concluded that artificial intelligence-based prediction algorithms have clear application in the transmission modeling of the COVID-19 epidemic and can be an important tool in the fight against pandemics. One of the most preferred estimation methods among these tools is regression analysis. When we look at the studies on the spread of COVID-19 in the literature, we see that the regression analysis gives successful results [11]. Some of the estimation studies with COVID-19 data are applications of regression methods only [12-14]. In the study, a regression-based hybrid machine learning model was proposed, a study was presented to estimate the size of COVID-19 cases by using data from different countries [15]. In the proposed approach to estimate the COVID-19 epidemic curve in Germany, data are stratified by age, and regression analysis is performed with a detailed dataset (Küchenhoff et al., 2020). Another part of the propagation models created by regression techniques is the comparison of prediction models [16, 17]. In addition to many international studies, there are COVID-19 spread prediction models applied in Turkey. We can generally group these studies as mathematical models and machine learning models of the epidemic. Analyzed by calculating the baseline reproduction number of the COVID-19 outbreak through a system of ordinary differential equations (ODEs) combined as a mathematical modeling exercise [18]. Widely preferred worldwide SIR [19] and SEIR [20] epidemiological epidemic models were applied to the COVID-19 data and the estimation results of the peak point of the epidemic in Turkey were evaluated. In another study, the spread of the COVID-19 pandemic in Turkey, the effects of the measures taken, and the number of tests performed on the number of cases were analyzed with a single Poisson regression technique [21]; the prediction success of the two methods was compared using Autoregressive Integrated Moving Average (ARIMA) and Artificial Neural Networks (ANN) for the COVID-19 spread in Turkey [22]. In this article, the dynamics of the spread of COVID-19 in Turkey are examined. By using the cumulative infected cases and death number data, prediction models were created with Linear Regression (LR), Random Forest Regression (RFR), Bayesian Ridge Regression (BRR) methods. This article consists of three parts; In the first part, brief information about COVID-19 and literature research on predicting the spread of COVID-19 are presented. In the second part, the materials and methods used in estimating the spread of the epidemic are introduced. Discussions and results are given in the last section.

3. MATERIAL AND METHOD

In this section, Linear Regression (LR), Random Forest Regression (RFR), Bayesian Ridge Regression (BRR) approach used to predict the spread of COVID-19 in Turkey are presented.

3.1. Data Source

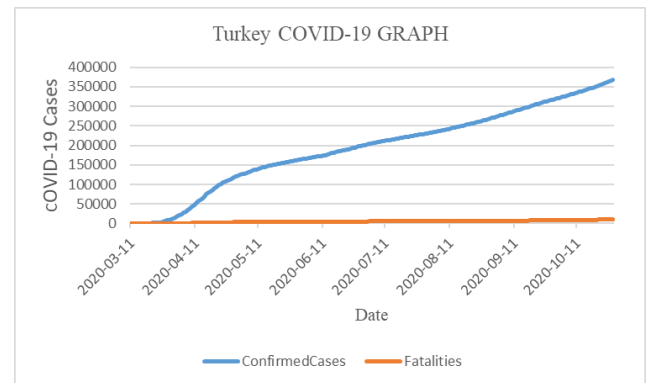


Figure 1. Turkey real data for cumulative confirmed cases and fatalities

Figure 1 shows the cumulative distribution of confirmed infected cases and fatal cases in Turkey for the period from March 11 to October 28, 2020. The data used in the study were obtained from the Johns Hopkins University database [23]. Due to the random spread of the disease, its future epidemiological progression is uncertain. In Figure 1, it is seen that the spread of COVID-19 in Turkey progresses in the form of exponential growth and that measures should be taken regarding this rapid increase.

3.2. Model Formulation

Machine learning algorithms include regression algorithms as a subset. Regression is a statistical analysis method that attempts to anticipate the relationship between a dependent variable and one or more independent variables by assuming a substantial correlation between the predicted variables and the findings [24]. In this section, Linear Regression (LR), Random Forest Regression (RFR), and Bayesian Ridge Regression (BRR) approaches are presented to estimate the spread of COVID-19 in Turkey.

3.3. Linear Regression

Linear Regression [25] finds the best fit across data features and labels. Mathematically, linear regression can be expressed as an equation of the line. A simple linear regression line equation is written as [26]:

$$y_i = \alpha_0 + \alpha_1 x_i + \epsilon_i, \quad i=1,2,\dots,n \quad (1)$$

Equation (1) can be expressed separately with the following equations:

$$y_1 = \alpha_0 + \alpha_1 x_1 + \epsilon_1, \quad i=1,2,\dots,n \quad (2)$$

$$y_2 = \alpha_0 + \alpha_1 x_1 + \epsilon_2, \quad i=1,2,\dots,n \quad (3)$$

⋮

$$y_n = \alpha_0 + \alpha_1 x_n + \epsilon_n, \quad i=1,2,\dots,n \quad (4)$$

The whole equation is added together to get the following equation:

$$\sum y_n = n\alpha_0 + \alpha_1 \sum x_i + \sum \epsilon_i \quad (5)$$

In linear regression, the sum of the residuals is zero $\sum \epsilon_i = 0$ when the following equation is obtained:

$$\sum y_n = n\alpha_0 + \alpha_1 \sum x_i \quad (6)$$

3.4. Random Forest Regression

The Random Forest approach proposed by Breiman is a machine learning algorithm with many decision trees. The random forest algorithm creates multiple decision trees and combines them to get a more accurate and stable prediction [27]. Random Forest algorithm, it contains parameters such as number of trees, number of attributes represented by B and M, respectively. Here the values of B and M are less than or equal to the dimensional values. T(i) represents the tree at index i. The tree(i) is created at each node by choosing a random value from the feature subset, taking into account only the splits in these features [28].

$$D = ((x_1, y_1), \dots, (x_n, y_n)) \quad (7)$$

D = observed data point,

B = Number of trees,

M = Number of features,

x_i d-dimensional vector, $B, M \leq d$, $T_i = \text{Tree}T_i$

This method has proven its success in both regression classification problems in recent years and is one of the best machine learning algorithms used in many different fields [29]. The dataset is first randomly divided into two portions in the random forest algorithm: training data for learning and validation data for testing the learning level. After dividing the dataset into two-thirds for training and one-third for validation, multiple decision trees are constructed at random from the dataset. Each tree's branching is determined by randomly chosen estimators from the nodes. The final estimate from the random forest approximation is the average of all results from each tree. Therefore, each tree affects the random forest estimate at certain weights [30].

3.4. Bayesian Ridge Regression

Bayesian Ridge Regression has all the features of Ridge regression and Bayesian linear regression as it is a special case of Bayesian linear regression and belongs to Ridge regression [31]. Ridge Regression reduces standard errors by adding one degree of deviation to its estimates. Ridge regression is a parameter estimation method used to address the common linearity problem that often arises in multiple linear regression [32].

If a sample set is defined as X, all samples of X are subtracted independently of a fixed but unknown probability density function p(y). Based on these examples recorded as P(y|X), it is necessary to estimate the probability distribution of y. The essence of Bayesian estimation is to get p(y|X) as close as the possible top(y).

Mathematically, the outcome variable y is assumed to be Gaussian distributed around Xw , X as follows to obtain a fully probabilistic model [33]:

$$P(y|X, w, \alpha) = N(y|Xw, \alpha) \quad (8)$$

One of the most useful types of Bayesian regression is Bayesian Ridge Regression, which estimates a probabilistic model of the regression problem. Here, the initial value of the coefficient w is given by the spherical Gaussian as follows:

$$p(w|\lambda) = N(w|0, \lambda^{-1}I) \quad (9)$$

This resulting model is called Bayesian Ridge Regression.

3.5. Regression Analysis

- Coefficient Determination

In this study, correlations between variables were tested using the Pearson correlation test. The correlation coefficient (R^2) of a regression model is calculated as follows [34]:

$$R^2 = \frac{\sum(\hat{Y} - \bar{Y})^2}{\sum(Y_i - \bar{Y})^2}$$

When $R^2=0$, there is no correlation between input and output variables. $R^2=1$, while there is a strong positive relationship between input and output variables; If $R^2=-1$, there is an inverse relationship between input and output variables.

- *Mean Absolute Error (MAE)*, is frequently used to measure accuracy in regression and time series models. The MAE is the average vertical distance between each true value and the line that best fits the data [35].
- *Mean Squared Error (MSE)*, tells you how close a regression curve is to a set of points. The MSE

- is always positive, and the model is said to perform well if the MSE is close to zero [36].
- *Root Mean Square Error (RMSE)*, a metric that measures the magnitude of the distance between the values predicted by a predictive model and their true values. The RMSE is the standard deviation of the estimation errors. A zero RMSE value means that the model does not make any errors. The advantage of RMSE is that it punishes big mistakes more. As a result, it is used more frequently than other procedures [35].

4. RESULT AND DISCUSSION

In this study, confirmed cumulative COVID-19 infected and death cases from the beginning of the epidemic in Turkey on March 11, 2020, to October 28, 2020, were modeled using three different regression methods. Estimates are provided to show the trend of COVID-19 spread and to help the government produce effective health policies. Linear Regression (LR), Random Forest Regression (RFR), and Bayesian Ridge Regression (BRR) analyzes were performed using the data mining libraries of the Python software package.

In this study, confirmed cumulative COVID-19 data between 11 March 2020 and 28 October 2020 were obtained from the Johns Hopkins University database. Only patients confirmed positive by laboratory tests are considered officially confirmed infected cases and validated data were used for the analyses. The data were divided into a training dataset (70%) to train the established model and a test dataset (30%) to validate the model.

In the figures below, the actual number of cases and deaths in Turkey between September 28, 2020 and October 28, 2020, and the values obtained as a result of the estimation models are included in the same graph comparatively.

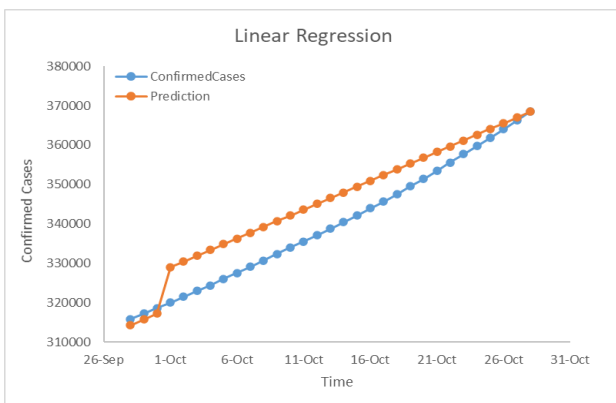


Figure 2. Turkey actual and prediction confirmed case data for LR

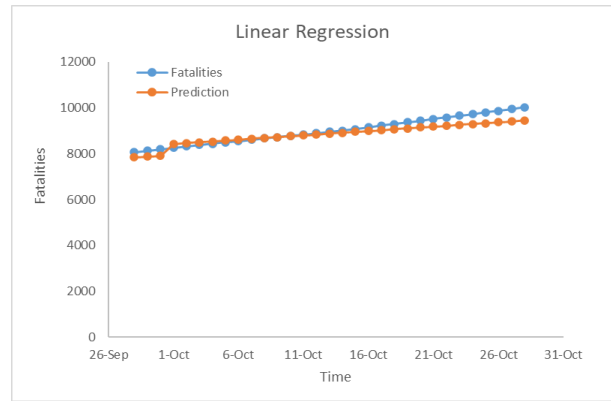


Figure 3. Turkey actual and prediction fatalities case data for LR

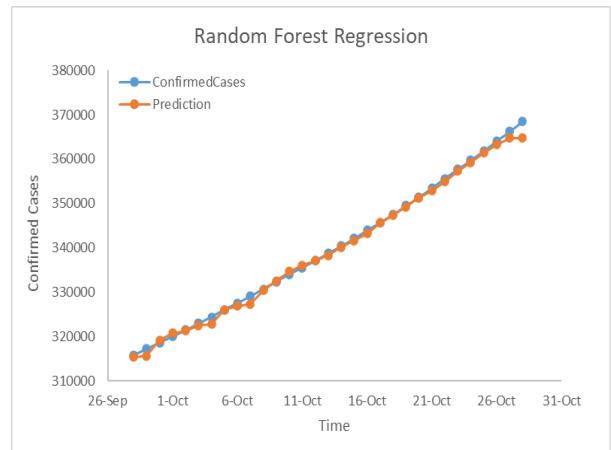


Figure 4. Turkey actual and prediction confirmed case data for RFR

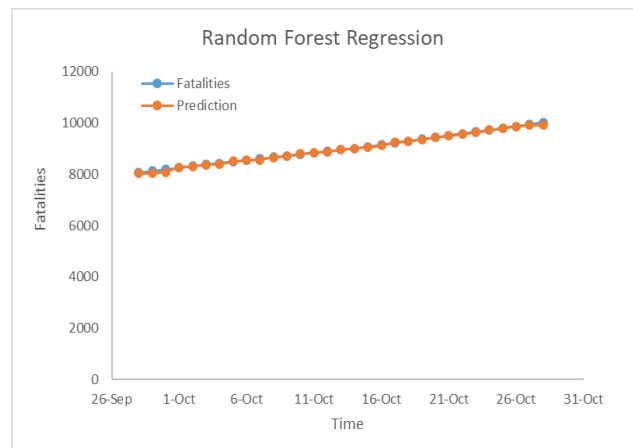


Figure 5. Turkey actual and prediction fatalities case data for RFR

In the study, the same procedures are applied for the cumulative number of infected cases and the estimation of the number of deaths. The Pearson correlation (R^2), RMSE, MSE, and MAE values obtained as a result of the regression analysis are shown separately in Table 1, Table 2, and Table 3. As can be seen from Table 2, the highest Pearson correlation value was obtained as a result of the RFR analysis. When the estimations of the number of infected cases and deaths are examined separately, it is

seen that the minimum error in the RMSE, MSE, and MAE values belongs to the RFR model.

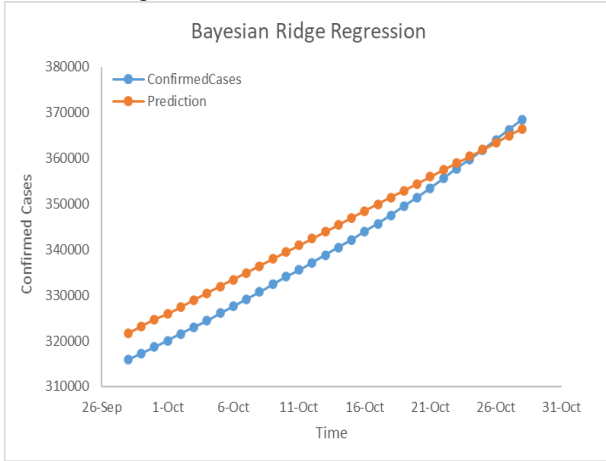


Figure 6. Turkey actual and prediction confirmed case data for BRR

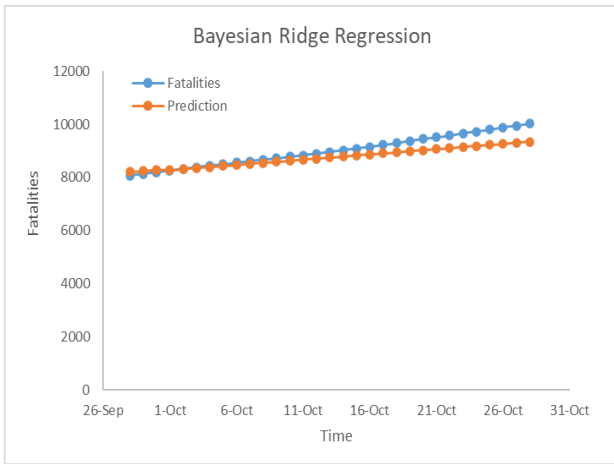


Figure 7. Turkey actual and prediction fatalities case data for BRR

Table 1. Linear regression results

LR	ConfirmedCase	Fatalities
R^2	0.97	0.94
$RMSE$	17700.77	608.11
MSE	313317161.6	369795.24
MAE	13491.74	521.99

Table 2. Random forest regression results

RFR	ConfirmedCase	Fatalities
R^2	1	1
$RMSE$	1236.08	33.49
MSE	1527893.16	1121.49
MAE	955.17	24.56

To determine the correlation between the cumulative number of infected cases and deaths in Turkey between March 11, 2020, and October 28, 2020, the correlation values between the date of the COVID-19 outbreak and the

number of cases are calculated. The correlation between the date and number of infected cases is 0.984939, the correlation between date and number of deaths is 0.970312. The correlation values between the date and the number of infected cases are quite close to 1, indicating that these variables have a significant statistical association.

Table 3. Bayesian ridge regression results

BRR	ConfirmedCase	Fatalities
R^2	0.97	0.94
$RMSE$	17830.11	620.01
MSE	317912784	384408.11
MAE	14164.44	526.46

Modeling the sudden increase-decrease in the number of cases during the COVID-19 epidemic and their effects on future cases is only possible with models made at certain periods. In this study, this variable process of the epidemic was analyzed with regression techniques and visible results were obtained. Table 1, Table 2, and Table 3 show how well-established regression models performed.

In a similar analysis, the expansion of the COVID-19 outbreak in Turkey was predicted [22]. Autoregressive Integrated Moving Average (ARIMA), and Artificial Neural Networks (ANN) techniques were applied to a data set with a sample size of 285 days. ConfirmedCases and Fatalities correlation coefficient 0.987, 0.996 for ARIMA; ConfirmedCases and Fatalities correlation coefficient 0.97, 0.99 values were obtained for ANN. In this study, with a smaller data set of 232 days, the RFR model Confirmed Cases and Fatalities correlation coefficient values are 1 for both. This result showed a strong and positive relationship between the actual values and the estimation results of the model. While modeling epidemics is very important, determining the appropriate forecasting method for the data, making early and accurate predictions are critical elements in minimizing the damage of the epidemic. The sample size has a direct relationship with the ANN's capacity to anticipate generalization. ANN usually converges at the local minimum with small sample sizes and performs poor predictive generalization [37]. However, the RFR algorithm is more powerful than other machine learning algorithms as it can generate random trees by taking random training data from subsets [38]. When the results in different areas where the RFR model is applied are examined, it is seen that it outperforms other machine learning models in terms of prediction accuracy [39-41]. The RFR model has minimal error values with a high correlation in estimating the cumulative number of infected cases and deaths. This study demonstrates that the RFR has a strong generalization ability with an acceptable degree of accuracy in predicting the spread of the COVID-19 outbreak. On the other hand, since the model applied in the study is tested with some of the training data, it is highly likely that the RFR will encounter an overfitting problem. For this reason, the fact that the training and test

data sets of the model are different from each other will allow more realistic estimation results to be obtained.

5. CONCLUSION

The COVID-19 epidemic, which adversely affected the whole world, has still been in effect since March 10, 2020, when the first case was seen. Using a variety of predictive modeling techniques to predict the spread of Covid-19 is important to effectively guide public health policymaking. The healthcare industry is an important potential area for predictive technical work to serve people in need of medical intervention. Especially during the epidemic period, thousands of people need to receive services from health systems at the same time. However, to optimize the allocation of the few resources available, it is vital to ensure that the estimations are accurate and precise, especially in resource-strapped societies. For this reason, it is vital to be prepared for interventions by planning the equipment and service capacities of health systems.

In this study, LR, RFR, BRR models estimate the cumulative number of infected cases and the cumulative number of deaths for Turkey between 11 March and 28 October 2020. Three regression techniques RMSE, MSE, MAE were compared with statistical metrics. The results revealed that the RF approach outperformed the others in terms of prediction accuracy. The correlation coefficient of the RFR model is too high to predict the spread of COVID-19. All findings point to the potential use of machine learning-based regression algorithms for epidemiological spread modeling and prediction in future pandemics.

If the machine learning models applied throughout the country in this study are applied separately for each region and each province data in future studies, it is thought that it will be effective analyzes in taking precautions against global epidemics such as COVID-19.

REFERENCES

- [1] X. Xu, C. Yu, L. Zhang, L. Luo, J. Liu, "Imaging features of 2019 novel coronavirus pneumonia", *Eur J Nucl Med Mol Imaging*, 47(5), 1022-1023, 2020.
- [2] R. Carrasco-Hernandez, R. Jácome, Y. López Vidal, S. Ponce de León, "Are RNA Viruses Candidate Agents for the Next Global Pandemic? A Review", *Ilar j*, 58(3), 343-358, 2017.
- [3] Internet: WHO, Naming the coronavirus disease (COVID-19) and the virus that causes it, [https://www.who.int/emergencies/diseases/novel-coronavirus-2019/technical-guidance/naming-the-coronavirus-disease-\(covid-2019\)-and-the-virus-that-causes-it](https://www.who.int/emergencies/diseases/novel-coronavirus-2019/technical-guidance/naming-the-coronavirus-disease-(covid-2019)-and-the-virus-that-causes-it), 30.07.2021.
- [4] Internet: J. Hopkins, Coronavirus Resource Center, <https://coronavirus.jhu.edu/map.html>, 11.12.2021.
- [5] M. V. Krishna, "Mathematical modelling on diffusion and control of COVID-19", *Infect Dis Model*, 5, 588-597, 2020.
- [6] Y. Ding, L. Gao, "An evaluation of COVID-19 in Italy: A data-driven modeling analysis", *Infect Dis Model*, 5, 495-501, 2020.
- [7] J. E. Gnanvi, K. V. Salako, G. B. Kotanmi, R. Glele Kakai, "On the reliability of predictions on Covid-19 dynamics: A systematic and critical review of modelling techniques", *Infect Dis Model*, 6, 258-272, 2021.
- [8] N. Chen et al., "Epidemiological and clinical characteristics of 99 cases of 2019 novel coronavirus pneumonia in Wuhan, China: a descriptive study", *Lancet*, 395(10223), 507-513, 2020.
- [9] D. J. Nokes, R. M. Anderson, "The use of mathematical models in the epidemiological study of infectious diseases and in the design of mass immunization programmes", *Epidemiol Infect*, 101(1), 1-20, 1988.
- [10] Y. Mohamadou, A. Halidou, P. T. Kapen, "A review of mathematical modeling, artificial intelligence and datasets used in the study, prediction and management of COVID-19", *Applied Intelligence*, 50(11), 3913-3925, 2020.
- [11] J. Musulin et al., "Application of Artificial Intelligence-Based Regression Methods in the Problem of COVID-19 Spread Prediction: A Systematic Review", *Int J Environ Res Public Health*, 18(8), 2021.
- [12] R. Sujath, J. M. Chatterjee, A. E. Hassanien, "A machine learning forecasting model for COVID-19 pandemic in India", *Stoch Environ Res Risk Assess*, 1-14, 2020.
- [13] R. O. Ogundokun, A. F. Lukman, G. B. M. Kibria, J. B. Awotunde, and B. B. Aladeitan, "Predictive modelling of COVID-19 confirmed cases in Nigeria", *Infect Dis Model*, 5, 543-548, 2020.
- [14] S. T. A. Shah, A. Iftikhar, M. I. Khan, M. Mansoor, A. F. Mirza, M. Bilal, "Predicting COVID-19 Infectious Prevalence Using Linear Regression Tool", *Journal of Experimental Biology and Agricultural Sciences*, 2020.
- [15] M. Saqib, "Forecasting COVID-19 outbreak progression using hybrid polynomial-Bayesian ridge regression model", *Applied Intelligence*, 51(5), 2703-2713, 2020.
- [16] L. A. Amar, A. A. Taha, M. Y. Mohamed, "Prediction of the final size for COVID-19 epidemic using machine learning: A case study of Egypt", *Infect Dis Model*, 5, 622-634, 2020.
- [17] S. Sengupta, S. Mugde, G. Sharma, "Covid-19 Pandemic Data Analysis and Forecasting using Machine Learning Algorithms", *medRxiv preprint*, 2020.
- [18] I. H. Aslan, M. Demir, M. M. Wise, S. Lenhart, "Modeling COVID-19: Forecasting and analyzing the dynamics of the outbreak in Hubei and Turkey", *medRxiv preprint*, 2020.
- [19] M. Ozdinc, I. Senel, S. Ozturkcan, A. Akgul, "Predicting the Progress of COVID-19: The Case for Turkey", *Turkiye Klinikleri Journal of Medical Sciences*, 2020.
- [20] Ş. Arslan, M. Y. Özdemir, A. Uçar, "Nowcasting and Forecasting the Spread of COVID-19 and Healthcare Demand in Turkey, A Modelling Study", *medRxiv*, 2020.
- [21] B. Taşdelen, D. Derici Yıldırım, "Predicting COVID-19 Cases in Turkey with Poisson Regression and the Effect of Preventions on Incidence Rate Ratio Estimation", *Turkiye Klinikleri Journal of Biostatistics*, 12(3), 293-302, 2020.
- [22] G. Toga, B. Atalay, M. D. Toksari, "COVID-19 prevalence forecasting using Autoregressive Integrated Moving Average (ARIMA) and Artificial Neural Networks (ANN): Case of Turkey", *J Infect Public Health*, 14(7), 811-816, 2021.

- [23] Internet: J. Hopkins. "COVID-19 Datasets from Johns Hopkins University", <https://github.com/CSSEGISandData/COVID-19>, 09.08.2021.
- [24] H. Singh, J. Dhar, **Mathematical Population Dynamics and Epidemiology in Temporal and Spatio-Temporal Domains**, CRC Press, 302, 2018.
- [25] X. Yan, X. Su, "Linear regression analysis: theory and computing", *World Scientific*, 2009.
- [26] A. K. Gupta, V. Singh, P. Mathur, C. M. Travieso-Gonzalez, "Prediction of COVID-19 pandemic measuring criteria using support vector machine, prophet and linear regression models in Indian scenario", *Journal of Interdisciplinary Mathematics*, 24(1), 89-108, 2020.
- [27] L. Breiman, "Random forests", *Machine learning*, vol. 45(1), 5-32, 2001.
- [28] A. Khakharia et al., "Outbreak prediction of COVID-19 for dense and populated countries using machine learning", *Annals of Data Science*, 8(1), 1-19, 2021.
- [29] C. M. Yeşilkanat, "Spatio-temporal estimation of the daily cases of COVID-19 in worldwide using random forest machine learning algorithm", *Chaos Solitons Fractals*, 140, 110210, 2020.
- [30] A. M. Prasad, L. R. Iverson, A. Liaw, "Newer classification and regression tree techniques: bagging and random forests for ecological prediction", *Ecosystems*, 9(2), 181-199, 2006.
- [31] M. Byrtek, F. O'Sullivan, M. Muzi, A. M. Spence, "An adaptation of ridge regression for improved estimation of kinetic model parameters from PET studies", *IEEE transactions on nuclear science*, 52(1), 63-68, 2005.
- [32] G. C. McDonald, "Ridge regression", *Wiley Interdisciplinary Reviews: Computational Statistics*, 1(1), 93-100, 2009.
- [33] Y. Yang, Y. Yang, "Hybrid prediction method for wind speed combining ensemble empirical mode decomposition and bayesian ridge regression", *IEEE Access*, 8, 71206-71218, 2020.
- [34] Q. Shi, M. Abdel-Aty, and J. Lee, "A Bayesian ridge regression analysis of congestion's impact on urban expressway safety", *Accid Anal Prev*, 88, 124-37, 2016.
- [35] T. Chai, R. R. Draxler, "Root mean square error (RMSE) or mean absolute error (MAE)? – Arguments against avoiding RMSE in the literature", *Geoscientific model development*, 7(3), 1247-1250, 2014.
- [36] O. Köksoy, "Multiresponse robust design: Mean square error (MSE) criterion", *Applied Mathematics and Computation*, 175(2), 1716-1729, 2006.
- [37] R. Mao, H. Zhu, L. Zhang, A. Chen, "A new method to assist small data set neural network learning", **in Sixth international conference on intelligent systems design and applications**, Jian, China, 1, 17-22, October 2006.
- [38] P. Panov, S. Džeroski, "Combining bagging and random subspaces to create better ensembles", **in International Symposium on Intelligent Data Analysis**, Ljubljana, Slovenia, 4723, 118-129, 2007.
- [39] K. B. Prakash, S. S. Imambi, M. Ismail, T. P. Kumar, Y. Pawan, "Analysis, prediction and evaluation of covid-19 datasets using machine learning algorithms", *International Journal*, 8(5), 2020.
- [40] D. S. Palmer, N. M. O'Boyle, R. C. Glen, J. B. Mitchell, "Random forest models to predict aqueous solubility", *Journal of chemical information and modeling*, 47(1), 150-158, 2007.
- [41] P. Zahedi, S. Parvande, A. Asgharpour, B. S. McLaury, S. A. Shirazi, B. A. McKinney, "Random forest regression prediction of solid particle Erosion in elbows", *Powder Technology*, 338, 983-992, 2018.

Impact of Feature Configurations on Speech Recognition Using Listen, Attend and Spell Network

Zekeriya TÜFEKÇİ¹, Gökay DİŞKEN²

¹Department of Computer Engineering, Çukurova University, Adana, Turkey

²Department of Electrical-Electronics Engineering, Adana Alparslan Türkeş Science and Technology University, Adana, Turkey
ztufekci@cu.edu.tr, gdisken@atu.edu.tr

Abstract— Conventional speech recognition systems heavily relied on Hidden Markov Models (HMM) as they model time-varying feature vector sequences effectively. State distributions were computed using Gaussian Mixture Models (GMM). More accurate models have been achieved with the aid of deep learning – HMM hybrid approach, where GMMs were replaced by neural network models. Recently, the paradigm shifted to end-to-end deep learning models as they have produced state-of-the-art results in many different speech-related applications. Listen, Attend and Spell (LAS) network is one of those models which can produce character sequences at its output. The LAS consists of two parts; the encoder network which receives acoustic features as inputs, and the attention-based recurrent decoder network which produces one character at a time step. In this work, speech recognition performance of the LAS network is examined using different configurations of acoustic features extracted from the TIMIT database. In general, context-windowing (stacking a few consecutive frames as the input of the network) is preferred in speech recognition systems. However, the experimental results indicated that a better performance can be obtained without the context-windowing. Furthermore, increasing the acoustic feature dimension did not make a high impact on the performance.

Keywords— deep learning, recurrent neural networks, speech recognition, timit

1. INTRODUCTION

Speech recognition systems include several components such as an acoustic model, a language model, a pronunciation model, etc. [1]. End-to-end systems, where different components trained jointly, gained popularity recently. Prior to these systems, Hidden Markov Model (HMM) method was the dominating approach for speech recognition [2]. The conventional HMM systems include Gaussian mixture models (GMMs) to fit each state of HMM to the speech feature frames. Deep neural networks (DNNs) were replaced with the GMMs, and achieved a higher performance on several speech recognition benchmarks [3]. Another important advancement for speech recognition is Connectionist Temporal Classification (CTC), which transforms the network outputs into a conditional probability distribution over label sequences [4]. Combined with a recurrent neural network (RNN), end-to-end training system was designed and tested in phoneme recognition experiments [5].

A more recent neural network architecture, namely Listen, Attend and Spell (LAS) [6], has been introduced for large vocabulary conversational speech recognition. One of the most important aspects of this network is that it can emit one character at a time at the output, without the need for

an explicit language model, a pronunciation model, state model such as HMM, etc. Contrary to the CTC, the LAS network does not make any independence assumption on the output probabilities. The LAS is actually a sequence-to-sequence network [7-9], which includes an encoder RNN, and a decoder RNN.

In this work, a network similar to the LAS is used for speech recognition on TIMIT database, which is a popular database that have been used in a variety of speech-related studies besides speech recognition [10-12]. The inputs of the network, i.e. filterbank energies, are investigated in terms of feature dimensions and context-windowing (frame stacking). Context-windowing is a term that describes a stack of adjacent frames. Considering the frame of interest at a given time step, it can be concatenated with a few previous time steps, or it can be concatenated with both previous and next time steps [13-15]. For the input feature size, a high number of dimensions is preferred in general, considering the static features with/without temporal derivatives (e.g., 123 in [5], 40 in [6], 80 (240 with frame stacking) in [14]). The effects of the frame stacking and feature size are investigated experimentally in this study. For the feature size, several values are considered from 20 up to 80. 3 consecutive frames are stacked for context-windowing, and the results are

compared with the single-frame version of the same feature size. The experimental results showed that neither frame stacking nor increasing input size is effective to increase the recognition performance.

2. LAS NETWORK

The network used in this work is a modified version of [6]. The main difference between the models is the types of the layers. In the original LAS network, pyramidal long short-term memory (LSTM) layers were used in the encoder. In this work, gated recurrent unit (GRU) [9] layers were preferred as they have one less gate than LSTM, and deliver a higher performance for small databases [16]. Similarly, GRU layers were also used in the decoder, and a fully-connected network with two layers was used as the attention layer [13]. The LAS network is summarized as a block diagram in Figure 1. The green rectangles represent the encoder part, which takes acoustic features as inputs. The blue rectangles represent the decoder part, where the inputs are phoneme labels (ground-truth for training, and predictions for testing). The fully-connected layer serves as an embedding layer on the input phonemes.

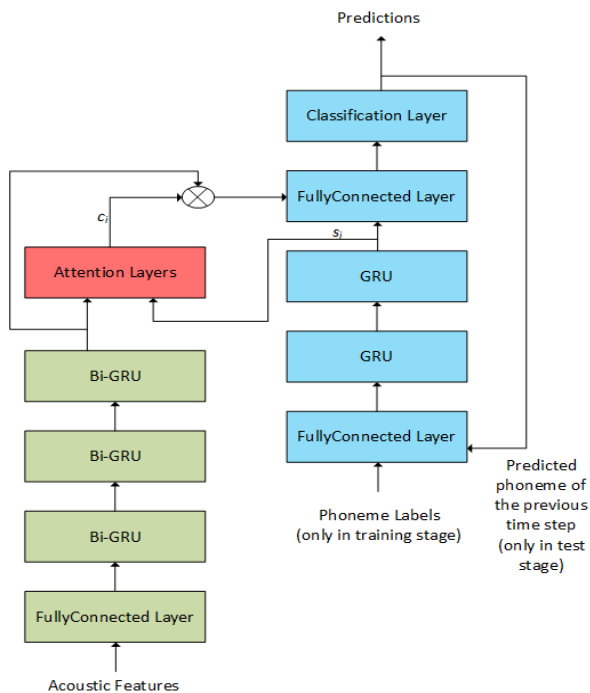


Figure 1. Block diagram of LAS network

The GRU layers of the encoder are bidirectional, hence both the input sequence and its reversed form are learned. The output of the decoder at time step i depends on current decoder state (s_i) and the context vector (c_i , obtained via attention layer). The attention layer consists of three fully-connected layers. One of the layers gets its inputs from the encoder, and another layer gets its inputs from the final GRU layer of the decoder (These inputs serve as source states and target states). Then, the layer outputs are added together, and ReLU activation function is applied. Note that the size of the inputs from the encoder layer will be

doubled due to the bidirectional operation (assuming all GRUs have the same hidden size). Hence, their sizes are matched at the layer outputs. A final layer is applied to the resulted vector, and the attention vector is obtained at its output. The main purpose of the attention layer is to focus the decoder on only a few frames of the encoder output [6]. The output of the attention layer is therefore can be thought of a mask that weights features based on their relation with the target output. So, encoder output states are multiplied with the attention mask.

The masked output and the last GRU layer of the decoder are concatenated to construct the input of the fully-connected layer before the classification layer. The classification is therefore can focus on the parts highlighted by the attention layer. Finally, softmax function is used in the classification layer which produces one phoneme at each time step until a special token, end-of-sentence, is produced.

As the acoustic models of the speech, log-mel filterbank features are used, since they perform better than the traditional cepstral features for neural network-based system. The reason for that may be the extraction of more useful information with the neural networks, which could not be possible with the low dimensional cepstral features [17].

3. EXPERIMENTS

The modified LAS network's performance for phoneme recognition was examined on TIMIT database, where 3696 sentences from 462 speakers were used for training (SA records were removed), 400 sentences were chosen as the validation data, and the test set contains 192 sentences from 24 speakers. There is no overlap between speakers of each partition. Transcriptions of TIMIT are based on 61 phonemes. However, typical approach is to map those phonemes into 48 for modelling. Further, confusions among some of the phones are not considered as error, so only 39 phoneme categories are used in the evaluation [18]. The same approach is followed in this work.

All of the GRU layers have 256 hidden states. The networks were trained for 200 epochs, using a batch size of 16, and ADAM optimizer. Learning rate was $3e-4$ initially, and reduced by half if there is no improvement after ten epochs. Also, dropout was used after the GRU layers with a 0.5 probability. Pytorch was used to build, test, and train the network.

For feature extraction, filterbank features were extracted using torchaudio, which provides Kaldi's feature extraction capabilities within Python language. Frame length was 25 ms, and frame shift was 10 ms. The number of triangular filters varied from 20 to 80. In case of frame stacking, 3 consecutive frames were concatenated, then the frame rate is down-sampled by 3. Hence, the temporal length is reduced by 3-fold, but none of the frame information was lost.

Table 1 presents the results in terms of phoneme error rate (PER). The best performance (19.17% PER) was achieved when 30-dimensional features were used without frame stacking. On average, individual processing of the frames led to a 3.92% relative PER reduction compared to the stacked frames. Also, single frame processing consistently outperformed stacked frames for each feature size. A possible reason for this result may be the fact that temporal resolution was decreased by stacking frames. More precisely, the network operates at every 30 ms for the stacked frames (as a single frame was produced every 10 ms).

Table 1. Phoneme error rates (PER,%) for different feature configurations

Feature Dimension	PER	
	Single Frame	Stacked Frames
20	20.12	20.73
25	20.15	21.03
30	19.17	21.52
35	20.54	21.12
40	20.85	21.01
60	20.67	21.40
80	21.05	21.55
Average	20.36	21.19

Another important information derived from the results is that more dimension did not guarantee more accuracy, whether the frames were stacked or not. Nevertheless, as the dimension increases, the PER slightly increases. As more filters were used to obtain high dimensional features, they can focus on much narrow bands, produce unnecessary details of the speech spectrum, hence sending some sort of unstable features to the network.

This situation is visually verified by comparing features obtained from the same frame. Figure 2 shows the log-Mel features of a voiced frame. The upper plot has a dimension of 30, and the resulted graph is smoother than the plot below, which has a dimension of 80. Hence, increasing the feature size produce abruptly changing features, which may not include discriminative information. The network will try to learn those small variations too, and if the training data is not sufficient, the resulted model will be poor. The absolute difference between PER values may not seem significant, but increasing the feature dimension will require more arithmetic operations in both training and test stages. Also, more memory will be needed to store the high dimensional features.

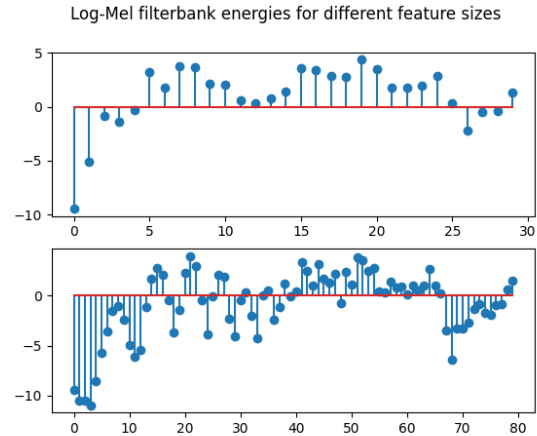


Figure 2. Comparison of 30-, and 80-dimensional features obtained from the same frame.

4. CONCLUSION

In this work, the effects of feature dimensions and frame stacking on phoneme recognition performance of the LAS network were examined using the TIMIT database. The results indicated that features with higher dimensions performed slightly worse than low dimensional ones. Visual inspection of the different dimensional features verified that more dimensions increase the fluctuation of the features. Less dimensions produce more stable features, and deliver slightly better performances. An optimum feature size can be found empirically. Similarly, single frame processing consistently performed better than the frame stacking approach. Nevertheless, frame stacking brings more computational and storage loads. Therefore, for relatively small databases, the LAS network can exploit low dimensional features without the frame stacking.

REFERENCES

- [1] C. Kim, D. Gowda, D. Lee, J. Kim, A. Kumar, S. Kim, A. Garg, C. Han, "A Review of On-Device Fully Neural End-to-End Automatic Speech Recognition Algorithms," **2020 54th Asilomar Conference on Signals, Systems, and Computers**, Pacific Grove, USA, 277–283, 1-4 November, 2020.
- [2] A. P. Varga and R. K. Moore, "Hidden Markov model decomposition of speech and noise," **International Conference on Acoustics, Speech, and Signal Processing**, Albuquerque, USA, 845–848, 3-6 April, 1990
- [3] G. Hinton, L. Deng, D. Yu, G. E. Dahl, A. Mohamed, N. Jaitly, A. Senior, V. Vanhoucke, P. Nguyen, T. N. Sainath, B. Kingsbury, "Deep Neural Networks for Acoustic Modeling in Speech Recognition: The Shared Views of Four Research Groups," *IEEE Signal Process. Mag.*, 29(6), 82–97, 2012.
- [4] A. Graves, S. Fernández, F. Gomez, and J. Schmidhuber, "Connectionist temporal classification," **Proceedings of the 23rd international conference on Machine learning - ICML '06**, Pittsburgh, USA, 369–376, 25-29 June, 2006.

- [5] A. Graves, A. Mohamed, and G. Hinton, "Speech recognition with deep recurrent neural networks," **2013 IEEE International Conference on Acoustics, Speech and Signal Processing**, Vancouver, Canada, 6645–6649, 26-31 May, 2013.
- [6] W. Chan, N. Jaitly, Q. Le, and O. Vinyals, "Listen, attend and spell: A neural network for large vocabulary conversational speech recognition," **2016 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)**, Shanghai, China, 4960–4964, 20-25 March, 2016.
- [7] I. Sutskever, O. Vinyals, and Q. V. Le, "Sequence to sequence learning with neural networks," **27th International Neural Information Processing Systems**, Montreal, Canada, 3104–3112, 8-13 December, 2014.
- [8] D. Bahdanau, K. Cho, and Y. Bengio, "Neural Machine Translation by Jointly Learning to Align and Translate," **3rd International Conference on Learning Representations**, San Diego, USA, 1–15, 7-9 May, 2015.
- [9] K. Cho *et al.*, "Learning Phrase Representations using RNN Encoder–Decoder for Statistical Machine Translation," **Proceedings of the 2014 Conference on Empirical Methods in Natural Language Processing (EMNLP)**, Doha, Qatar, 1724–1734, October, 2014.
- [10] M. T. S. Al-Kaltakchi, W. L. Woo, S. S. Dlay, and J. A. Chambers, "Comparison of I-vector and GMM-UBM approaches to speaker identification with TIMIT and NIST 2008 databases in challenging environments," **2017 25th European Signal Processing Conference (EUSIPCO)**, Kos, Greece, 533–537, 28 Aug. - 2 Sept., 2017.
- [11] T. Drugman, Y. Stylianou, Y. Kida, and M. Akamine, "Voice Activity Detection: Merging Source and Filter-based Information," *IEEE Signal Process. Lett.*, 23(2), 252–256, 2016.
- [12] I. Cohen and B. Berdugo, "Speech enhancement for non-stationary noise environments," *Signal Processing*, 81(11), 2403–2418, 2001.
- [13] T. Luong, H. Pham, and C. D. Manning, "Effective Approaches to Attention-based Neural Machine Translation," **Proceedings of the 2015 Conference on Empirical Methods in Natural Language Processing**, Lisbon, Portugal, 1412–1421, 17-21 September, 2015.
- [14] H. Sak, A. Senior, K. Rao, and F. Beaufays, "Fast and accurate recurrent neural network acoustic models for speech recognition," **Interspeech 2015**, Dresden, Germany, 1468–1472, 6-10 September, 2015.
- [15] J. Yang, L. Liu, and Q. He, "Discriminative feature based on FWMW for playback speech detection," *Electronic Letters*, 55(15), 861–864, 2019.
- [16] S. Yang, X. Yu, and Y. Zhou, "LSTM and GRU Neural Network Performance Comparison Study: Taking Yelp Review Dataset as an Example," **2020 International Workshop on Electronic Communication and Artificial Intelligence (IWECAI)**, Shanghai, China, 98–101, 12-14 June, 2020.
- [17] F. Seide, G. Li, X. Chen, and D. Yu, "Feature engineering in Context-Dependent Deep Neural Networks for conversational speech transcription," **2011 IEEE Workshop on Automatic Speech Recognition & Understanding**, Waikoloa, USA, 24–29, 11-15 December, 2011.
- [18] K.-F. Lee and H.-W. Hon, "Speaker-independent phone recognition using hidden Markov models," *IEEE Trans. Acoust.*, 37(11), 1641–1648, 1989.

Performance of YOLO on a Custom Dataset of Infrared Laser Images

Emre Ege EŞER¹, Firat HARDALAC²

¹ Aselsan A.S. Ankara, Turkey

² Department of Electrical and Electronics Engineering, Gazi University, Ankara, Turkey
eeser@aselsan.com.tr, fiat@gazi.edu.tr

Abstract— YOLO (You Only Look Once) is a state-of-the-art, real-time object detection algorithm, first developed by Redmon, Divvala, Girshick and Farhadi. Which was later worked upon and developed further by others. In this paper, we compare the different versions of YOLO algorithm, on a completely different dataset that only contains one class of objects. In total, 3 algorithms will be compared on a dataset containing 74 infrared laser images taken from an infrared camera. While the mAP results were fairly close, FPS value of YOLO v4 tiny method was superior to others. These experiments will be followed by automated laser location finder in a future work. The 3 algorithms that will be used are YOLOv3, YOLOv4 and YOLOv4 tiny. Infrared laser sources are used in multiple areas such as thermal imaging, targeting, and tracking in military applications, temperature measurement, astronomy and so on. We also use a GPU with middle class performance, along with a CPU, which will give the information on performance of YOLO with lower end computers.

Keywords— YOLO, object detection, laser images, deep learning, laser guidance, infrared lasers.

1. INTRODUCTION

Object detection has been a popular subject of multiple work done since the development of deep neural networks. Even though there are multiple attempts at improving the performance of these neural networks, attempts at a higher speed network has not been as successful. That is, until the YOLO algorithm. Fast-RCNN [1] and faster-RCNN [2] algorithms were a definite improvement that cannot be ignored, although there is no argument that the YOLO algorithm is the superior model, speed-wise.

Older architectures use classifiers to detect objects in an image. This results in a slower model, resulting in object detection processes taking an unreasonable time for real-time object detection. Most of the newer algorithms, such as RCNN and faster versions of RCNN, work using region proposals. These region proposals help the algorithm to be a bit more accurate while decreasing the speed and making these models more resource-intensive. Even being faster than CNN and RCNN, these algorithms are not suitable for real-time detection at higher FPS. With YOLO, there is no need to use region proposals. The image is divided into sectors, which are then rated for possible object locations, called the “class probability map”. After which, bounding boxes of objects are finalized. Because of this simple algorithm,

YOLO runs more efficiently compared to other algorithms and allows real-time object detection. Base YOLO model can process up to 45 frames per second on a TITAN-X GPU, while YOLOv4 tiny model can reach processing speed of up to 1774 FPS [3].

In systems containing more than one camera or containing both a camera and a laser source, a process called “boresight” must be done. This is commonly referred to calibrating a rifle so that the sights match with where the bullet is going to land. Boresight in laser systems is not much different. The laser must be calibrated in such a way, that the spot of the laser source lands exactly in the middle of the camera, or the crosshair. This can and usually is done by a person, which opens a whole new window for errors. By using deep learning and neural networks, error margins can be lowered to almost non-existent. Our future work will be related to real-time determination of laser spot location and because of that, there must be a comparison as to which algorithm will be used. For this work, speed is needed as much as low amounts of error. Since YOLO seems like the fastest algorithm while keeping up with the other methods at error-performance-wise, only YOLO algorithms will be compared.

The wavelength of infrared (IR) waves starts from 700 nm, just above the visible light spectrum, and end with

1 mm, before microwaves [4]. They are commonly used in everyday applications as heating, cooling, communications etc. IR waves can also be used to measure heat, used almost daily for temperature measurements of people. Other uses of IR waves include night vision [5], thermal imaging [6], and thermal guidance of laser-guided ammunition, in military applications [7]. Because IR radiation lies beyond the visible spectrum of light, a human eye cannot see IR lasers without special equipment. The most common way to detect laser spot is to use an IR camera. These cameras or sensors work in the same spectrum of light, allowing us to see where our laser source falls on the target. But possible noisy backgrounds of some scenarios might make it impossible to detect laser spot on the image to a human. This can be avoided with the use of computers. Better yet, we can even use artificial intelligence to create an ever-learning network for better detection. This detection method can be used during military missions to minimize possible laser boresight errors, or to aid in personnel during infrared laser production. It can be considered to use a simple solution for this problem such as detecting the brightest spot on an image, instead of using a whole artificial intelligence algorithm, though this simple method falls short with the increase of background noise, in which the brightest spot might not be, and usually is not, the laser itself.

2. YOLO ALGORITHMS AND NETWORK PERFORMANCE

2.1. Origin of YOLO

YOLO algorithm was first developed as a model aimed at acquiring higher speed and real time detection. Image is divided into an $S \times S$ grid, and each grid cell predicts bounding boxes and confidence for these bounding boxes. Final prediction is outputted as a $7 \times 7 \times 30$ tensor. The model is based on convolutional neural networks and evaluated on PASCAL VOC detection data set. The network has 24 convolutional layers with 2 fully connected layers. Final layer activation function chosen to be a linear function. Also, sum-squared error is used, allowing the network to weight errors equally across all bounding boxes [3]. It was compared with other networks at the time. Looking at PASCAL VOC 2012 leaderboard, we see that YOLO performed successfully even though it is the only real-time detector on the table. There was another attempt with combining YOLO and Fast-RCNN. Since YOLO performs a lot better in background errors, they were able to eliminate background errors caused by Fast-RCNN and got a 71.8% mAP score.

These specifications of this algorithm make it one of the fastest object detection methods to date. But to provide this speed benefit, there must be some limitations. YOLO's worst drawback might be the localization errors, compared to other methods. Even then, this drawback is justified with a better performance on

background errors, and surely with the faster detection speed.

2.2. Scoring of Performance

There are multiple parameters to use while measuring the performance of a deep learning algorithm. Precision shows what percentage of our predictions are correct. Recall presents how good the positives are found. Equations are shown as follows:

$$Precision = \frac{True\ Positive}{Total\ Positives} \quad (1)$$

$$Recall = \frac{True\ Positive}{All\ Cases} \quad (2)$$

Another method is to use IoU (Intersection over Union). This measures the amount of overlap between the original bounding box and predicted bounding box. This is a good way to find out how precise the network is, compared to original data.

AP or average precision is calculated by finding the area under a precision-recall curve. Because this method combines both precision and recall, it is considered a better alternative compared to others mentioned here. mAP (mean Average Precision) indicates the average of AP across all classes of a network. Since we will be using only one class, mAP and AP will be referred as the same method from now on. All the mAP scores showed in this paper will be considered as measured at %50.

$$mAP = \frac{\sum_{q=1}^Q AP(q)}{Q} \quad (3)$$

2.3. Progress and Other Work

After the success of the original YOLO, Redmon and Farhadi continued to improve this method's performance, by developing YOLO v2, also known as YOLO9000. This algorithm can detect over 9000 categories, thus the name YOLO9000. This model was aimed at improving YOLO's shortcomings of localization errors and low amount of recall [8].

Because this version is aimed mostly at increasing the amount of object categories and since we will have only one object category during our testing, we will not be working with YOLO v2.

Redmon and Farhadi didn't stop with YOLO v2. They also worked on an even more improved model, YOLO v3. This model improved the already-fast algorithm by making it even faster and surprisingly more accurate than the last [9]. This algorithm was measured to have 57.9 mAP score on COCO data set, with IoU taken as 0.5. While it falls behind when compared to RetinaNet, it is far better than YOLO v2 and somewhat better than older models like faster-RCNN. Though the mAP score doesn't take speed into account. We need to give up some of our precision to gain the speed YOLO algorithms can achieve.

YOLO v4 was developed by Wang, Bochkovskiy and Liao, in February 2021. This one is the most recent method mentioned in our paper. There are three main sub-types within this method; large, normal, and tiny. V4 Large performed 55.5% AP on MS COCO dataset, at 16 FPS while the tiny version performed an AP score of 22.0% at 443 FPS [10]. This is important for our work since we will measure the speed on a lower-end processing capacity. And 443 FPS is more than enough for a live video application, in which the laser images will be used on. It is stated that YOLO v4 tiny can even reach a computing speed of 1774 frames per second. With this information, YOLO tiny appears to be the fastest performing method, even while not losing too much of precision.

There are multiple works related to YOLO and its successors. Recently, people applied YOLO to aid in COVID-19 pandemic, mostly aimed at detecting medical face masks worn by the public [11], [12], [13]. One of them features a YOLO v2 model based on ResNet-50 deep transfer learning model, tested with both adam and SGDM. This achieved promising results, with an AP score of 82% [11]. YOLO-LITE [14] is another fast network like v4 tiny, developed in November 2018. This one achieved 12.26% score on COCO dataset at 21 FPS, running on a non-GPU computer. It also achieved a mAP of 33.81% on PASCAL VOC dataset. This was not taken into comparison for our work, since YOLO v4 tiny is the more recent one with a much higher FPS rate. YOLO algorithms aren't limited with COCO or PASCAL datasets. There are also examples of custom dataset applications such as ours. One of them is poisonous mushroom detection, conducted by Cengil and Çınar [15]. They created a dataset containing 8 categories of poisonous mushrooms, specifically. This means there are 8 classes in training. The images were taken from the internet, but the dataset is small by the authors. Worst performing class, webcaps, had 0.556 AP while the best performing one, that being Fly Agaric, scored 0.993 AP. mAP score for this dataset was 0.778. The algorithm for this was YOLO v5. Another work features YOLO v3 and several other networks, tested on thermal cameras for the purpose of finding illegal immigrant or terrorist movements. The results once again proved that YOLO is the fastest among the others, with minimal drop of detection success [16]. This work gives an idea as to how our work will perform because we also use thermal imaging as our source. The difference will be that they used people as objects whereas we use laser as objects.

3. EXPERIMENTAL WORK

Our data set contains 74 infrared laser images taken from an IR camera, simultaneously and heavily used in military aviation. These images are taken from several different wavelengths under diverse conditions and intensities. Final goal will be to use the algorithm with

15 to 30 FPS processing capability on a low-end or middle-end GPU. Out of the 74-image data set, 38 of the images are used in training while 26 are used in testing. Training is done on a i5-4690K CPU. YOLO v3, YOLO v4 and YOLO v4 tiny algorithms are trained and compared within each other.

3.1. YOLO v3

First training was done with YOLO v3. Maximum epoch size is 100, each epoch containing 10 steps. This process of training took 2 hours to complete on the CPU mentioned above. The result is a mAP score of 88.462 at 1.89 FPS. The FPS value is lower than expected from a YOLO algorithm, which is caused by the usage of a CPU instead of a GPU.

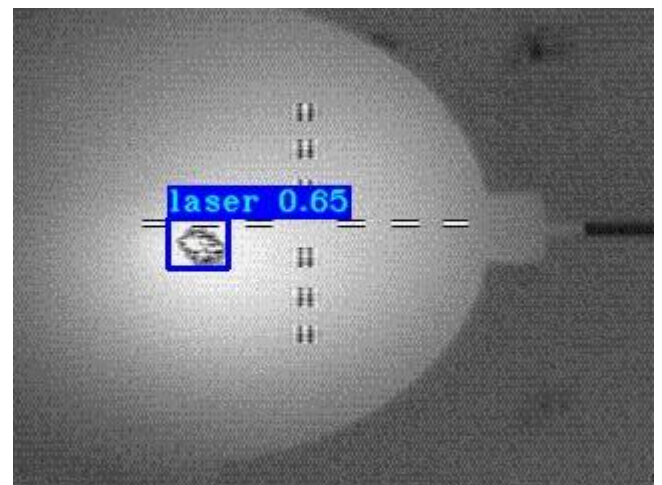


Figure 1. First sample of laser image on YOLO v3

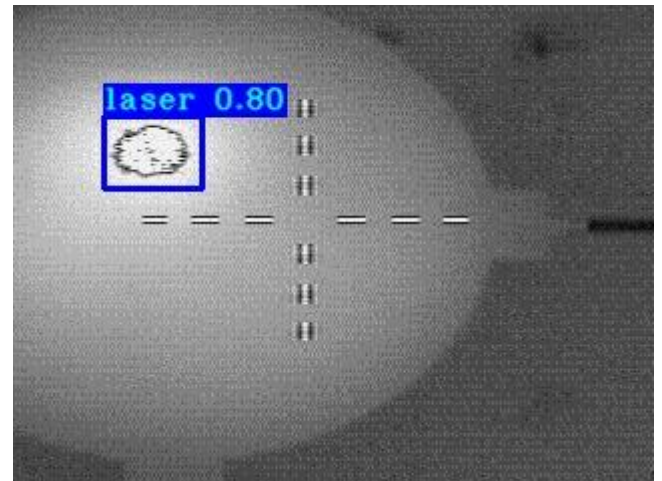


Figure 2. Second sample of laser image on YOLO v3

3.2. YOLO v4

YOLO v4 is applied on the same data set. Training parameters are kept same as YOLO v3 experiment for comparison. This training took just below 2 hours to complete. The result is a mAP score of 88.154 at 2.34 FPS. Frames processed per second increased compared

to YOLO v3. mAP score seems to be about the same as YOLO v3. The difference is considered negligible.

On one of the test images, this algorithm falsely found two lasers on top of each other. This didn't happen in either of the other two algorithms. This is thought to be one key difference from YOLO v3 in our experiments.

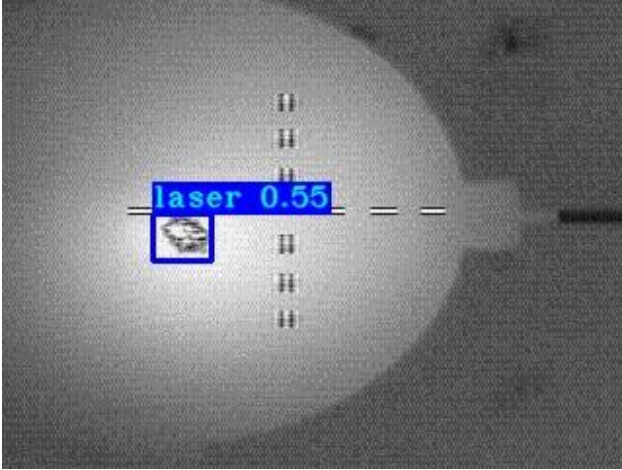


Figure 3. First sample of laser image on YOLO v4

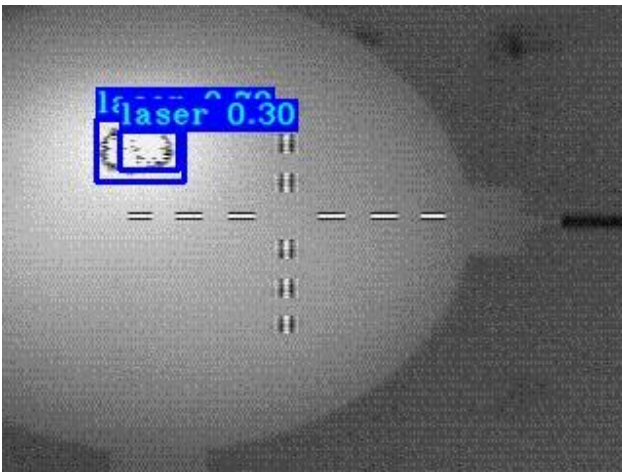


Figure 4. Second sample of laser image on YOLO v4

3.3. YOLO v4 Tiny

With the tiny algorithm of YOLO, we were expecting a higher FPS with a lower mAP score compared to rest. This was indeed the case, YOLO v4 Tiny performed significantly faster, at 11.89 FPS with a mAP score of 87.865.

YOLO v4 Tiny seems to be the overall best performing algorithm among our three YOLO versions.

3.4. Compared Results

With every test completed, we can finally compare the results of these methods. As expected, YOLO v4tiny is the fastest algorithm with the lowest mAP score. (Though the score is not much different than the others.)

The comparison shows that there is no need to use slower and more resource-intensive algorithms if speed of detection is more important than precision of detection. But even if precision was the key parameter for a work, the difference of precision is so minor that it can be ignored.

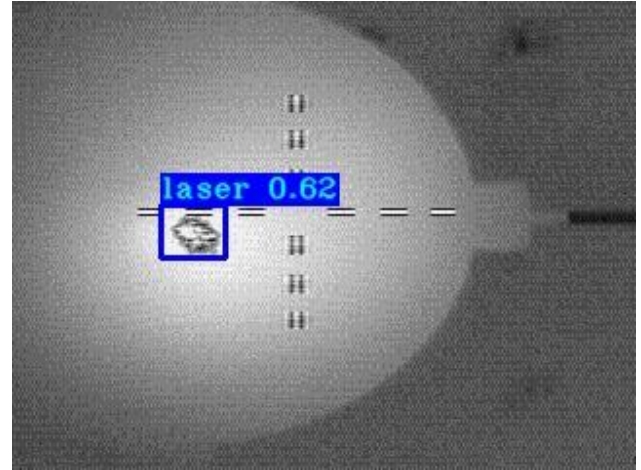


Figure 5. First sample of laser image on YOLO v4 tiny

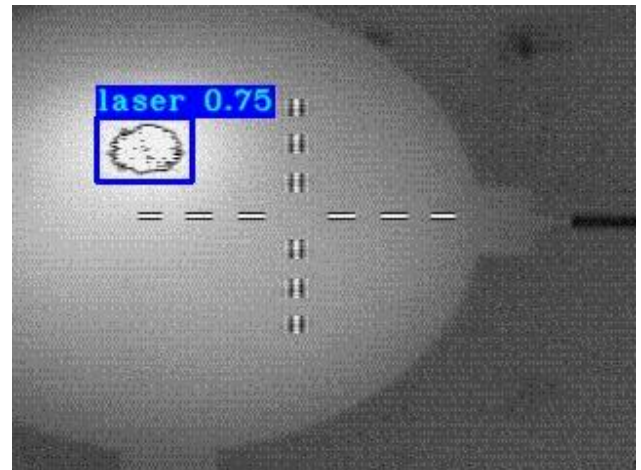


Figure 6. Second sample of laser image on YOLO v4 tiny

Table 1. Results of three different YOLO networks

Method	mAP	FPS
YOLO v3	88.462%	1.89
YOLO v4	88.154%	2.34
YOLO v4 tiny	87.865%	11.89

4. CONCLUSION

Experiments were done under the same conditions, same equipment, and the same parameters. Three YOLO versions were compared on a data set containing 74 images. Training time was recorded for comparison. After the training, 3 images were tested to see differences among the versions.

YOLO v3 seemed to be the least efficient. While it has the highest mAP score, it also has the lowest detection speed by a large margin. Compared to YOLO v4, the training time was almost the same. mAP scores are also almost identical. YOLO v4 outperformed YOLO v3 on detection speed. YOLO v3 has been the least desired choice in the experiment. On the other hand, YOLO v4 tiny detects objects much faster than the others, while keeping up with detection precision.

It should be kept in consideration that we only have one class, that is laser, in our data set. Our results don't reflect what these algorithms can perform on a multi-class data set. It should also be considered that laser images are simple to find, especially on a flat surface as used in our experiments. Future work will be conducted to test how these networks compare under more noisy backgrounds. The results really show which method is the best for one class such as a laser image taken from an infrared camera. YOLO v4 or maybe even v3 might have performed significantly better than the v4 tiny version, in a different set containing more than one class. In our case, YOLO v4 tiny is the best choice for our work.

5. FUTURE WORK

The aim for this paper and the experiments done are to determine which version of YOLO algorithm is the most suitable for determining the location of laser spots on an infrared camera. This version of YOLO algorithm will be used in the making of a thesis paper in the future.

ACKNOWLEDGEMENT

The code used in the learning was mostly borrowed from <https://github.com/pythonlessons>. We wish to thank Redmon and Farhadi for creating the original YOLO and for their great contribution to the object detection literature.

REFERENCES

- [1] R. Girshick, "Fast R-CNN", **Proceedings of the IEEE international conference on computer vision**, Boston, MA, USA, 1440-1448, June, 2015.
- [2] H. Jiang, E. Learned-Miller, "Face detection with the faster R-CNN", **2017 12th IEEE international conference on automatic face & gesture recognition**, Washington DC, USA, 650-657, June, 2017.
- [3] J. Redmon, S. Divvala, R. Girshick, A. Farhadi, "You only look once: Unified, real-time object detection", **Proceedings of the IEEE conference on computer vision and pattern recognition**, Las Vegas, NV, USA, 779-788, June, 2016.
- [4] Internet: National Aeronautics and Space Administration, Science Mission Directorate, http://science.nasa.gov/ems/07_infraredwaves, 03.11.2021.
- [5] Internet: American Technologies Network Corporation, <https://www.atncorp.com/hownightvisionworks>, 04.11.2021.
- [6] J. M. Lloyd, **Thermal imaging systems**, Springer Science & Business Media, Berlin, Germany, 2013.
- [7] S. P. Mahulikar, H. R. Sonawane, G. A. Rao, "Infrared signature studies of aerospace vehicles", *Progress in Aerospace Sciences*, 43(7-8), 218-245, 2007.
- [8] J. Redmon, A. Farhadi, (2017). "YOLO9000: better, faster, stronger.", **Proceedings of the IEEE conference on computer vision and pattern recognition**, Honolulu, HI, USA, 7263-7271, July, 2017.
- [9] J. Redmon, A. Farhadi, (2018). "Yolov3: An incremental improvement", *arXiv preprint arXiv:1804.02767*, 2017.
- [10] C. Y. Wang, A. Bochkovskiy, H. Y. M. Liao, "Scaled-yolov4: Scaling cross stage partial network." **Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition**, 13029-13038, Online, June, 2021.
- [11] N. S. Punn, S. K. Sonbhadra, S. Agarwal, G. Rai, "Monitoring COVID-19 social distancing with person detection and tracking via fine-tuned YOLO v3 and Deepsort techniques", *arXiv preprint arXiv:2005.01385*, 2020.
- [12] C. R. Rodriguez, D. Luque, C. La Rosa, D. Esenarro, B. Pandey, "Deep learning applied to capacity control in commercial establishments in times of COVID-19", **2020 12th International Conference on Computational Intelligence and Communication Networks (CICN)** 423-428, Bhimtal, India, September, 2020.
- [13] M. R. Bhuiyan, S. A. Khushbu, M. S. Islam, "A deep learning based assistive system to classify COVID-19 face mask for human safety with YOLOv3", **2020 11th International Conference on Computing, Communication and Networking Technologies (ICCCNT)** 1-5, Kharagpur, India, July, 2020.
- [14] R. Huang, J. Pedoeem, C. Chen, "YOLO-LITE: a real-time object detection algorithm optimized for non-GPU computers", **2018 IEEE International Conference on Big Data (Big Data)**, 2503-2510, Seattle, WA, USA, December, 2018.
- [15] E. CENGİL, A. ÇINAR, "Poisonous Mushroom Detection using YOLOV5", *Turkish Journal of Science and Technology*, 16(1), 119-127, 2021.
- [16] M. Krišto, M. Ivasic-Kos, M. Pobar, "Thermal object detection in difficult weather conditions using YOLO", *IEEE Access*, 8, 125459-125476, 2020.

ISO/IEC 27001 Information Security Management System and Cybersecurity Maturity Model Certification-CMMC Applications

İsmail Yiğit COŞAR, Aslıhan TÜFEKÇİ

Informatics Institute, Gazi University, Ankara, Turkey
yigitcsr@gmail.com, asli@gazi.edu.tr

Abstract— With the developing technology, threats to information security are also increasing rapidly. Accordingly, institutions especially need to strengthen their information security measures. Today, ISO/IEC 27001 Information Security Management System Standard, which is the successor of the first information security management system standard BS7799-1 published by BSI in 1995, is still widely used in this field. However, changing security needs with the developing technology have created new applications specific to the kind of information and the sector. One of the most recent examples of these is the U.S. Cyber Security Maturity Model Certification. In this study, the two applications mentioned were compared in terms of general features, scope, requirements and certification, and an evaluation was made for the future of information security applications.

Keywords— cyber security, information security, maturity model, ISO/IEC 27001, management system, CMMC

1. INTRODUCTION

"Digitalization", which has an impact on every aspect of our lives with the developing technology, has increased the variety and size of information assets, and has made it mandatory for individuals and institutions to take precautions regarding information security. In the last 50 years, especially with the great developments in the field of technology, the control of the real world has been transferred to the cyber world to a significant extent, so the concept of "information", which stands out in every field, has started to be used almost synonymously with the concept of "power" [1]. As it can be understood from here, information is the basis of the power required to ensure security, and any weakness that may occur in the protection of information assets will pose a significant threat to the sustainability of the assets. In this respect, the cost of not giving importance to information and data processing security is considerably higher than expected [2].

For example, according to a study conducted in the United States in 2018, it is stated that harmful cyber activities only cost the US economy between 57 and 109 billion dollars in 2016 [3]. When we look at the world, it is seen that this amount has increased to 1% of the total gross domestic product of the world countries, that is, to 600 billion dollars in 2016 [4]. In this respect, the reasons for cyber-attacks, which have a very high damage potential, are very diverse,

but the dimensions of such attacks can be on a personal, institutional or national scale. And also, it is possible to encounter international attacks with developed countries behind them in many cases. From this point of view, it can be said that the international cold wars of the past are taking place in the cyber environment today [5].

In order to protect their information security infrastructure, developed countries have been continuing their efforts to reduce risks and ensure security in the cyber environment, starting with the activities of preparing standards in the field of information security, since the 90s. In this field, in 1995, BSI (British Standards Institution) published the information security management standard BS 7799-1:1995 in the UK and laid the foundations of ISO/IEC 27001, which is the widely used information security management system standard today [6]. Since then, many international organizations (AQAP, IEEE, ETSI, ITU, NIST, ISACA etc.) have published many standards in the field of information security and put them at the service of individuals and institutions [7]. Some of the security measures implemented in this context can be applied only in the relevant sector as they are developed specifically for certain sectors in line with their purpose, while others can be applied in all sectors as they are prepared for basic activities.

In this study, two separate information security management applications, which can be shown as an example of this situation; the EN ISO/IEC 27001:2017 Information Security Management System Standard, which is prepared considering the general activities of the organizations, and the Cyber Security Maturity Model Certification (CMMC) applications developed for the defense industry, both have been evaluated in terms of general features, scope, requirements and certification, and a prediction has been made about how information security practices will change in the future.

2. ISO/IEC 27001:2017 INFORMATION SECURITY MANAGEMENT STANDARD

2.1. General Features

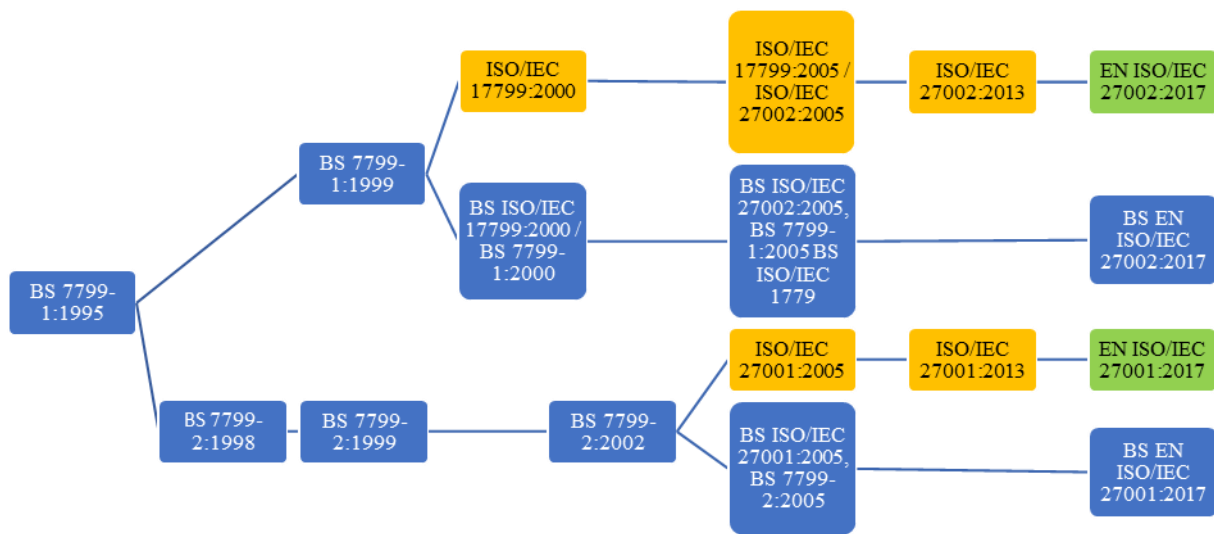


Figure 1. EN ISO/IEC 27001 development stages

The first version of the ISO/IEC 27001 Information Security Management System standard was prepared in 2005 by the Information Security, Cyber Security and Privacy Protection Subcommittee SC27, affiliated to the Information Technologies Joint Technical Committee JTC1 within ISO [9].

The EN standard with the number EN ISO/IEC 27001:2017 is the current version of the ISO/IEC 27001:2013 standard published by CEN (European Committee for Standardization) in 2017 [10]. There is no difference in requirements between this latest version published by CEN and the previous version published by ISO. And the version published by CEN was also accepted as “Turkish Standard” by TSE in 2017, translated into Turkish and published in our country in 2018 with the title of TS EN ISO/IEC 27001 “Information technology-Security techniques- Information security management systems- Requirements” [11]. In this study EN version of the standard which is published by CEN is used for refer ISO/IEC 27001.

Until the 1990s, the responsibility of taking precautions against threats to information systems, which increased with the spread of information technologies, was only on the information technology units of the organizations. After with the concept of “Information Security Management System” which is foreseen in the BS 7799-1 standard published by BSI in 1995, the responsibility of ensuring information security and related duties spread throughout the organization [8]. As can be seen in the development stages shown in Figure 1 this standard is also the basis of the ISO/IEC 27001 Information Security Management System Standard, which is published by ISO (International Organization for Standardization) and defines the conditions that an information security management system that can be created independently from the fields of activity in organizations.

2.2. Scope

Within the scope of the EN ISO/IEC 27001:2017 standard, requirements are defined for the establishment, implementation, maintenance and continuous improvement of an information security management system appropriate to the context of the organization. The aforementioned requirements have been created in such a way that they have a general nature due to the structure of the standard and are applicable to all sectors.

Within the standard, the requirements are defined under clauses 4 to 10. However, in order to fully meet the "Planning" requirements expressed in Article 6, a total of 112 control conditions under 18 main headings are included in the standard annex.

Although there is no obligation for the implementation of the standard, legislation, contract, etc. Due to other special conditions, situations may arise in organizations where it is necessary to apply the standard.

The EN ISO/IEC 27001:2017 standard also has the feature of being the most widely accepted information security management system standard in the international arena, as it has been accepted by CEN, which is a member of 34 countries in Europe, and ISO, an independent organization of which 165 countries are members in the world [12].

2.3. Requirements

In order to access the standard content that includes the requirements, organizations must purchase the standard document for a fee. The Turkish and English versions of the standard are sold by TSE, and versions in other foreign languages can be accessed digitally on the websites of ISO and CEN.

In the standard content; the requirements and other issues regarding the establishment, implementation, maintenance and continuous improvement of the information security management system are expressed under a total of 11 main titles. These titles consist of general expressions used by ISO according to the "high-level structure" HLS applied in management system standards [13].

The requirements that must be met within the scope of the standard due to the "high-level structure" in question are defined in the headings specified in articles 4 to 10. The requirements defined by these articles are applicable to all types of organizations and information assets, regardless of any criteria, level or sector.

Accordingly, an organization shall carry out;

- In accordance with the "4. Context of the organization" article; identifying and scoping internal and external issues that may affect the objectives of the information security management system in line with its objectives,
- In accordance with the "5. Leadership" article; providing senior management to show their support for the information security management system, establish the information security policy and define the institutional roles,
- In accordance with the "6. Planning" article; identifying risks and opportunities, evaluating risks, defining the risk handling process, and determining the objectives of the information security management system and planning how to achieve it,
- In accordance with the "7. Support" article; defining the resources and competencies for the information security management system, and providing controls on awareness, communication and written information,

- In accordance with the "8. Operation" article; ensuring the planned operation of the information security management system

- In accordance with the "9. Performance evaluation" article; performing monitoring, measurement, analysis-evaluation, internal audit, management review,

- In accordance with the "10. Improvement" article; in case of non-compliance, taking corrective action and maintaining continuous improvement of the information security management system,

activities in sustainable and effective way in order to be certified according to the standard [14].

In the table in the annex to the standard, the control objectives and controls that are included in the "6. Planning" article and will be used as a reference in the implementation of the selected information security risk processing activities are listed directly with reference to the expressions specified in the articles 5 to 18 of the ISO/IEC 27002:2013 standard [15].

2.4. Certification

ISO does not carry out certification activities due to its structure and purposes, but certification activities according to ISO standards are carried out by independent certification bodies. In line with the requirements of the EN ISO/IEC 27001 standard, the organizations that create the information security management system are required to apply to these certification bodies in order to carry out third-party audits and certify their systems.

It is important that the certification body referenced here is accredited by an accreditation body that is a member of the IAF (International Accreditation Forum) according to the (TS) (EN) ISO/IEC 17021-1:2015 standard, which defines the conditions for the organizations that provide the audit and certification of management systems within the scope of conformity assessment. In this way, the national and international validity of the acquired document will be guaranteed [16].

3. CYBER SECURITY MATURITY MODEL CERTIFICATION (CMMC)

3.1. General Features

Cyber Security Maturity Model Certification is an information security management model that the United States Department of Defense (DoD) expects from its suppliers to strengthen the information security infrastructure in the defense industry projects it carries out, and the requirements of the model are expressed in 5 different maturity levels depending on the application.

The US defense industry sector (DIB), which is the world's largest defense industry [17], consists of a globally influential complex industrial structure that provides the design, manufacture, delivery, and maintenance as well as research and development of military weapon systems, subsystems, and components or parts to meet U.S. military requirements.

The US defense industry consists of more than 100,000 defense industry companies and their subcontractors operating under contract with the US Department of Defense, including manufacturing assets located in many countries and domestic and foreign organizations, and the product required to mobilize, deploy and maintain military operations. and provides services. However, the sector does not include the commercial infrastructure of service providers such as power, communications, transportation or utilities that the US Department of Defense uses to meet military operational requirements [18].

Due to the fact that the sector is so large and comprehensive, it is considered by the US Department of Defense that significant national security risks will arise if the security of the information in the defense industry sector, expressed as CUI (Controlled Unclassified Information) and FCI (Federal Contract Information), is not properly ensured [19].

While the CUI mentioned here is defined as information which is government-created or owned, or an entity created or owned by or on behalf of government, non-confidential but requires protection or dissemination control in accordance with and consistent with applicable laws, regulations, and government-wide policies information other than those classified under Executive Order 13526-Classified National Security Information or the Atomic Energy Act, FCI means non-public information provided or created for the government under a contract to provide or develop a product or service to the government.

The relations of two different types of information with each other and with public information are shown in Figure 2 [20].

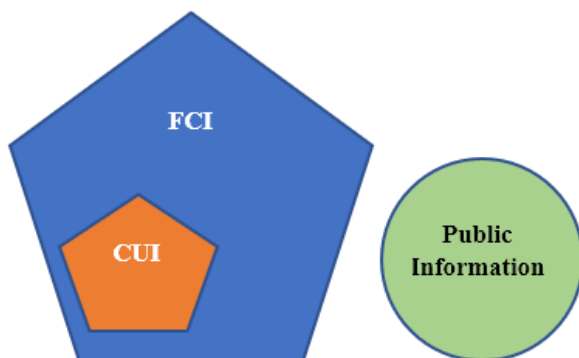


Figure 2. Types of information.

As it can be understood from here, while FCI covers all the information that is not expected to be disclosed to the public after the conclusion of the contract between the US

Department of Defense and defense industry organizations, CUI refers to the information that needs to be protected with additional protection measures.

To reduce the risk of compromise of this information and ensure information security in non-confidential communication networks, The Office of the Under Secretary of Defense for Acquisition and Sustainment-OUSD(A&S) has carried out joint studies with U.S. Department of Defense stakeholders include university affiliated research centers (UARC), federally funded research and development centers (FFRDC) and the sector. Finally, as a result of these studies, it has developed an information security management model called CMMC by taking advantage of many information security guidelines and standards such as NIST SP 800-171, NIST SP 800-53, NAS9933, CERT RMM v1.2, FAR 52.204-21, DFARS 252.204-7012 and FIPS 200 in order to define common information security criteria.

3.2. Scope

The aim of CMMC is to set out a framework of information security measures to be applied to protect the CUI and FCI defined above, considering the information flow to subcontractors in a multi-layered supply chain in defense industry sector organizations. Thus, it is aimed to provide more assurance to the US Department of Defense that the relevant defense industry organization can adequately protect sensitive information [21].

Within the scope of CMMC, there are 5 different maturity levels, 27 different domains and a total of 171 practices defined in line with the above-mentioned objectives and targets.

Regarding CMMC applications in the sector, the legislation "Defense Federal Procurement Regulation Addendum: Evaluation of the Contractor's Implementation of Cyber Security Requirements" (DFARS Case 2019-D041) was published in the Federal Register on 19/09/2020 to be valid on 30/11/2020.

In accordance with this legislation, before 1/10/2025, US Department of Defense contractors and subcontractors are required to make their information security infrastructures comply with CMMC requirements. After that date, the bidder or contractor will not be eligible to participate in U.S. Department of Defense contracts unless they hold a valid (i.e. not older than three years) certification for the required CMMC level. In addition, CMMC certification requirements are expected to be passed on to subcontractors at all levels, depending on the sensitivity of non-confidential information conveyed to each subcontractor [22].

As it can be understood from here, although CMMC practices are national in terms of legislation, it is seen that the implementation of the CMMC model will be mandatory for all defense industry organizations working

with the US Department of Defense and their sub-suppliers worldwide as of the specified date.

3.3. Requirements

Unlike the information security management system standards, the current version of the framework document for the implementation of the CMMC model, CMMC V1.02 (CMMC_ModelMain_V1.02_20200318) and other reference documents are publicly available on the website of the Procurement and Maintenance Office of the US Department of Defense [23].

In the main document of the model, the cyber security maturity model is discussed at 5 different levels. The maturity model mentioned here is an evaluation approach consisting of a number of key process areas and different levels of maturity to evaluate the organization's performance in managing the business process [24] and accordingly, for each level, separate processes and practices are included in the document according to the type and sensitivity of the information that needs to be protected and is the subject of the threat [25]. For example, the document defines "basic protection" requirements for FCI and more advanced "security" requirements for CUI.

Within the CMMC model, there are a total of 17 "fields" and 43 expressed "capabilities" based on the Federal Information Processing Standards (FIPS) Publication 200 and NIST SP 800-171.

In this framework, "processes" and "practices" were defined within these "domains" and distributed to 5 different maturity levels. Accordingly, the applications expected to be realized within the scope of process maturity and capabilities for each domain within the scope of the model are grouped into 5 different levels as shown in Figure 3. Thus, for each level, a total of 171 practices were defined to be implemented in the relevant domain. These practices are distributed between levels with; 17 for Level 1, 72 for Level 2, 130 for Level 3, 156 for Level 4, and 171 for Level 5.

The defense industry organization can perform CMMC applications in its entire organization or only in the unit/section where the information that needs to be protected is used and stored. In this direction, in order to meet the aimed level requirements, the organization shall meet the process maturity requirements and also perform the practices at the aimed and preceding levels in all domains. If the process maturity level and the level of the implemented practices are different, it is accepted that the lower-level requirements are met.

As can be understood from the above, each level determines the specifications of the practices, while also expressing the extent to which the processes are institutionalized. Here, it is foreseen that the organization will carry out the necessary activities in a sustainable,

effective and high quality as the institutionalization scale of the processes increases.

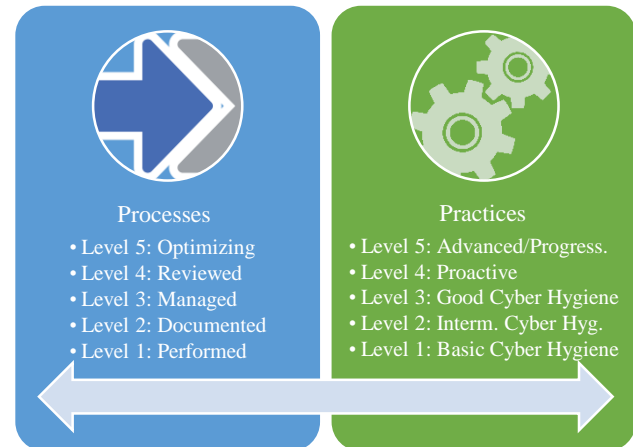


Figure 3. CMMC levels

3.4. Certification

The US Department of Defense or its Procurement and Maintenance Office does not carry out any certification activities within the scope of the CMMC model.

In addition, the US Department of Defense has authorized the CMMC Accreditation Body (CMMC-AB), which is an independent organization and accredited according to the ISO/IEC 17011 standard which defines requirements for bodies accrediting conformity assessment bodies, to act as the regulatory authority in the execution of the business and transactions defined within the scope of the CMMC model [26].

As the CMMC-AB regulatory authority, it authorizes and accredits CMMC third party audit organizations (C3PAO) to carry out tasks such as third-party auditing and certification, and CMMC auditor and trainer certification organizations (CAICO) for certification of CMMC auditors and trainers.

Thus, in order for an organization to be authorized as C3PAO, it must be accredited according to the ISO/IEC 17020 standard which defines the conditions for the operation of various types of inspection bodies, and to be authorized as CAICO, it must be accredited according to the ISO/IEC 17024 standard which defines the general conditions for personnel certification bodies. However, CMMC-AB can authorize without accreditation if the organization confirms that it meets the relevant requirements. In such a case, the organization must be accredited within the scope of the standard related to its field of activity within 27 months from the date of authorization [27].

Organizations that fulfill the CMMC model requirements in line with the above should apply to one of the C3PAO organizations authorized by CMMC-AB for third party

audits and certification. The list of C3PAO organizations is published on the CMMC-AB website.

4. CONCLUSION

With the increasing importance of information, threats to information security are also increasing. Accordingly, existing information security defense infrastructures must be constantly updated and developed and new security models must be created in order to maintain their effectiveness. In this study, the EN ISO/IEC 27001 standard and CMMC applications, which are two different information security management models that can set an example for different stages of this development process, are discussed.

As stated here, the studies for enterprise-scale information security which are started with BSI's first information security standard BS 7799 in 1995 [28], has revealed the current version of EN ISO/IEC 27001, which is the most widely used information security management standard today, in line with the developing and changing threats and the security needs arising from them.

With the EN ISO/IEC 27001 standard, it is aimed to establish the necessary requirements for the establishment and certification of an information security management system that can provide adequate protection for all types of organizations and all types of information assets, independent of the sector. In this respect, it is seen that the scope of application and certification of the standard is very wide, but it includes general measures that can now be considered basic in terms of requirements.

However, today, threats to information security are changing much faster than in the past, and potential targets are increasing accordingly. This situation requires taking more advanced, more detailed and more effective information security measures, which are customized according to both the type of information assets and the sector, unlike previous practices at the point of information security management.

In this context, the US Department of Defense has developed an information security management model for its supplier organizations, together with related communities and primarily to protect national security.

This model, called CMMC, focuses only on a certain sector and certain types of information in terms of scope, unlike the EN ISO/IEC 27001 standard. In this way, it is seen that the requirements defined in CMMC are more detailed and effective, specific to the relevant sector and information types.

From this point of view, it is considered that the studies to increase or protect the existing security levels about information security will focus on sectoral and information types specific gradual management models rather than

general and comprehensive management models in the future, and accordingly, the effectiveness of basic information security management standards will decrease over time.

And also, it is thought that it is critical to create and implement information security management models with sector and information type specific requirements, especially in the public and defense industry sector similar to CMMC, for eliminating cyber threats to the national security of our country.

REFERENCES

- [1] A. Akçoraoglu, "Yeni Kapitalizm Teorileri, Dijital Devrim ve Türkiye Kapitalizmi", *Mülkiye Dergisi*, 43(3), 525–575, 2019.
- [2] G. Canbek, Ş. Sağıroğlu, **Bilgi ve Bilgisayar Güvenliği: Casus Yazılımlar ve Korunma Yöntemleri**, Grafiker Yayıncılık, Ankara, Turkey, 2006.
- [3] McAfee, **The Economic Impact of Cybercrime—No Slowing Down Executive Summary**, 2018.
- [4] The Council of Economic Advisers, **The Cost of Malicious Cyber Activity to the U.S. Economy**, 2018.
- [5] Ş. Sağıroğlu, M. Alkan, R. Samet, **Siber Güvenlik ve Savunma-Farkındalık ve Caydırıcılık**, Grafiker Yayıncılık, Ankara, Turkey, 2018.
- [6] E. Humphreys, "Information Security Management System Standards", *Datenschutz und Datensicherheit - DuD*, 35(1), 7–11, 2011.
- [7] Ş. Sağıroğlu, O. Aktaş, O. Alkan, **Siber Güvenlik ve Savunma-Standartlar ve Uygulamalar**, Grafiker Yayıncılık, Ankara, Turkey, 2019.
- [8] E. Humphreys, "Information Security Management Standards: Compliance, Governance and Risk Management", *Information Security Technical Report*, 13(4), 247–255, 2008.
- [9] Internet: ISO Technical Committees, <https://www.iso.org/technical-committees.html>, 05.06.2021.
- [10] Internet: CEN, CEN/CLC/JTC 13- Cybersecurity and Data Protection, <https://standards.cen.eu/>, 05.06.2021.
- [11] Internet: TSE, "TSE.NET Standard Detayı TS EN ISO/IEC 27001:2017", <https://intweb.tse.org.tr/Standard/Standard/Standard.aspx?081118051115108051104119110104055047105102120088111043113104073083043047076078043057108043104101>, 05.06.2021.
- [12] H. Susanto, M. N. Almunawar, Y. C. Tuan, "Information Security Management System Standards: A Comparative Study of the Big Five", *International Journal of Engineering and Computer Science*, 11(5), 23–29, 2011.
- [13] Internet: ISO, ISO Management System Standards, <https://www.iso.org/management-system-standards.html>, 05.06.2021.
- [14] TSE, **TS EN ISO/IEC 27001**, TSE, Ankara, Turkey, 2018.
- [15] ISO, **ISO/IEC 27002**, ISO, Geneva, Switzerland, 2013.
- [16] Internet: ISO, Certification, <https://www.iso.org/certification.html>,

- 06.06.2021.
- [17] SIPRI, **World military spending rises to almost \$2 trillion in 2020**, 2021.
- [18] Internet: CISA, Defense Industrial Base Sector, <https://www.cisa.gov/defense-industrial-base-sector>, 06.06.2021.
- [19] Internet: National Archives, About CUI, <https://www.archives.gov/cui/about>, 06.06.2021.
- [20] Internet: National Archives, FCI and CUI what is the difference?, <https://isoo.blogs.archives.gov/2020/06/19/fci-and-cui-what-is-the-difference/>, 06.06.2021.
- [21] Internet: OUSD(A&S), CMMC Home Page, <https://www.acq.osd.mil/cmmc/index.html>, 06.06.2021.
- [22] Department of Defense, “DFARS Case 2019-D041”, *Federal Register*, 29/09/2020, 61505–61522, 2020.
- [23] Internet: OUSD(A&S), “CMMC Model and Assessment Guides”, <https://www.acq.osd.mil/cmmc/draft.html>, 06.06.2021.
- [24] X. Meng, M. Sun, M. Jones, “Maturity Model for Supply Chain Relationships in Construction”, *Journal of Management in Engineering*, 27(2), 97–105, 2011.
- [25] OUSD(A&S), **CMMC v1.02 Model Main Document**, OUSD(A&S), 2020.
- [26] Internet: CMMC-AB, CMMC-AB Home Page, <https://cmmcab.org/>, 12.06.2021.
- [27] Internet: OUSD(A&S), CMMC Faq, <https://www.acq.osd.mil/cmmc/faq.html>, 12.06.2021.
- [28] E. Ersoy, M. Alkan, “Bilgi Güvenliğinin Kurumsal Bazda Uygulanması”, Bilgi Güvenliği ve Kriptoloji Konferansı, Ankara, Turkey, 13-14 December, 2007.

Natural Language Processing Based Recommendation System to Improve Patent Preliminary Search

Yavuz YILMAZ¹, Tahsin ÇETİNYOKUŞ²

¹ Department of Management Information Systems, Gazi University, Ankara, TURKEY

² Department of Industrial Engineering, Gazi University, Ankara, TURKEY
yavuzdesign@gmail.com, tahsinc@gazi.edu.tr

Abstract—In recent years, there has been an increase in the amount of patent applications in parallel with technological developments. Patent researches with traditional methods cannot keep up with this rate of increase. In this direction, it is essential to develop tools that will help the patent examiner and the patent applicant by processing the patent data. In this study; using deep learning, machine learning and natural language processing techniques, a model have been presented that can help users in patent research. A model has been developed that aims to obtain documents that are closest in content to the patent document to be researched. In the training of this model, the description, abstract and claim texts of the 376,609 patent applications made to the American Patent and Trademark Office in 2018 are used. The developed system is tested in 8,085 patent applications published in the first week of 2019. In the model; cosine similarity, euclidean and manhattan distance are used. The patent classification code is chosen as the performance evaluation criterion of the developed model. 82% success is achieved as a result of testing the model.

Keywords— natural language processing, nlp, machine learning, deep learning, patent

1. INTRODUCTION

In patent documents, all the features of the invention that are desired to be protected are written clearly without being hidden. In addition, since these documents contain classified technical data, they contain important information that cannot be found in any other document type. Moreover, eighty percent of the information found in patent documents cannot be found anywhere else [1]. Another feature of patent documents is that they contain structured data. This open-source and classified nature of patent data makes it suitable for data science processing.

In order to prevent possible patent infringements, it is necessary to investigate the patent applications related to the subject of the study (aka state-of-the-art) before R&D work is carried out in any technological field. For people unfamiliar with industrial property issues, this research process can be quite challenging. In addition, while evaluating the patentability of patent applications, patent experts have to examine all documents in the state of the art about that subject. If access to these documents is done by traditional methods, it is achieved by manually scanning in patent databases.

With the increase in awareness of industrial property and developing technology in recent years, there has been a significant increase in the number of patent and utility

model applications [2]. Compared to this rate of increase, conventional research methods are quite slow and require a lot of labor. Accordingly, it is an essential need to develop tools that will assist the patent researcher and the applicant by processing the patent data.

In this study, a recommendation system is created that can support patent experts and applicants in patent search. In the following parts of the study, studies in the literature are presented. In the next section, the model developed within the scope of the study is explained and the test results are presented. In the last part, the results of the study are interpreted and suggestions are made.

1.1. Studies on Natural Language Processing Based Patent Search

Article by Wu et al. (2020) is about the examination of patent documents related to information processing technologies (ICT) with natural language processing and deep learning techniques [3]. Patent data is trained with MLP (Multilayer Perceptron) deep learning model after pre-processing and vectorization stages. The results were compared with other machine learning models.

Chen et al. (2020) study is about extracting meaningful information from patent documents [4]. After the pre-

processing of the patent data, technological information extraction is done with deep learning. This information is processed and used for technological trend analysis and detection of patent infringement.

Article by Kim et al. (2020) is about clustering of patent documents with deep embeddings [5]. Two data sets consisting of 3 and 5 categories were used in the study. In the developed model, abstracts of twenty thousand patent documents were used as data. Patent documents have been vectorized with the Doc2Vec algorithm after text preprocessing. In the next stage, documents are clustered with the DEC (Deep embedding clustering) algorithm. With this technique, the distances between different documents are increased and the distances between similar documents are reduced. In this way, documents with close content are better categorized.

Studies based on the processing of patent data in the literature can be grouped under four main headings: automatic classification of patents, technological trend analysis based on patent data, detecting the similarity of patent texts, patent valuation. There are studies about processing the patent data and they can be beneficial to the patent specialist. However, these works, which generally remain at the theoretical stage, have not turned into a form that the patent researcher can use in practice.

A patent recommendation system study in which an artificial intelligence-based approach is adopted instead of a method in which traditional human behaviors are imitated has not been found in the literature. In addition, it has been determined that there is a gap in the literature about the studies in which the patent text is not used in its raw form, but after the pre-processing stages.

In this study, it is aimed to achieve the following objectives:

- Obtaining the information of the closest documents in terms of the technical field and content of the invention to be patented, much faster and more consistently than the existing methods.
- The algorithms used will be compared and the models with the highest performance will be determined.
- This recommendation system is aimed to support both experts and applicants in patent research.

2. METHOD

In the study, the Red Book dataset presented by the American Patent and Trademark Office (USPTO) is used [6]. This dataset is published weekly in various categories (full text, bibliographic information, etc.). In this study, 52 patent publications of 2018 is used in the training set. The total size of this data set, which includes the full text (description, claim, abstract) of 376,609 patent

applications, is 37 GB. Since this data set is not directly usable, firstly, the parsing of the XML files is performed.

In the developed model, 3 different data types are used: description, claim, abstract. The effect of using these text types in models on performance is investigated. In order for text files to be processed, they must first be converted to digital format. Before the text is digitized, it goes through a number of pre-processing steps. In this way, it is aimed to prevent misunderstandings in the text.

2.1. Text Preprocessing

The pre-processing and vectorization steps of the patent text are presented in Figure 1. In the first stage, the text is arranged in such a way that it consisted of only uppercase and lowercase English letters. In this way, the numbers and symbols in the text have been deleted. The aim here is to simplify the text to be used in the model. Subsequently, all letters are converted to lowercase. Subsequently, all letters are converted to lowercase.

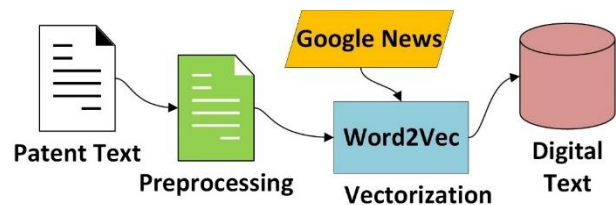


Figure 1. Vectorization stages of patent text

In the last step of the text pre-processing, the stopwords are eliminated. These words are not meaningful for the analysis of the patent document. In this elimination, words that can be found in every document such as conjunctions, articles, prepositions and which are not specific to documents have been deleted. In addition, the most common words in the data set are determined and a stopwords list is prepared. After these stages, the remaining text is decomposed into words. Word sets are created for each patent document.

2.2. Vectorization

After the pre-processing stages, the text is simplified and made suitable for vectorization. This process can be realized with architectures such as word2vec, glove, fastText, TF-IDF weighting [7,8].

In the vectorization phase of the model developed within the scope of the study, the "Google News Corpus" is used which is trained by using the Word2Vec architecture. This corpus is a 300-dimensional, pre-trained structure containing 3 million words [9]. The dimension is the number of coordinates that determine the position of the point. Thanks to this corpus structure, it is possible to skip the process that requires high hardware capacity and long training time. In this dictionary-like structure, each of the 3 million words has a numerical equivalent. In this way, the desired text can be quickly converted to vector.

During the vectorization phase of the words in 376,609 patent application documents in 2018, it was determined that approximately 3% of these words were not found in the Google News corpus. These words are generally meaningless words caused by spelling mistakes or after being cleared of symbols during text simplification process. 97% of the remaining words are vectorized. After this step, a vector table of 376,609 documents with a total of 900 dimensions (300 dimensions each for the description, abstract and claims) was obtained.

At the stage of choosing the patent classification code to be used in the model, the IPC code, which is in the first place in the patent documents, is selected. Because in American patent documents, the IPC code, which is in the first place, is in a superior position compared to the other lines and shows the technical area to which the invention is closest [10]. The first 4 digits (subclass) of the IPC code are used in model. Because, in the applications, it was observed that the training period of the model extended and the performance decreased after this IPC level.

2.3. Model

The digital text created as a result of the vectorization stages presented in Figure 1 is processed in the models indicated in Figure 2 at this stage.

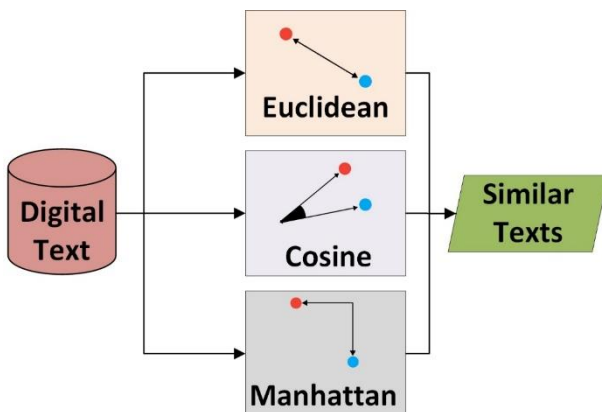


Figure 2. Model

In the developed model, an architecture is created that can measure patent text similarities. 300-dimensional vector database is used in the model. After vectorizing the test dataset with the same method as the training dataset, it was compared with all the documents in this database one by one. During this comparison, the distances between the vectors are measured. These measurements are made by three different methods used in the literature, cosine similarity and manhattan and euclidean distance calculations.

The three different measurement methods mentioned above are used in the models and their performances are compared. After the distance measurement process on all documents in the database is completed, the closest

documents are listed. Thus, a good starting position in patent research is established for the patent examiner.

3. EXPERIMENTS

In the testing phase of the created models, the data set published in the first week of 2019 is used. This set includes 8085 patent applications. As in the training phase, XML parsing is performed on this data as well.

During the vectorization phase of the patent text to be tested, 3 different models are created from 3 different texts in the data set. In the study, the effect of text lengths and types on success is investigated. Since the training time of the model will increase as the text size increases, the effect of this size increase on the performance is investigated.

The accuracy of the IPC subclass is taken as a performance metric in the model. In fact, manual accuracy control by the patent examiner is another alternative in measuring the proximity in terms of content. However, this control was not preferred due to the difficulty of evaluating more than eight thousand documents one by one and the fact that this evaluation can be subjective.

4. RESULTS

In the model, the texts (description, abstract, claim) of the patent applications published in the first week of 2019 are compared with all the applications published in 2018 one by one. During the comparison, cosine similarity values and euclidean and manhattan distance values are measured. The results obtained in line with the criteria explained in the performance criteria section are presented in Table 1. The numbers in the table indicate the test success percentages of the models.

Table 1. Text similarity model results

	Cosine	Euclidean	Manhattan
Abstract	0,731	0,732	0,732
Claims	0,754	0,754	0,751
Description	0,822	0,822	0,822

5. CONCLUSION

From the results, it is seen that the accuracy of the results increases as the text size used increases. In addition, the performances of abstract and claims models are very close to each other. This is due to the fact that applicants often copy the content of the first claim to the abstract section. It is seen that the specification model gives much better results than the other two models. The reason for this is that the text length of this model is much longer than the other two. It is also been understood that some parts of the description, which are included in the state-of-the-art sections and are not directly related to the relevant patent document, do not adversely affect the model performance.

It is seen that cosine similarity and euclidean and manhattan distance values give almost the same results. In other words, in the distance measurement of vectors, choosing any of these methods does not affect the model performance.

Patent classification code is used as a performance evaluation metric in testing the model. This preference is for the evaluation to be more objective. The fact that the obtained document is in the same technical field as the researched document shows that the content of the documents is close. However, it may not be at the level of technical detail that would be useful to the patent examiner. At this stage, manual checking by the examiner will be more efficient. The patent examiner will examine the documents obtained from the model outputs in the order of document proximity during the research process and will be able to easily take the first steps of research. This developed model will be a very useful tool in taking the first step of the research, which can be described as quite challenging in patent research.

REFERENCES

- [1] L. O. Levine, **Patent search strategies and techniques**, Pacific Northwest Lab. Report, 1986.
- [2] Internet: TÜRKPATENT Statistics, <https://www.turkpatent.gov.tr/TURKPATENT/statistics/>, 10.11.2021.
- [3] H. Wu, G. Shen, X. Lin, M. Li, B. Zhang, C. Z. Li, “Screening patents of ICT in construction using deep learning and NLP techniques”, *Eng. Constr. Archit. Manag.*, 27(8), 1891–1912, 2020.
- [4] L. Chen, S. Xu, L. Zhu, J. Zhang, X. Lei, G. Yang, “A deep learning based method for extracting semantic information from patent documents”, *Scientometrics*, 125(1), 289–312, 2020.
- [5] J. Kim, J. Yoon, E. Park, S. Choi, “Patent document clustering with deep embeddings”, *Scientometrics*, 123(2), 563–577, 2020.
- [6] Internet: USPTO, Bulk Data Storage System (BDSS), <https://bulkdata.uspto.gov/data/patent/application/redbook/fulltext/2018/>, 10.11.2021
- [7] T. Mikolov, I. Sutskever, K. Chen, G. S. Corrado, J. Dean, “Distributed Representations of Words and Phrases and their Compositionality”, *Advances in Neural Information Processing Systems*, 3111–3119, 2013.
- [8] Internet: Stanford University, GloVe, <https://nlp.stanford.edu/projects/glove/>, 10.11.2021.
- [9] Internet: Google, word2vec, <https://code.google.com/archive/p/word2vec/>, 10.11.2021.
- [10] P. Zuniga, D. Guellec, H. Dernis, M. Khan, **Patent Statistics Manual**, OECD, Paris, France, 2009.

Hair Removal and Lesion Segmentation with Fully Convolutional Neural Network and Image Processing in Images of Skin Cancer

Cihan AKYEL^{1*}, Nursal ARICI²

¹Graduate School of Informatics, Management Information Systems, Gazi University, Ankara, Turkey,

² Faculty of Applied Science, Department of Management Information Systems, Gazi University, Ankara, Turkey
cihan.akyel1@gazi.edu.tr, nursal@gazi.edu.tr

Abstract— Nowadays, one of the most common types of cancer is skin cancer. As in other types of cancer, early diagnosis is vital in the treatment process of skin cancer. Today, besides traditional methods, computer technology-based methods such as Image Processing and Deep Learning are frequently used in the diagnosis of skin cancer. The most important advantage of these methods is that they do not contain human faults during the diagnosis process. On the other hand, one of the most critical problems in using such automated methods is the inaccuracy in cancer diagnosis because the hair removal and lesion segmentation can not be performed correctly. This study presents a new FCN8 based approach for hair removal and segmentation in skin cancer images. Two datasets of International Skin Imaging Collaboration (ISIC) and PH2 datasets were used in the study. In the hair removal stage, training accuracy was %87.280, and validation accuracy was %88.310. In lesion segmentation, Training accuracy was %97.700, and validation accuracy was %92.100.

Keywords—deep learning, hair removal, skin cancer, fcn8

1. INTRODUCTION

Cancer can be defined as a disease due to the uncontrolled proliferation of cells in different organs [1]. Cancer is the second-highest cause of death. An estimated 9.6 million people died of cancer in 2018 [2]. One of the most common types of cancer in the world is skin cancer. In recent years, the number of skin cancer cases has been increasing rapidly with environmental conditions. Melanoma skin cancer is the deadliest. Approximately %75 of deaths related to skin cancer are caused by melanoma [3]. According to the American Cancer Society, the number of new melanoma cases in the USA in 2019 is 96,480, and the number of deaths is 7230. Early diagnosis is vital in the treatment of skin cancer. If the diagnosis is made early and treatment is initiated, the survival rate over five years is %92 [4]. Melanoma-type skin cancers tend to metastasize and spread to other tissues. The possibility of metastasis is high in melanoma tumors with an invasion depth of more than 2mm [5]. Traditional methods such as the ABCD rule, biopsy, and dermoscopy are used in skin cancer diagnosis. The ABCDE Rule is a common method. This method can help the dermatologist assess the skin lesion to check the risk of having malignant melanoma. Dermatologists can predict melanoma with %75 accuracy, which is low [6].

Another method that is used in the diagnosis of skin cancer is dermoscopy. Dermatologists assess dermoscopy images, which tends to be a very time-consuming and error-prone process [7]. Many serious steps such as dermoscopy and biopsy are required for the visual examination and conclusion of skin cancer. This causes a long time for the diagnosis of skin cancer. The success rate, especially during the visual inspection, can decrease significantly depending on the skill level of the dermatologist. Considering the importance of early diagnosis of skin cancer, the importance of the situation becomes apparent. The visual examination of skin cancer shows that the success rates of even the best dermatologists remain around %80 [8].

Today, image processing and deep learning algorithms are frequently used to diagnose skin cancer [9]. In detecting skin cancer with Deep Learning and Image Processing, cleaning the lesion from hair-like noise is crucial. If the hair removal is not done correctly, the success rate will decrease in the classification of the lesion [10]. It is seen that image processing features such as color spaces (such as YUV), pixel values, features, edge detection, thresholding, contrast enhancement are used in such studies [3, 11].

As Deep Learning and Image Processing Algorithms eliminate the human factor, they can provide more reliable

results than traditional methods. One of the essential advantages of these methods is that they do not need any surgical operations, and these advantages help reduce diagnosis time significantly. Surgical techniques take a long time and disturb patients. These algorithms ensure to eliminate human faults and to receive successful results as an expert. In addition, these methods reduce the cost at the diagnosis stage to almost zero and the error rates, primarily by reducing the human factor.

In the study of Talavera-Martinez et al., they created a hair dataset. Their dataset includes only simulated hair noises. They used the CNN-Based model with 512×512 input size [12]. In another study, Wei Li et al. proposed a method based on Unet for hair removal. They have prepared a hair dataset and archived about %96 dice-coef with Unet in the hair removal stage and %86.5 in segmentation [13].

When the literature is examined, deep learning and image processing methods are not alone in hair removal and segmentation stages. Using deep learning and image processing together is a more suitable solution to increase the success rate at these stages. In addition, the use of data augmentation techniques is beneficial in increasing the success rates.

This study has been carried out by taking these two approaches together. It is thought that it will contribute to the literature with this aspect. The following sections of our study are planned: The second section explains the material and method we apply to test the suggested approaches. In the third and fourth sections, the method results are shared, discussed, and concluded. In the third and fourth sections, the method results are shared, discussed, and concluded.

2. MATERIAL AND METHOD

2.1. Material

Two ISIC datasets that are taken from the ISIC challenge 2018 were used. In the first dataset, there are 10015 RGB images containing lesions in jpeg format. From this dataset, 2000 images containing hair noise (they may also have other noises) are randomly selected and increased to 8000 by data augmentation in the hair removal stage. This dataset will now be referred to as the cleaning dataset (8000 RGB images and masks created in the hair removal stage) [14]. A hardware system including a GTX1060(6GB) video card and 24GB DDR5 Ram was used in stages.

The second ISIC dataset(1300 images in jpeg format and mask in png format) and PH2 dataset (200 images and masks in bmp format) were used for lesion segmentation. This dataset was increased to 52800 by data augmentation. This dataset will now be referred to as the segmentation dataset. We chose ISIC 2018 because ISIC 2019 and ISIC 2020 do not have masks of images [15, 16]. In both stages, data augmentation was applied to increase learning success and prevent the system from falling into overfitting (90 and

180-degree rotation, horizontal and vertical flip operations).

2.2. Method

Our study consists of two stages: Hair removal and Lesion Segmentation. In both stages, the data are divided into %70 training, %20 validation, and %10 test set. These rates are acceptable values used in deep learning studies. Our study trained the system with these rates during the training phase, and there was no overfitting. For this reason, the rates that are given for training and test data were used.

FCN is chosen for the training stages. Training accuracy can be increased with other deep learning methods like Unet, LinkNet. This study shows that skin cancer diagnosis can be made with high accuracy rates with deep learning methods.

2.2.1. FCN

We used the FCN8 algorithm. FCN8 algorithm has 27 layers, imagenet weights, and input size (224×224). The optimizer was determined as adam. We used the Dice coefficient and mean squared error (MSE) loss function. The general architecture of FCN is presented in figure 1.

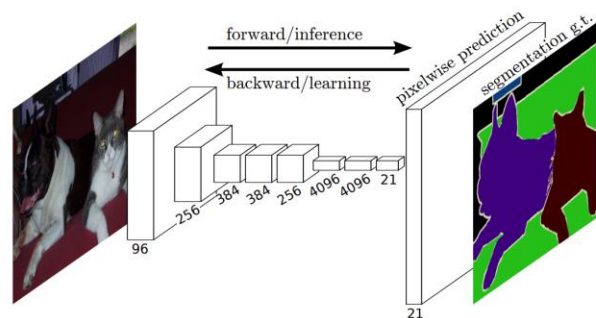


Figure 1. FCN Model [17].

Each layer of data in a convnet (Convolutional Neural Network) is a three-dimensional array of size $h \times w \times d$ (h and w are spatial dimensions, and d is the feature or channel dimension). The first layer is the image, with pixel size $h \times w$, and d color channels. Convnets are built on translation invariance. Basic components: convolution, pooling, and activation functions operate on local input regions and depend only on relative spatial coordinates. FCN uses 1×1 convolutional layers [17].

FCN8 includes upsampling operation. Additional predictions from pool3 at stride 8, providing even further precise boundaries. In FCN8, 8x upsampling is used. Thus, each pixel value is multiplied by 8 to equalize the size of the input and output images. FCN8 has more features with pool 3 than FCN16, 32. So it has the best results. It is seen in figure 2 [17].

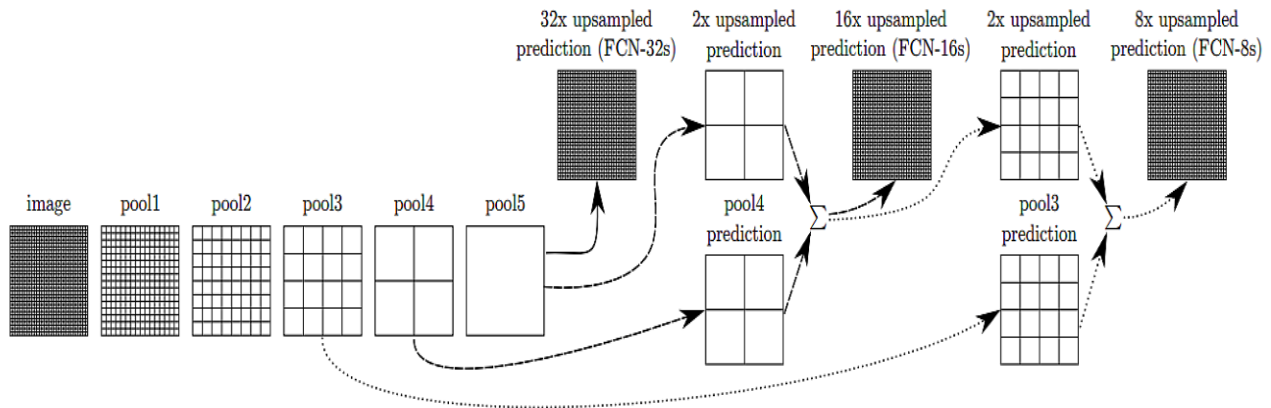


Figure 2. Comparison o FCN 8, 16, 32 [17].

Parameters that are used in training stages are given in Table 1.

Table 1. Parameters used in training stages

Parameters	Hair Removal Stage	Lesion Segmentation Step
Batch Size	8	8
Learning Rate	0.01	0.01
Epoch Number	500	500
Input Size	192*192	192*192
Optimizer	Adam	Adam

2.3. First Stage: Hair Removal Stage

The hair masks of the images in the cleaning dataset were created at this stage, and the system was trained with these images and masks. The Hair Removal Algorithm includes the following steps.

Step 1: Images were resized to 224 * 224*3 (Total 2000 RGB Images).

Step 2: The hair masks of the images were created by Adaptive Thresholding.

Step 3: In this step, tiny bubbles and fine hairs were removed from the image with the median filter. Then, morphological operations (opening and closing) were applied to remove the noise in the hair masks.

Step 4: Hair masks that emerged in the first three steps were corrected by examining the hair noise in the lesion using the Adobe fireworks program (Figure 3).

In this stage, a total of 2000 hair masks were created. Then the number of images and masks was reached 8000 with data augmentation.

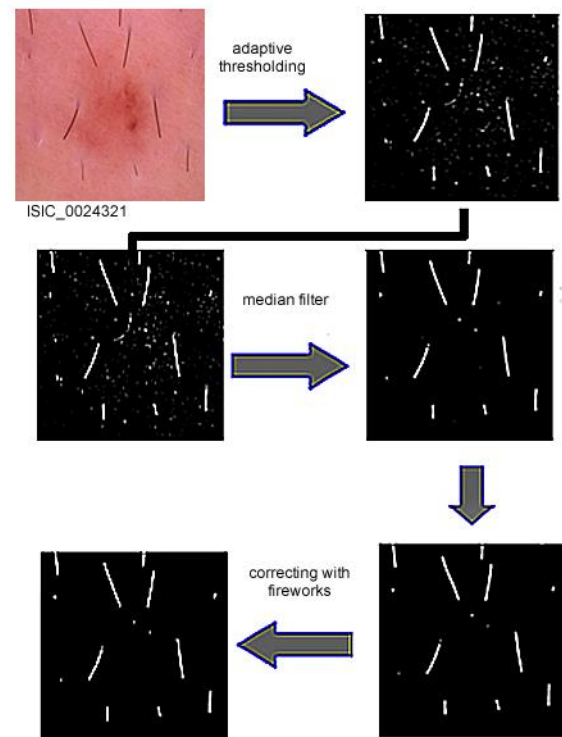


Figure 3. Adaptive threshold process.

Step 5: Cleaning datasets are divided into 5600 training, 1600 validation, and 800 test set.

Step 6: The model was trained for 500 epochs.

Step 7: Testing the model was created on test data. Figure 4 shows the sample mask and the cleaned images estimated by using the test data. Masks were cleaned with morphological operations and a median filter. Hair noise predicted mask, and INPAINT function in the image have been cleared. At this stage, the INPAINT function in the OpenCV library was used.

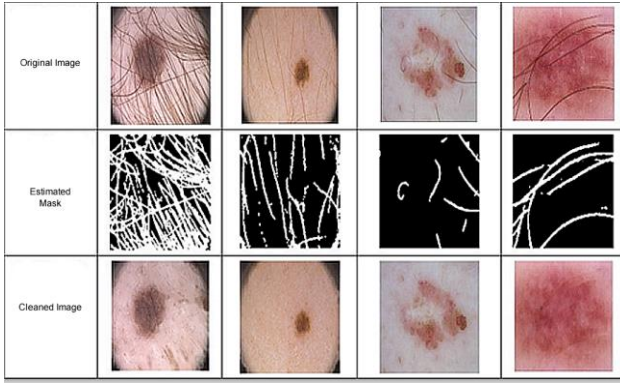


Figure 4. Results of test data.

2.4. Second Stage: Lesion Segmentation Stage

The system was trained with the FCN8 algorithm at this stage, using a segmentation dataset consisting of 52800 images and masks. As a result of the training, a success rate of approximately %92 was achieved. Hair noise in the dataset was cleared with the model created in the previous step. Images containing different noises such as band-aid were not excluded from the dataset. Lesion Segmentation Algorithm includes the following steps.

Step 1: With the model obtained in the previous step, the hair noise in the dataset was cleaned. And the images and masks have been resized to $224 * 224 * 3$ channels (RGB images).

Step 2: In this stage, the number of images was reached 52800 with data augmentation.

Step 3: Segmentation datasets are divided into 36960 training, 10560 validation, and 5280 test set.

Step 4: The model was trained for 500 epochs.

Step 5: Clearing the segmentation results from noise with morphological operations (median filter, opening, closing, dilated). The mask estimated at this stage has been optimized by image processing. The noise was cleaned on the mask with the median filter. Later, the opening and closing operations were applied, and the noises in the segmentation were cleaned. With the dilation function, area reductions due to previous operations were corrected. And tiny regions were removed. In this way, images of the lesions can be segmented more accurately.

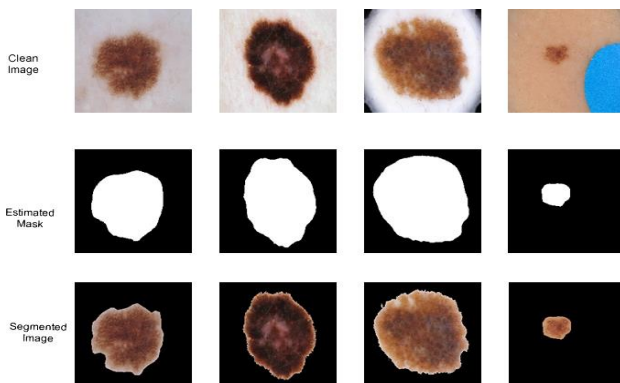


Figure 5. Lesion segmentation results of test data

3. RESULTS AND DISCUSSION

In this study, the FCN8 algorithm for Hair removal and Image Segmentation with the specified parameters obtained approximately %88 and %92 validation accuracy rates. The algorithm was run 500 epochs for two steps (Hair removal and Image Segmentation). Since there was no significant decrease in the validation loss rate after 500 epochs, training on this number of epochs was terminated.

Model results obtained in both stages can be seen in Table 2. Due to the graphics card limitation, we used input images as $192 * 192 * 3$. Using different input sizes and epoch numbers, better results can be obtained.

Parameters	Hair Removal Stage	Lesion Segmentation Step
Training Accuracy	87.280%	97.700%
Traning Loss	12.765%	10.220%
Validation Accuracy	88.310%	92.100%
Validation Loss	7.720%	6.15%

In the literature research, it was seen that thresholding and image processing algorithms were used in many studies in hair removal. The common problem encountered in these studies is that the hair mask can not be estimated correctly, and data loss occurs due to hair removal. Hair noise affects lesion segmentation accuracy. Figure 6 shows that when the mask is estimated with the model without hair removal, the success decreases.

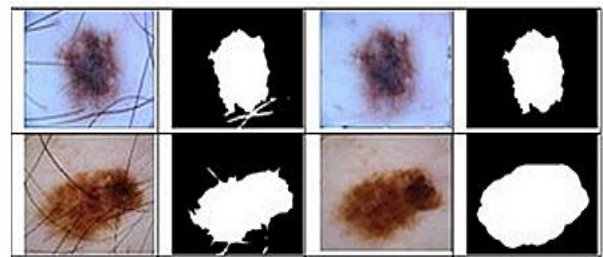


Figure 6. The Effect of hair noise in segmentation

In literature, some studies show FCN8 has higher accuracy than UNET in hair removal and segmentation [18, 19]. In the 1st stage, noises such as band-aid, pencil marks, black frames are removed from the lesion. In the 2nd stage, more accurate results were obtained by re-segmentation with image processing on the noiseless image used as input. In figure 7, the comparison is given between our model and document [20]. In the mentioned study, %96 success rate was achieved. These images are complex examples of segmentation. Because they have a low contrast transition

and locally high brightness. It is seen that our model gives more successful results.







Model	Image number	Original image	Ground truth mask	Method result
Other model	IMD088			
Our model	IMD088			

Figure 7. Comparison of models

In the study of Talavera-Martinez et al., similar to our work, they used a hair dataset. Their dataset includes only simulated hair noises. It is easy to find noises in these images. But in reality, images may contain more different hair noises (different color, contrast, etc.). As a result, we don't know how their model will work at real hair noises [12]. Also, the Study of Li et al. is similar to our study. They have created hair mask datasets (306 images and masks in ISIC 2018). They used a small hair dataset [13].

Different from the current literature in hair removal and segmentation stages, we presented new approaches. In the hair removal stage, a dataset containing hair masks was created. When the literature is examined, it is seen that there is no such like our dataset. We used both deep learning and image processing. In this way, better elimination of different hair noises has been provided. It is seen that the new approach used in the hair removal phase gives successful results. Using this dataset, training was completed with deep learning. In this stage, training accuracy was %87.280, and validation accuracy was %88.310. In the segmentation stage, training accuracy was %97.700, and validation accuracy was %92.100. Images with different noises such as band-aid were not excluded from the dataset. Although this situation caused a slight decrease in the success rate, the sensitivity of the model increased.

REFERENCES

- [1] O. Baykara, "Current Modalities in Treatment of Cancer", *Bahkesir Health Sciences Journal*, 5(3), 154-165, 2016.
- [2] Internet: WHO (2020). Cancer, <https://www.who.int/news-room/fact-sheets/detail/cancer>, 15.10.2021.
- [3] Z. Faisal, N. Abbadi, "New Segmentation Method for Skin Cancer Lesions," *Journal of Engineering and Applied Sciences*, 12(21), 5598-5602, 2017.
- [4] H. M. Unver, E. Ayan, "Skin Lesion Segmentation in Dermoscopic Images with Combination of YOLO and GrabCut Algorithm", *Diagnostics Journal*, 9(72), 1-21, 2019.
- [5] K. H. Güngör, Dkykő, Tıpta Uzmanlık Tezi, Ankara Üniversitesi Tıp Fakültesi, Ankara, Turkey, 2016.
- [6] F. Nowroozipour, **Skin Lesion Segmentation Techniques For Melanoma Diagnosis: Comparative Studies**, Master degree these, Toronto, Ontario, Canada, 2016.
- [7] Internet: Skin Cancer Segmentation and Classification with NABLA-N and Inception Recurrent Residual Convolutional Networks, <https://arxiv.org/ftp/arxiv/papers/1904/1904.11126.pdf>, 25.02.2021.
- [8] M. A. Kadampur, S. A. Riyae, "Skin cancer detection: Applying a deep learning based model driven architecture in the cloud for classifying dermal cell images," *Informatics in Medicine Unlocked Journal*, 18, 1-6, 2020.
- [9] M. Senan, M. Jadhav, "Classification of Dermoscopy Images for Early Detection of Skin Cancer – A Review," *International Journal of Computer Applications*, 178(17), 2019.
- [10] Internet: Skin Cancer Diagnosis, <https://www.sciencedirect.com/science/article/pii/S1877050916305865>, 11.05.2021.
- [11] S. Jain, V. Jagtap, N. Pise, "Computer-aided Melanoma skin cancer detection using Image Processing," **International Conference on Intelligent Computing, Communication & Convergence**, Bhubaneswar, Odisha, India, . 735-740, 2015.
- [12] Internet: L. Talavera-Martínez, P. Bibiloni and M. González-Hidalgo, *IEEE*, 26.10.2021.
- [13] L. Wei, N.J.R. Alex, T. Tardi, Z. Zhemin, "Digital hair removal by deep learning for skin lesion segmentation," *Pattern Recognition*, 117, 1-15, 2021.
- [14] Internet: Task 3: Lesion Diagnosis: Training, <https://challenge2018.isicarchive.com/task3/training/>, 20.10.2019.
- [15] Internet: ISIC 2018: Skin Lesion Analysis Towards Melanoma Detection, <https://challenge2018.isic-archive.com/>, 15.10.2019.
- [16] Internet: PH2 Dataset, <https://www.fc.up.pt/addi/ph2%20database.html>, 15.10.2021.
- [17] Internet: Fully Convolutional Networks for Semantic Segmentation, <https://arxiv.org/pdf/1411.4038.pdf>, 10.05.2021, 15.10.2021.
- [18] S. Zafar, S. Gilani, A. Waris, A. Ahmed, M. Jamil at al. "Skin Lesion Segmentation from Dermoscopic Images Using Convolutional Neural Network", *Sensors Journal*, 20(6), 2020.
- [19] K. Hasan, L. Dahal, P. N. Samarakoon, F. I. Tushara, R. Marti, "DSNet: Automatic Dermoscopic Skin Lesion Segmentation," *Computers in biology and medicine*, 120, 2020.
- [20] M. Abdulhamid, M. Şahiner, J. Rahebi, "New Auxiliary Funtion with Properties in Nonsmooth Global Optimization fir Melonoma Segmentation", *Hindawi Biomed Research International*, 1, 2020.

The Importance and Challenges of Twenty-First Century Skills for the Information and Communication Sector

Isra' Ala'Addin DWIEK¹, Hasan ÇAKIR²

¹Information System, Gazi University, Ankara, Turkey

² Computer Education and Instructional Technology, Gazi University, Ankara, Turkey
asraa.dweik@gazi.edu.tr, hasanc@gazi.edu.tr

Abstract— This review article aims to draw attention to twenty-first century skills and the determinants that influence their availability among students and workers. Several frameworks have been mentioned that have worked to define the skills of the twenty-first century, including (technology, information, communication, collaboration, critical thinking, creativity, and problem-solving skills). We also mentioned the importance of acquiring these skills for people, as well as the challenges we face in acquiring 21st century skills.

Keywords— 21st century skills, competencies, information and communication (ICT) sector, students, workers, challenges.

1. INTRODUCTION

The twenty-first century differs greatly from the twentieth century in terms of the competencies and knowledge needed to continue competing at the international level for nations and individuals to be successful in their personal, social, political and economic life [5]. There's been a shift in the economy from an industrial to a knowledge economy. This transformation has imposed a set of requirements and competencies that individuals must acquire through education systems [27, 45]. It has been stated that ICT has created a new set of competencies and knowledge required by teachers and employers [27]. In research, the terms "skills" and "competencies" are used. Competency may differ from skill; Competence is more than just knowledge or skill, as it involves meeting complex demands through reliance on psychosocial resources. Skill is the ability to perform precise tasks to solve problems. However, competence is seen as the ability to appropriately apply knowledge in a specific context (education, work, personal or professional development). Competences are not limited to knowledge components involving concepts, the use of theory procedural knowledge, or implicit knowledge. They also include aspects of career including technical skills as well as personal traits such as social skills or organizational skills, and ethical values [36].

Educators, employers, and organizations agree that 21st century skills are important and necessary to meet the demands of employment, integration, and successful participation in today's society, and to contribute positively to the well-being of nations' economies [41]. According to the National Association of Colleges and Employers' Survey of Job Outlook (2015), more than 80% of employers are looking for leadership skills in employees, they also want written and verbal communication skills, problem-solving skills, and a strong work ethic [31, 32]. The development of digital skills is also seen as necessary to achieve economic growth, digital transformation and increase the welfare of citizens in the world [28].

A literature review is conducted to collect the studies in the English language concerned with the availability of twenty-first century skills among university students and on workers, and is there a relationship between gender, educational level and other variables, and individuals' possession of twenty-first century skills. We expect that this article will practically help designers, developers of technology programmes, course designers, and researchers in identifying shortcomings in current programs as well as in training courses, and work to avoid them, and strengthen the strengths. It also helps teachers reorganize course content and its incorporation of 21st century skills.

2. THEORETICAL REVIEW

The Fourth Industrial Revolution is characterized by the emergence of several technologies and scientific engines that began to change the face of the world, including artificial intelligence techniques, the Internet of things, genetics (the human genome), 3D printing, quantum computer, nanotechnology, biotechnology, self-driving cars, big data, virtual currencies, and energy storage [49]. This revolution differs from previous revolutions in several aspects represented in the speed of spread of the technologies accompanying this revolution (exponential growth versus linear growth in the speed of spread of technologies associated with previous revolutions) [49]. Technology has come to replace manual labour and is being instilled in most aspects of people's life and work. Jobs involving routine manual and cognitive skills has seen a steady decline over the past 50 years, whereas a corresponding increase in jobs that require non-routine and subjective analysis has been seen. Several factors have caused to these trends such as accelerated automation and digitization of work that can be done in routine procedures [9, 10, 24, 29].

Currently, information, which is an essential economic resource that has its own properties, is a strategic resource in the new economic life, and information technology in our time represents the main element in economic growth. The transition to the knowledge economy requires the development of the local community by achieving high levels of knowledge, efficiency, and technical skill. Which leads to a change in the role of the teacher, the learner, and methods of learning, the educational environment, its equipment, as well as the educational curriculum [44]. In order to thrive in today's innovation industry, employees need a different skill set than they did in the past [7]. In addition to basic skills such as reading, writing and arithmetic, they need skills such as creativity, collaboration, and problem-solving as well as having personal qualities such as perseverance, curiosity and initiative [9, 20].

2.1. 21st Century Skills

The concept of twenty-first century skills may not be new. However, the content of these skills varies profoundly in technology-rich environments. The term twenty-first century skills were first used to list the general set of skills required for workers to thrive in the current job market [26]. Twenty-first century skills appeared in 2002, which aims to support students in career and university life in terms of mastering both content and skills, and advocacy for these skills began in all disciplines through the Partnership for 21st Century skills. This partnership has now become one of the most important leaders in the development and teaching of twenty-first century skills in

the world [8]. The twenty-first century skills framework for the twenty-first century skills partnership contains three skills, including: learning and innovation skills, which include critical thinking and problem solving, communications and collaboration, creativity and innovation, Digital literacy skills including information literacy, media literacy, Information and communication technologies (ICT) literacy, and finally life and career skills and it contains flexibility and adaptability, initiative and self-direction, social and cross-cultural interaction, productivity and accountability [8]. Besides, the North Central Regional Educational Laboratory's (NCREL) has classified twenty-first century skills into four sections including (Digital-Age Literacy, Inventive Thinking, Effective Communication, High Productivity) [1, 11, 25]. The classification of the United Nations Educational Scientific and Cultural Organization divided the skills of the twenty-first century into four sections (learning to know, learning to do, learning to live with others, and learning to be) [14, 46]. The classification of the Twenty-First Century Skills Assessment and Teaching of 21st Century Skills Project, (ATC21S) included four skills (methods of thinking, ways of working, tools of work, skills for living in the world) [14]. The Organization for Economic Co-operation and Development has classified twenty-first century skills as information, communication, ethics, and social impact [37].

3. METHOD

Twenty-first century skills are increasingly gaining traction in both academia and the professional world [12]. The current study aims to enhance our understanding of the availability of twenty-first century skills for students and workers in the information and communication sector. Document review and content analysis methods were used as the primary methodological method for this study.

3.1. Criteria for Inclusion

The studies reviewed in this study were selected based on several criteria, including the year of publication, and were conducted on high school students, university students in information and communication majors, and workers in information and communication sector as the main focus of the study. Articles published from 2016 to this year are included. Previous studies in English and Turkish are included. The participants are students in various technological disciplines and workers in information and communication sector. Only studies of twenty-first century skills or one of the twenty-first century skills were included. The article focused only on information and communication sector. Because we live in an evolving and constantly changing era that requires follow-up and renewal in people's skills constantly, so the author chose to start from 2016 in choosing studies that can promise

modern studies. Most of the articles examined the availability of twenty-first century skills for students of information and communication majors and for those working in information and communication sector or one of the skills, which are considered independent variables. These studies were included in this review.

3.2. Data Source

The existence of twenty-first century skills has attracted the interest of researchers in applying them as well as educational program developers in their studies. Many studies have been conducted on the extent to which it is possessed by students and workers. The scope of the current study is to understand the availability of skills and the importance of having them in school curricula and in jobs. To achieve the goal of the study, the studies published from 2016 to the present day were searched through the use of "21st century skills", "students", "employees", "information and communication sector", "information and communication disciplines", "Skills", "competencies", "soft". The articles published in English and Turkish were selected.

By searching for the year of publication and the above-mentioned keywords, "16" studies were obtained. The researcher classified into 3 classifications, the first for school students, the second for university students who study in technological disciplines, and the third for workers in information and communication sector. In addition, the studies were examined based on their objectives and variables to determine whether they are closely related to possessing twenty-first century skills for both students and workers. Studies that included students studying in other disciplines, as well as workers working in any sector different from information and communication sector, were excluded. The table (1) shows the studies reviewed.

4. THE IMPORTANCE OF ACQUIRING 21ST CENTURY SKILLS

Without the right skills, people cannot make contribution to the society, technological progress will not evolve into welfare, and countries cannot compete in the global economy [6, 40, 45]. So those responsible for education must shape education systems in this situation so that they can give students twenty-first century skills, which not only enable them to acquire knowledge, but also give them the ability to produce and apply knowledge in different aspects of life [17]. For decades, students thought that basic content classes were not relevant to the real world because they weren't asked to apply these concepts in real-world situations and skills were not entangled. As a result, the skills of the twenty-first century must be integrated into the classroom so that teachers can continue to prepare students to be successful after graduation from the educational institution [43] thus their success in college and in their career and also when they enter the labor market [21].

People must be able to collaborate, think critically, communicate orally and in writing. For now, members of the current workforce must be familiar with technology, be able to constantly adapt and reinvent themselves [33].

Companies are looking to recruit people with twenty-first century skills to increase profitability, keep pace with rapid technological developments, reduce the cost of recruiting and training new staff, and focus more on the global economy [6]. Labour market skills are increasingly linked to current jobs and employers expect their employees to excel in applying these skills in their work for the company's success [22, 38]. [48] explained that there is a lack of readiness among university graduates regarding the lack of twenty-first century skills. In addition, discovered a gap between the workforce and learning environments, explained that there is a difference between what students learn at school and what is expected of them in the workplace. Business managers have identified necessary skills not found in the new workplace, such as critical thinking, problem solving, creative thinking, collaboration with others, and professional and effective communication among employees [21, 47].

5. CHALLENGES FACING EDUCATION IN THE TWENTY-FIRST CENTURY

Acquiring skills develops individuals' awareness of what is going on around them, and qualifies them to enter the world of knowledge, encourages them to learn of all kinds, and provides them with the opportunity to connect with different learning sources. But there are challenges facing education in the twenty-first century, including [34]:

5.1. The Labor Market

The stereotyping and traditionalism of education made it unable to keep pace with the developments of the twenty-first century, which made it suffer from crises such as the traditional curricula, the outputs that are not suitable for the labor market.

5.2. Scientific and Technological Progress and the Information and Communication Revolution [13]

The huge amount of knowledge and information in science and technology imposed many difficulties in the field of the quality of knowledge provided to learners, through the contents of the curricula and courses, and the imposition of new trends in the study systems, and the use of technologies that did not exist before, but the matter became more difficult for educational institutions because of difficulties anticipate and prepare for change. Therefore, changing the goals and philosophies of education systems has become imperative to meet these future challenges.

5.3. Social and Cultural Challenges

These changes are directly linked to, influence and affected by population, technology and environmental factors. Social change - in contemporary times and in the future - is the product of the technological revolution, and this means a change in values and standards, social relations, behavioral patterns, and others. Here comes the role of education in “cultural development, which is seen today and, in the future, as a means for comprehensive development and as an end for it at the same time. Without it, it is difficult for economic, social and scientific development to reach its full goals.” Where cultural development is based on education, education, science, and technology.

5.4. Student Needs

Globalization and technological change combine to have a tremendous impact on students in their schools today. Global economic opportunities and the need to acquire twenty-first century skills have raised academic requirements, and the prevailing model of human capital is no longer satisfied with prolonging school years to prepare students for industrial jobs. What is required today is a mastery of the content and a deep understanding of the discipline, in addition to, the acquisition of soft skills represented in communication, initiative, adaptability and abilities related to finding solutions to the issues at hand.

5.5. New Expectations from School Education

Expectations of school education vary across national contexts within each country. The voices and demands of key partners, employers and parents are all rising high, demanding what they expect schools to develop in students, and this is of course not a separate development from the challenges arising from economic globalization. In advanced economies such as the United States, the United Kingdom, Canada and Australia, there are concerns in addressing the shortage of literacy and numeracy skills, with a high shortage of science and mathematics graduates needed for technological development.

6. LITERATURE REVIEW OF ICT SECTOR

Many studies have been conducted on school and university students and workers in various ICT professions and occupations. This study aimed to examine the availability of all twenty-first century skills or one skill among the twenty-first century skills to student and workers in ICT sector, as well as the availability of these skills with other factors in previous literary studies. [50] found that the skills of the twenty-first century do not differ significantly according to the gender variable, while problem-solving skills, critical thinking skills, information and technology literacy skills, entrepreneurship and innovation skills differ according to the gender of the

participants, as well as according to the levels of students for their classes. [50] also found that there is no relationship between job awareness, social responsibility, and leadership skills according to grade level. Although there is a positive relationship between students' skills in the twenty-first century and their educational attainment. They found [15] that students who participated in the certification program had a high level of skill in terms of 21st century skills. Additionally, there was no effect of gender for students' participation in the technology-based certification program. A study [35] found that there is no difference between students studying in the Anatolian Vocational Training Program and those studying in the Anatolian Technical Program according to their computer thinking skills. However, when calculating the average scores, it was observed that the students of the Technical Anatolia Program had higher scores in terms of total scores and they had a higher level in sub-skills compared to the Anatolian Vocational Students. [23] found that an individual's cognitive skills (how to think, solve problems and learn) have a greater impact on a person's ability to function in a technically rich society than simply knowing specific software packages or equipment. In research [4] it was found that computer science students showed a clear development in media literacy skills. [39] found that students in general have a moderate knowledge of digital literacy according to the concept of (IC3), which includes online living, the main applications, and the basics of computing. The study also emphasized the need to educate students about digital literacy in order to prepare themselves to enter the digital society and career path with their own ability. [42] found that computer science students perform much better in solving problems than students in any other discipline. The results also showed that only computer science students had an indication of improvement in problem solving from lower to higher levels.

The results of the study [30] showed an improvement in students' critical thinking in academic writing after they joined an academic writing class. The researcher [28] also found that this method can help students hone their 21st century skills such as collaboration and communication. The results of the study [12] showed that communication skills are the most requested skill for young graduates, followed by teamwork and problem solving. The researcher [16] conducted a study entitled “Graduate Candidates' 21st Century Skills and Challenge for the Faculty and the University in Online Learning Era”. The results showed that students in the College of Science and Technology rated their overall skill in the twenty-first century at an average level. Students also assessed soft skills; Communication skills, IT skills, Arithmetic, Learn to learn, Problem solving skills, Intermediate level teamwork skills. These findings suggest that science technology students were not adequately prepared for twenty-first century skills; This in turn has led to the faculty and university facing a difficult challenge to reform the learning practices process, and such a policy must be taken to give more emphasis on twenty-first skills in the curriculum and teaching and learning practices. The

researchers did not find [2] meaningful relationship between the education of students studying in the ICT department, informational leadership qualifications, and twenty-first century skills. The results of the researcher's study [48] showed that industries have the most expected competencies in the category of professional skills and information technology knowledge, followed by the information technology category and the information technology management and support category. It also turns out that the skills required are lifelong learning, personal attitude, teamwork and dependability, IT foundations, however there are slightly different requirements between IT industry and IT in non-IT industries. They concluded [19] that the problem-solving skills of the elderly and females were lower, because higher competencies in reading, writing and arithmetic lead to an increase in problem-solving skills. Also, people with more experience in exchanging work-related information have higher skills in solving problems. While employees who make presentations have lower problem-solving skills, those who frequently use ICT especially in sending and receiving email or conducting real-time discussions have a higher ability to solve problems. They concluded [3] that the abilities differ from one group to another. Where leadership, strategy, planning and negotiation capabilities are most important to the professional profile of working supervisors and managers. While problem solving and teamwork skills are paramount to the operational core. The results of the study [27] showed that the most important skills for new IT professionals are personal skills, specifically personal characteristics of problem solving, critical thinking and team skills. Additionally, the study found that technical skills were important, particularly database knowledge, proficiency, knowledge of programming languages, object-oriented knowledge, and web development skills. The study also showed skills that are not considered very important, especially software specialized for applications such as decision support systems, expert systems, geographic information systems and enterprise resource planning systems. It was also found that there are some important differences in the importance of some skills depending on the age of the respondent, gender, years in the field of information technology, and administrative level. In addition to a significant difference in importance between males and females for fourteen of the 42 skills in this study, especially in competencies (females considered all skills except one more important than males), indicating a gender influence. [18] The results of the study showed that the required skills are divided into two parts: The first section is hard skills such as project management, technical support, and requirements management. It was found that these skills are in great demand. While the second section is the personal skills related to the team (such as teamwork and team player), communication and problem solving, which are among the most important interpersonal skills required of employers. There is a growing demand for professionals in emerging fields such as data analytics, cyber security, cloud computing, and programming languages such as SQL, Python, which support data management and analytics.

7. CONCLUSION

The aim of this article is to review research studies that examined the availability of twenty-first century skills for both students of information and communication majors and those working in information and communication sector. According to the Partnership for 21st Century skills, the 21st century skills were divided into "learning and innovation skills, IT and media skills, life skills". Participants in the studies were students studying in information and communication disciplines and workers in information and communication sector. The studies included the skills of the twenty-first century, including (learning and innovation skills - Digital literacy skills, life and career skills / Digital-Age Literacy, Inventive Thinking, Effective Communication, High Productivity/methods of thinking, ways of working, tools of work, skills for living in the world / learning to know, learning to do, learning to live with others, and learning to be). Recent research studies from 2016 to the present day were reviewed. It was found through research studies that the subject of twentieth century skills is one of the important topics in scientific research. Of the research studies that found no effect of gender among participants in the extent to which they possess skills, and others that found an effect of gender in the availability of some skills, some studies showed that IT students have a high level of skills. This indicates that the literature needs more studies to know the extent to which the skills of the twenty-first century are possessed by workers, especially in information and communication sector.

REFERENCES

- [1] A. Arslan, "Öğretmen Adayları Perspektifinden Pandemi Öncesi ve Sonrası Öğrencilere Kazandırılması Gereken 21. Yüzyıl Becerilerinin Belirlenmesi", *Millî Eğitim*, 49(1), 553–571, 2020.
- [2] A. Korucu, M. Ünüvar, "Bilgisayar ve Öğretim Teknolojileri Öğretmenliği Bölümünde Okuyan Öğretmen Adaylarının Aldıkları Eğitim ile Bilişim Liderliği Vasfı ve 21. Yüzyıl Becerileri Arasındaki İlişki", *Bilgi ve İletişim Teknolojileri Dergisi*, 2(1), 45–54, 2020.
- [3] A. Llorens-Garcia, X. Llinas-Audet, F. Sabate, "Professional and interpersonal skills for ICT specialists", *IT professional*, 11(6), 23–30, 2009.
- [4] A. Vasilchenko, D.P. Green, Q. Haneen, P. Anne, B. Tom, and B. Madeline, "Media literacy as a by-product of collaborative video production by CS students", In **Proceedings of the 2017 ACM conference on innovation and technology in computer science education**, 58–63, 2017.
- [5] A. Velez, **Preparing students for the future—21st century skills**, Doctor's thesis, Faculty of The USC Rosier School of Education East Carolina University, 2012.

- [6] A. Schleicher, "Developing Twenty-First-Century Skills for Future Jobs and Societies", **Anticipating and Preparing for Emerging Skills and Jobs**, Editör: Brajesh Panth, Rupert Maclean., Springer, Singapore, 47-55, 2020.
- [7] B. E. Penprase, "The Fourth Industrial Revolution and Higher Education. In N. W. Gleason (Ed.)", **Higher Education in The Era of The Fourth Industrial Revolution**, Editör: Nancy W. Gleason, Singapore, Singapore, 10, 207–229, 2018.
- [8] B. Freilandika, **The Development of 2013 Curriculum Textbook Based On 21st Century Skills İn Senior High School**, Master's Thesis, Department of Arabic Education, 2020.
- [9] Boston Consulting Group (USA), World Economic Forum, **New vision for education: Unlocking the potential of technology**, 2015.
- [10] C. B. Frey, M. A. Osborne, "The future of employment: How susceptible are jobs to computerization?", *Technological Forecasting and Social Change*, 114, 254-280, 2017.
- [11] C. Lemke, "Engage 21st Century Skills: Digital Literacies for a Digital Age", 2002.
- [12] D. McGunagle, L. Zizka, "Meeting Real World Demands of the Global Economy: An Employer's Perspective", *Journal of Aviation/Aerospace Education & Research*, 2018.
- [13] F. Youssef, "Modern teacher and the challenges of the twenty-first century", *Journal of Educational and Psychological Sciences*, 4(19), 110-130, 2020.
- [14] Global partnership for education, **21st century skills: what potential role for the global partnership for education? A landscape review**, 2020.
- [15] H. Çakır, O. Delialioğlu, E. Özgül," Evaluation of information technology certificate programs in terms of 21st century skills: Bilgi teknolojileri sertifika eğitim programlarının 21. Yüzyıl becerileri açısından değerlendirilmesi", *Journal of New Results in Science*, 16, 998-1013, 2019.
- [16] H. Hadiyanto, N. Noferdiman, S. Syamsurizal, M. Muhaimin, Y. Yuliusman, S. Ramli, U. Sulistyono," Graduate Candidates' 21st Century Skills and Challenge for the Faculty and the University in Online Learning Era", *The Turkish Online Journal of Educational Technology*, 201-216, 2018.
- [17] J. Brennan, R. King, Y. Lebeau, **The role of universities in the transformation of societies**, Synthesis Report, Centre for Higher Education Research and Information/Association of Commonwealth Universities, UK, 2004.
- [18] J. Nwokeji, R. Stachel, T. Holmes, R.Orji," Competencies Required for Developing Computer and Information Systems Curriculum", **In 2019 IEEE Frontiers in Education Conference (FIE)**, IEEE Press, 1–9, 2019.
- [19] J. Yoon, E. Hur, M. Kim, "An Analysis of the Factors on the Problem-Solving Competencies of Engineering Employees in Korea", *Sustainability*, 2020.
- [20] J.W. Rojewski, & Hill, R. B. Hill, "A framework for 21st century career-technical and workforce education curricula", *Peabody Journal of Education*, 92(2), 180–191, 2017.
- [21] Kh. Alshare, M. Sewailem, "A Gap Analysis of Business Students Skills In the 21st Century A Case Study of Qatar", *Academy of Educational Leadership Journal*, 2018.
- [22] K. Robbins, **21st century skills: student readiness, opportunities, and development**, Master's Thesis, Faculty of the Department of Human Development and Family Science, 2017.
- [23] M. Ahmad, A. Karim, R. Din, I. S. M. A. Albakri, "Assessing ICT competencies among postgraduate students based on the 21st century ICT competency model", *Asian Social Science*, 9(16), 32, 2013.
- [24] M. Arntz, T. Gregory, U. Zierahn, **The risk of automation for jobs in OECD countries**, OECD Social, Employment and Migration Working Papers, No. 189. Paris, France: OECD Publishing, 2016.
- [25] M. Çevik, C. Şentürk, "Multidimensional 21st century skills scale: Validity and reliability study", *Cypriot Journal of Educational Science*, 14(1), 11-28, 2019.
- [26] M. Davidekova, J. Hvorecký, "ICT Collaboration Tools for Virtual Teams in Terms of the SECI Model", *Int. J. Eng. Pedagog*, 7(1): 95-116, 2017.
- [27] M.E. McMurtrey, J. P. Downey, S. M. Zeltmann, W. H. Friedman, "Critical skill sets of entry-level IT professionals: An empirical examination of perceptions from field personnel", *Journal of Information Technology Education: Research*, 101-120, 2008.
- [28] MEB, **Bilgisayar Bilimi Dersi Öğretim Programı Kur1-Kur2**, Millî Eğitim Bakanlığı Ortaöğretim Genel Müdürlüğü, Ankara, 2016.
- [29] M. Falk, F. Biagi, "Relative demand for highly skilled workers and use of different ICT technologies", *Applied Economics*, 49(9), 903-914, 2017.
- [30] M.R. Putri, S. T. Sela, "Capturing Students' Critical Thinking Skills in English for Academic Writing Course: A Case in Information System Students", **In Proceedings of the 2018 2nd International Conference on Education and E-Learning**, 146-149, 2018.
- [31] Internet: National Association of Colleges and Employers, Job outlook 2016, 2015, <http://www.naceweb.org>, 10.05.2021.
- [32] Internet: National Association of Colleges and Employers, Job outlook 2019,2016, <https://www.odu.edu/content/dam/odu/offices/cmc/docs/nace/2019-nace-job-outlook-survey.pdf>.
- [33] N.K. Hixson, J. Ravitz, A. Whisman, "Extended Professional Development in Project-Based Learning: Impacts on 21st Century Skills Teaching and Student Achievement", *West Virginia Department of Education*, 2012.
- [34] N. Subhi, "The extent to which the skills of the twenty-first century are included in the developed science curriculum for the first intermediate grade in the Kingdom of Saudi Arabia", *Journal of Educational Sciences*, 2016.
- [35] N. Yalçın, A. İkinci," Meslek Liseleri Bilişim Teknolojileri Alan Öğrencilerinin Bilgisayarca Düşünme Beceri Düzeylerinin Eğitim Program Türüne Göre İncelenmesi", *Abant İzzet Baysal Üniversitesi Eğitim Fakültesi Dergisi*, 20(4), 1639-1656, 2020.
- [36] R. Al-Hariri, "Twenty first century skills", *International Journal of Pedagogical Innovations*, 8(1), 70-91, 2020.
- [37] R. Cansoy, "Uluslararası Çerçvelere Göre 21. Yüzyıl Becerileri ve Eğitim Sisteminde Kazandırılması", *İnsan ve Toplum Bilimleri Araştırmaları Dergisi*, 7(4), 3112-3134, 2018.

- [38] S. Boyaci, N. Atalay, "A Scale Development for 21st Century Skills of Primary School Students: A Validity and Reliability Study", *International Journal of Instruction*, 9(1), 133-148, 2016.
- [39] S. Daungtod, "A study of digital literacy of 1st year computer education students faculty of education Nakhon Phanom University" **In Proceedings of the 10th International Conference on E-Education, E-Business, E-Management and E-Learning**, 241-244, 2019.
- [40] S. Eryilmaz, Ç. Uluyol, "21. Yüzyıl Becerileri Işığında FATİH Projesi Değerlendirmesi", *Gazi Eğitim Fakültesi Dergisi*, 35(2), 209-229, 2015.
- [41] OECD, **Lessons from PISA for the United States**, Strong Performers and Successful Reformers in Education, OECD Publishing, 2011.
- [42] S. Salehi, K. D. Wang, R. Toorawa, C. Wieman, "Can Majoring in Computer Science Improve General Problem-solving Skills?" **In Proceedings of the 51st ACM Technical Symposium on Computer Science Education**, 156-161, 2020.
- [43] S. Yalçın "21. Yüzyıl becerileri ve bu becerilerin ölçülmesinde kullanılan araçlar ve yaklaşımlar", *Journal of Faculty of Educational Sciences*, 51(1), 183-201, 2018.
- [44] T. Al-Qudah, F. Al-Abidin, S. Al-Anazi, M. Al-Anabtawi, "The social factors affecting the selection of academic specialization and its impact on the attitudes of university students towards their specialization", *IUG Journal of Humanities Research*, 27(2), 240-262, 2019.
- [45] T. Varis, "New technologies and innovation in higher education and regional development. RUSC", *Universities and Knowledge Society Journal*, 4(2), 2007.
- [46] UNESCO, **Arab region outcome statement, Towards inclusive and equitable quality learning opportunities for all**, 2018.
- [47] V.L. Bassols, **ICT skills and employment**, 2002.
- [48] V.Siddoo, J. Sawattawee, W. Janchai, P. Yodmongkol, "Exploring the competency gap of IT students in Thailand: The employers' view of an effective workforce", *Journal of Technical Education and Training*, 9(2), 2017.
- [49] Internet: World Economic Forum (WEF), The Fourth Industrial Revolution: what it means, how to respond, 2019,2016, <https://www.weforum.org/agenda/2016/01/the-fourth-industrial-revolution-what-it-means-and-how-to-respond/>, 6.10.2021.
- [50] Y.S. Göktepe, "Lise Öğrencilerinin 21. Yüzyıl Becerilerinin Bazı Demografik Değişkenler Açısından İncelenmesi", *Ulakbilge Sosyal Bilimler Dergisi*, 884-897, 202

Appendix-1

Table 1. List of reviewed studies

	Author	year	Title
1.	Y. S. GÖKTEPE	2020	Lise Öğrencilerinin 21.Yüzyıl Becerilerinin Bazı Demografik Değişkenler Açısından İncelenmesi
2.	H. ÇAKIR, Ö. Delialioğlu, & E. Özgül,	2019	Evaluation of information technology certificate programs in terms of 21st century skills
3.	N. YALÇIN, & A. İKİNCİ,	2020	Meslek Liseleri Bilişim Teknolojileri Alan Öğrencilerinin Bilgisayarca Düşünme Beceri Düzeylerinin Eğitim Program Türüne Göre İncelenmesi
4.	M.G. Rafique, & H.A.Khan,	2018	" Information Literacy Skills of Management Sciences' Students
5.	A.Vasilchenko, , D.P.Green, H.Qarabash, A.Preston, T.Bartindale , & M.Balaam,	2017	Media Literacy as a By-Product of Collaborative Video Production by CS Students
6.	S.Daungtod,	2019	A study of digital literacy of 1st year computer education students faculty of education Nakhon Phanom University
7.	S.Salehi, K.D.Wang, R.Toorawa, & C.Wieman,	2020	Can Majoring in Computer Science Improve General Problem-solving Skills?
8.	M.R.Putri,, & S.T.Sela,	2018	Capturing Students' Critical Thinking Skills in English for Academic Writing Course: A Case in Information System Students
9.	D.McGunagle,, & L.Zizka,	2018	Meeting Real World Demands of the Global Economy: An Employer's perspective
10.	H. Hadiyanto, N. Noferdiman, S. Syamsurizal, M. Muhaimin, Y.Yuliusman, S.Ramli, & U.Sulistyo,	2018	Graduate Candidates' 21st Century Skills and Challenge for the Faculty and the University in Online Learning Era
11.	A.T.KORUCU, & M.ÜNÜVAR,	2020	Bilgisayar ve Öğretim Teknolojileri Öğretmenliği Bölümünde Okuyan Öğretmen Adaylarının Aldıkları Eğitim ile Bilişim Liderliği Vasfı ve 21. Yüzyıl Becerileri Arasındaki İlişki
12.	E.van Laar, A. J.van Deursen, , J.A.Van Dijk, &J. de Haan,	2019	Twenty-first century digital skills for the creative industries workforce: Perspectives from industry experts
13.	V.Siddoo, J.Sawattawee, W.Janchai, & O.Thinnukool,	2019	An exploratory study of digital workforce competency in Thailand
14.	J.Yoon ,E.J Hur, & M.Kim,	2020	An Analysis of the Factors on the Problem-Solving Competencies of Engineering Employees in Korea
15.	M.KASAP,	2020	Yüzyıl Yönetici Yeteneklerinin Belirlenmesine Yönelik Bir Çalışma; İstanbul ve Bursa Örneği
16.	J.CNwokeji, R.Stachel, T.Holme, & R.O.Orji,	2019	Competencies Required for Developing Computer and Information Systems Curriculum

A Decision Support System Proposal for the Assignment of a Lawyer-Case in Law Offices

Nadide Gizem GÜRSON DOLAR¹, Başak GÖK², Hadi GÖKÇEN³

¹Management Information Systems Department, Başkent University, Ankara, Turkey

²Management Information Systems Department, Gazi University, Ankara, Turkey

³Industrial Engineering Department, Gazi University, Ankara, Turkey

nggdolar@baskent.edu.tr, basakgok@gazi.edu.tr, hgokcen@gazi.edu.tr

Abstract— The increase in the number of lawyers in our country in recent years has led to the emergence of large law offices where many lawyers are employed. These offices have a greater number of cases and diversity of cases compared to offices with a small number of lawyers. It is seen that the said offices are faced with problems such as which of the many lawyers to employ for a case they have received, and which lawyer will achieve more successful results in this type of case. In this study, it is aimed to create a decision support system that offers a recommendation to the manager about which lawyer should be appointed, taking into account the target effectiveness and constraints in a case that comes to large law offices, based on the effectiveness of lawyers. In this decision support system, the criteria determined by the expertise and expert opinions were used in the classification of the types of cases. In this direction, a “0-1 integer programming model” was developed for the model base, the model was solved with IBM ILOG CPLEX optimization software package, and a desktop application interface could integrate with CPLEX solver in C# Visual Studio environment was prepared. Thanks to the developed decision support system, it has aimed to solve the problem optimally, which office managers approach intuitively.

Keywords— decision support system, assignment problem, lawyer-case assignment

1. INTRODUCTION

The increase in number of lawyers in our country in recent years has led to the emergence of large law offices where many lawyers are employed. These offices have greater number of cases and diversity of cases compared to offices with a small number of lawyers.

When the large law offices operating today are examined, it has seen that the offices are faced problems such as which of the many lawyers they will employ for, which lawyer will achieve more successful results in accepted cases.

At the beginning of the study, the relevant literature was examined, and it was seen that a decision support system was created to be used in the field of law and the said study was related to the decisions to the unification of legal precedents by the Supreme Court (Yargıtay) [1]. In other words, there is no study in the literature on the assignment of lawyer-case.

In this study, a decision support system has been developed in law offices that have a large number of lawyers and litigation, in line with criteria such as the type of case,

workload, and constraints, which proposes the appointment of a lawyer to the managers.

2. CONCEPTUAL FRAMEWORK

2.1. Decision Support Systems

It is considered that the concept of decision support system actually arises from the inadequacy of management information systems and is a combination of information technology and operations research. If we need to make a general definition for the decision support system, we can say that the systems that help the decision makers in the managerial position to make decisions. A typical decision support system has three components. These components are [2];

- Data Management,
- Model Management and
- Dialogue Management.

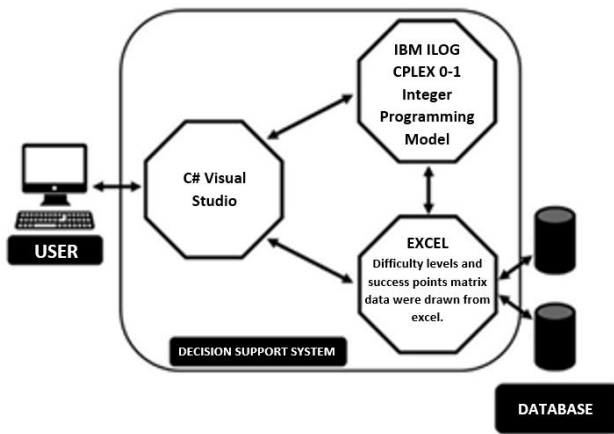


Figure 1. DSS Components

As seen in Figure 1, IBM ILOG CPLEX 0-1 integer programming model was used as the model management, C# Visual Studio was used as the dialog management, and excel application was used as the data management at decision support system.

2.2. Integer programming

Mathematical symbols and expressions used to represent a real situation are called models. Models are representations of real objects or situations and can be presented in a variety of forms. In modelling terminology, the representation of a problem by a system of symbols and mathematical relations or expressions are called mathematical models and are a critical part of any quantitative approach to decision making. Integer linear programming is an approach used for problems that can be set up as linear programs, provided that some or all of the decision variables are integer values [3].

Integer programming types are divided into three groups: pure, mixed, and 0-1 integer. In pure integer programming, all decision variables are required to be integers, in this group the questions are pure integer programming problems. Mixed integer programming is problems where some variables are desired to take integer values. Programming with 0-1 integers, decision variables can only take two values, such as 0 or 1 [4].

2.3. Classification of Cases by Subject

The resolution of a legal dispute by independent courts using judicial methods is called judicial activity. Independent judicial bodies have been established to carry out judicial activities. The Turkish judicial organization can be examined under three headings [5]:

- Constitutional jurisdiction
- Judicial jurisdiction
- Administrative jurisdiction.

The main task of the constitutional judiciary is to supervise the conformity of laws with the Constitution. The court

established for this purpose is the Constitutional Court. Apart from supervising the conformity of laws to the Constitution, the Constitutional Court also has duties such as adjudicating individual applications and adjudicating as the Supreme Court (Yargıtay) [6]. In cases that fall under the subject of constitutional jurisdiction, since law firms are not directly applied to in practice, they are not included in the classification of cases made in this study.

The judicial jurisdiction consists of courts with general jurisdiction within the judicial system, in other words, all matters other than the constitution and administrative jurisdiction are within the scope of the judicial jurisdiction [6].

Judicial jurisdiction is divided into two branches, civil jurisdiction and criminal jurisdiction. While private law disputes are resolved in civil jurisdiction, in criminal jurisdiction, persons who commit acts that are considered crimes according to penal laws are tried [5]. In the judiciary, the supreme court is the Supreme Court (Yargıtay). The Supreme Court (Yargıtay) decides on the files that come through appeal. There are fifteen civil and sixteen criminal chambers in the Supreme Court (Yargıtay) [5]. The Supreme General Assembly of the Supreme Court (Yargıtay Büyük Genel Kurulu) decides on which issues these chambers are responsible for dealing with cases. According to the division of labor decision of the Grand General Assembly of the Supreme Court (Yargıtay Büyük Genel Kurulu) of Appeals dated 2020, the cases that come to the civil and criminal departments in the judicial judiciary can be classified according to the their topics listed in Annex-1 [7].

Administrative jurisdiction is the branch of judiciary in which disputes arising from the actions and transactions of the administration are handled. In administrative jurisdiction, the defendant is always the administration [8]. The highest court in the administrative jurisdiction is the Council of State (Danıştay). The Council of State (Danıştay) decides on the decisions made by the administrative and tax courts and the decisions made at the end of the cases that it handles as a court of first instance, in case of appeal. The Council of State (Danıştay) consists of twelve chambers, eleven of which are litigation chambers and one is administrative chamber [8]. The Presidency Board of the Council of State (Danıştay Başkanlık Kurulu) decides on which subjects the litigation chambers are responsible for dealing with cases. Administrative lawsuits can be classified according to their subjects as in Annex-1, within the framework of the decision of the Presidency of the Council of State (Danıştay Başkanlık Kurulu) on the division of labor of the departments of the Council of State (Danıştay) dated 2020 [9].

In this context, a total of thirty types of cases have been determined, ten for civil cases, ten for criminal cases, and ten for administrative cases.

3. CURRENT PRACTICES FOR LAWYER-CASE ASSIGNMENT

In our country, there are two applications prepared by the Turkish Bar Associations (Türkiye Barolar Birliği) for lawyer-case. The first of these applications is the CPC Automatic Lawyer Assignment and Follow-up Program, which is used to appoint a lawyer to be made by the Bar Association, upon the request of the authority that took the statement during the criminal investigation phase or the judge who made the interrogation according to the Criminal Procedure Code, or upon the request of the court during the criminal prosecution phase.

The second is the Legal Aid Lawyer Assignment and Follow-up System, which is a lawyer Assignment system for legal aid requests made by the Legal Aid Offices of Bar Associations, which is prepared for the healthy implementation of the legal aid service provided by the state to people who need legal protection but not have economic opportunity to benefit from the legal service.

3.1. CPC Automatic Lawyer Assignment and Follow-up Program (CMK Otomatik Avukat Atama ve Takip Programı)

The working features and parameters of CPC Automatic Lawyer Assignment and Follow-up Program are as follows [10]:

- The CPC Lawyer Assignment and Follow-up Program, prepared by the Union of Turkish Bar Associations (Türkiye Barolar Birliği), was put into use as of September 2013.
- The CPC Assignment program has been made for the purpose of common use of all bar associations, except for some settings, in order to establish a certain standard in bar associations.
- Since the change requests from the bar associations will affect all the bar associations included in the system, the condition of complying with the regulations and generally approved rules is made by reviewing.
- The program is offered to bar associations free of charge.
- Lawyers who want to be included in the assignment list must have an identity card with Barocard features.
- Institutions can enter their requests with the Virtual Barocard number and password given to them by the bar associations.
- During the request registration phase, the contact information is automatically withdrawn from “MERNİS”.
- In the system, the bar associations are given the settings menu, and they are given the opportunity to shape the following settings according to their own applications.

- Lawyers were given the opportunity to choose day-night, investigation-prosecution, the district they want to be in charge of, and active-passive transactions by informing the bar association while they were registered in the system.
- At the stage of registration of lawyers in the system, in addition to the average score and average score according to the districts, additions and subtractions can be made according to the scores determined by the bar associations themselves.
- Lawyers can log in via the Barocard website, see the assignment details, points and penalties, perform active-passive transactions, change their e-mail address and GSM number.
- Our bar associations who want to switch to the program are given the opportunity to test.

3.2. Legal Aid Lawyer Assignment and Tracking System (Adli Yardım Avukat Atama ve Takip Sistemi)

The Legal Aid Lawyer Assignment and Follow-up System, prepared by the Union of Turkish Bar Associations (Türkiye Barolar Birliği), was put into service in October 2012 in order to ensure that the procedures and principles related to Legal Aid are implemented properly and to eliminate the problems experienced by the Bar Association Legal Aid Offices in recording applications, appointing lawyers and following up the applications.

The features of the Legal Aid Lawyer Assignment and Tracking System are as follows [11]:

- The system can be used in two different ways to make assignments on the basis of Scoring and Number of Tasks, thus ensuring a fair distribution.
- In order for the applicant’s information to be automatically entered into the system, the Lawyers can choose the district in charge and follow up their transactions on www.barokart.com.tr.
- In order to provide ease of follow-up in the system, the lines “Subject of Litigation-Authority-Legal Reason” are automatically activated.
- If requested by the Bar Association Legal Aid Board, a report can be obtained by logging into the system.
- Through this application, payment-collection transactions can be easily recorded and reported for both lawyers and Bar Association Legal Aid Units.
- Quarterly reports sent to the Turkish Bar Associations (Türkiye Barolar Birliği) can still be tracked and recorded through the system.

3.3. Comparison of the Proposed Decision Support System with Existing Applications

When the features and parameters given above regarding the “CPC Automatic Lawyer Assignment and Follow-up Program” and the “Legal Aid Lawyer Assignment and Follow-up System” are examined, it is possible to say that both applications are actually an assignment application. But these applications ignore the assignment efficiency while performing the assignment process. That is, there is a point system on the basis of both applications, but the points are determined according to the number of assignments and the place of duty, not the possible success or performance of the lawyers in the cases they have accepted. The system tries to distribute the number of duties fairly by assigning the lawyer with the lowest score, while doing this, it does not take into account the efficiency of the assignment. When the purposes of these applications are examined, it is seen that the first purpose of both applications is to ensure that the distribution of the number of tasks has is fair, and the second purpose is to facilitate the follow-up of payment and collection.

The decision support system for the assignment of lawyers and selecting case types, which will be proposed within the scope of our study, differs from the existing practices as the assignment which is made effectively based on lawyer success scores and case difficulty level.

4. RESEARCH METHODOLOGY

In the study, first, the types of cases were classified according to their subjects in line with the “recent decisions regarding the division of labor” of the Supreme Court of Appeals Grand General Assembly (Yargıtay Büyük Genel Kurulu) in 2020 and the Council of State Presidency (Danıştay Başkanlık Kurulu) in 2016. In cases that fall under the subject of constitutional jurisdiction, since law firms are not directly applied to in practice, they are not included in the classification of cases made in this study. In this context, a total of 30 types of cases have been determined, 10 for civil cases, 10 for criminal cases and 10 for administrative cases. In the case type difficulty scoring, first, the criteria affecting the level of difficulty were determined by the experts. The content validity of the determined criteria was submitted to expert opinion with the “Criteria for Determining the Difficulty Levels of Case Type”. Next, experts scored difficulty rate between 1 and 5, for 30 case types, considering the criteria. Thanks to these degrees of difficulty, it is possible to ensure a fair distribution of the caseload for each lawyer.

In order to determine the estimated effectiveness of a lawyer in a new case type accepted by the office, first of all, the criteria that affect the success of the lawyer at case types were determined. The content validity of the determined criteria was presented to the expert opinion with the “Lawyer Success Determination Criteria Expert Opinion Form”. Then experts scored rating between 1 and 9 for the probable success of 10 lawyers in each of the 30 case types. Thus, the estimated effectiveness of each lawyer in a new case has been tried to be determined by grading in the light of criteria such as talent, expertise or

experiences. Thanks to this rating, it has been possible to assign the lawyer with the highest probability of success in the case type in accepted cases.

In this study, a “0-1 integer programming model” was developed in line with the objectives and constraints determined for the problem; the model was solved with the IBM ILOG CPLEX optimization software package that is desktop application interface could integrate with the CPLEX solver in C#; the Visual Studio environment was prepared; and a decision support system proposal was presented in the study.

4.1. Determination of Criteria

For the experts to determine the degree of difficulty of the types of cases, first the criteria affecting the degree of difficulty in question were determined. The content validity of the criteria, are affecting the degree of difficulty of case type, was presented to the expert opinion with the “Criteria for Determining the Difficulty Levels of Case Types”.

Next, experts scored difficulty rate between 1 and 5, for 30 case types, considering the criteria. Each score of litigation difficulties is constructed as follows:

- 1 - Very easy
- 2 - Easy
- 3 - Medium
- 4 - Difficult
- 5 - Very Difficult

Thanks to these degrees of difficulty, it is possible to ensure a fair distribution of the caseload on each lawyer.

In order to determine the estimated effectiveness of a lawyer in a new type of case that comes to the office, first of all, the criteria that affect the success of a lawyer against any type of case are determined. The content validity of the determined criteria was presented to the expert opinion with the “Lawyer Success Determination Criteria Expert Opinion Forum”. The scope validity of the criteria that affect the success scores of lawyers against the types of cases will be determined by expert opinion.

The experts were then asked to give a score of 1-9 for the probable success of the 10 lawyers against each of the 30 types of cases. Thus, the estimated effectiveness of each lawyer in the office in a new case has been tried to be determined by grading in the light of criteria such as talent areas and areas of expertise.

Thanks to this rating, it will be possible to appoint the lawyer with the highest probability of success in each case that comes to the office against that type of case.

4.2. Developing the Model

In order to maximize Lawyer-Case assignment efficiency, a 0-1 integer programming model was developed [12] [13] and solved with the test data using the IBM ILOG CPLEX optimization software package. The created assignment model and codes are given below:

Indices

i: number of lawyers

j: number of cases

Parameters

A_{ij} : Total benefit value obtained in case the i. lawyer is appointed to the j. case

B_j : difficulty value of j. cases

Variables

X_{ij} : 1 if the i. lawyer is assigned to the j. case, 0 otherwise

$$\text{Objective function: } \max Z = \sum_{i=1}^I \sum_{j=1}^J A_{ij} X_{ij} \quad (4.1)$$

Constraints:

$$\sum_{i=1}^{10} X_{ij} = 1 \quad j = 1, 2, \dots, 30 \quad (4.2)$$

$$\sum_{j=1}^{30} X_{ij} B_j = 1 \quad i = 1, 2, \dots, 10 \quad (4.3)$$

$$X_{ij} \in \{0, 1\}$$

The CPLEX codes used are as follows:

```
int avukat=10;
```

```
int dava =30;
```

```
range avukatlar=1..avukat;
```

```
range davalar=1..dava;
```

```
int A[avukatlar][davalar]=...;
```

```
dvar boolean X[avukatlar][davalar];
```

```
maximize sum (i in avukatlar, j in davalar)
(X[i][j]*A[i][j]);
```

```
subject to
```

```
{
forall (j in davalar) sum (i in avukatlar) X[i][j]==1;
forall (i in avukatlar) sum (j in davalar) X[i][j]<=10;
}
```

```
For the new case;
```

```
int avukat=10;
```

```
int dava =30;
```

```
int yenidava=1;
```

```
int yuk=10;
```

```
range avukatlar=1..avukat;
```

```
range davalar=1..dava;
```

```
range yenidavalar=dava+1..dava+yenidava;
```

```
range yukler=1..yuk;
```

```
int B[davalar]=...;
```

```
int C[avukatlar][yenidavalar]=...;
```

```
int D[yenidavalar]=...;
```

```
int E[yukler]=...;
```

```
dvar boolean Y[avukatlar][yenidavalar];
```

```
maximize sum (i in avukatlar, j in yenidavalar)
(Y[i][j]*C[i][j]);
```

```
subject to
```

```
{
forall (j in yenidavalar) sum (i in avukatlar) Y[i][j]==1;
forall (i in avukatlar) sum (j in yenidavalar)
Y[i][j]*D[j]+E[i]<=10;
}
```

4.3. Desktop Application (User Interface)

Sample pictures of the results obtained by using the decision support system interface and input data created for lawyer-case assignment are as follows. First of all, the operation of the decision support system was tested with

the test data. The input data that desktop application pulls from excel is shared in Annex-2.



Figure 2. DSS splash screen

Figure 2 represents the opening screen of DSS. During the establishment of DSS to the office, firstly, the “initial assignment screen” is accessed from the opening screen.

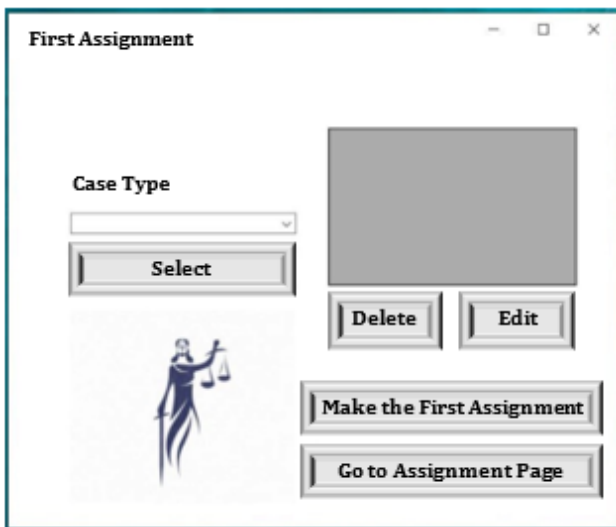


Figure 3. The screen where the first assignment operations are performed

At the initial assignment, it was assumed that the office currently has 30 cases and 10 lawyers. If the assignments related to the current cases in the office were made with the created DSS, the result would be as seen on the console screen in Figure 5 and the “initial assignment report screen” in the desktop application in Figure 6.

In the “initial assignment report screen”, it is seen that the lawyers with the highest possible success in the case type in question are assigned to the cases. As a result of the first assignment, the burdens on the lawyers arising from the difficulty levels of the types of cases arose.

In the ongoing process, a new case coming to the office is assigned based on these burdens and the lawyer’s success

score. Figure 7 and Figure 8 represent the assignment based on the current burdens and success score of a new case.

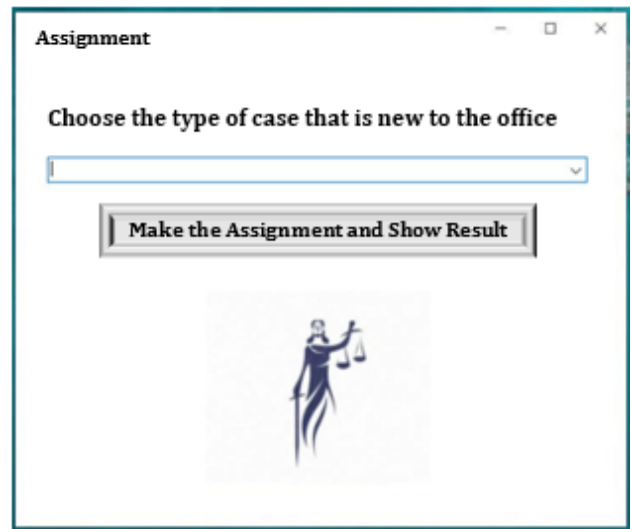


Figure 4. The screen where the assignment operations are performed

```
Parallel b&c, 12 threads:
Real time      = 0,00 sec. (0,00 ticks)
Sync time (average) = 0,00 sec.
Wait time (average) = 0,00 sec.
-----
Total (root+branch&cut) = 0,09 sec. (4,42 ticks)

OBJECTIVE: 252
// solution (optimal) with objective 252
// Quality Incumbent solution:
// MILP objective                                2,5200000000e+02
// MILP solution norm |x| (Total, Max)          3,00000e+01 1,00000e+00
// MILP solution error (Ax-b) (Total, Max)     0,00000e+00 0,00000e+00
// MILP x bound error (Total, Max)            0,00000e+00 0,00000e+00
// MILP x integrality error (Total, Max)       0,00000e+00 0,00000e+00
// MILP slack bound error (Total, Max)        0,00000e+00 0,00000e+00
//
X = [[0 0 0 0 0 1 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 1 0 1 0 0 0 0]
[0 0 1 0 0 0 0 0 0 0 0 0 0 1 0 0 0 0 0 0 1 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0]
[1 1 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 1 0 0 1 0 0 0 0 0 0 0 0 1]
[0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0]
[0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0]
[0 0 0 0 0 0 0 1 0 0 0 0 0 0 1 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0]
[0 0 0 0 0 0 0 1 0 0 0 0 0 0 1 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0]
[0 0 0 0 0 1 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0]
[0 0 0 0 1 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0]
[0 0 0 1 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0]
[0 0 0 1 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0]];
Çıkış için enter'a bas
```

Figure 5. First assignment result screen 1

In the “assignment report screen”, it is seen that the new incoming case is assigned to the lawyer with the highest possible success among the lawyers whose case load is suitable, according to the difficulty level.

ILK ATAMA RAPORU

F1	H1	H2	H3	H4	H5	H6	H7
Ankara1	0	0	0	0	0	1	0
Ankara2	0	0	1	0	0	0	0
Ankara3	1	1	0	0	0	0	0
Ankara4	0	0	0	0	0	0	0
Ankara5	0	0	0	0	0	0	0
Ankara6	0	0	0	0	0	0	0
Ankara7	0	0	0	0	0	0	0
Ankara8	0	0	0	0	0	0	1
Ankara9	0	0	0	0	1	0	0
Ankara10	0	0	0	1	0	0	0

Figure 6. First assignment result screen 2

```

Parallel b&c, 12 threads:
  Real time      = 0,00 sec. (0,00 ticks)
  Sync time (average) = 0,00 sec.
  Wait time (average) = 0,00 sec.
-----
Total (root+branch&cut) = 0,00 sec. (0,01 ticks)

OBJECTIVE: 5
// solution (optimal) with objective 5
// Quality Incumbent solution:
// MILP objective                    5,0000000000e+00
// MILP solution norm |x| (Total, Max) 1,00000e+00 1,00000e+00
// MILP solution error (Ax=b) (Total, Max) 0,00000e+00 0,00000e+00
// MILP x bound error (Total, Max) 0,00000e+00 0,00000e+00
// MILP x integrality error (Total, Max) 0,00000e+00 0,00000e+00
// MILP slack bound error (Total, Max) 0,00000e+00 0,00000e+00
//
V = [[0]
      [0]
      [1]
      [0]
      [0]
      [0]
      [0]
      [0]
      [0]
      [0]];
Çıkış için enter'a bas

```

Figure 7. Assignment result screen 1

ATAMA RAPORU

Avukatlar	Atama Sonucu
Av1	0
Av2	0
Av3	0
Av4	1
Av5	0
Av6	0
Av7	0
Av8	0
Av9	0
Av10	0
*	

Figure 8. Assignment result screen 2

5. CONCLUSION AND DISCUSSION

With the decision support system developed for lawyer-case assignment, it has been possible to both appoint the lawyer with the highest possible success in a new case, and to ensure a fair distribution of the caseload on each lawyer. To the best knowledge of the author, this study is the first DSS application carried out in this context. The decision support system created to present different scenarios (what-if) to the manager will be considered in the continuation of the study. In the continuation of the study,

lawyer-case assignments in a real law firm will be carried out, and the positive difference between the current assignment and the assignment made by the model will be evaluated. Again, model validation will be carried out with the feedback received from the experts.

REFERENCES

- [1] A. Güneş, **Yargıtay Kararları Birleştirilmesi İçin Uzman Sistemler Aracılığı ile Yargıtay İçtihatlarının Birleştirilmesinde Karar Destek Sistemi Geliştirilmesi**, Master's Thesis, Türk Hava Kurumu Üniversitesi Fen Bilimleri Enstitüsü, 2016.
- [2] H. Gökçen, **Yönetim Bilgi/Bilişim Sistemleri: Analiz ve Tasarım**, Palme Yayınevi, Ankara, 2011.
- [3] D. R. Anderson, D. J. Sweeney, T. A. Williams, J. D. Camm and K. Martin, **An Introduction to Management Science: Quantitative Approaches to Decision Making**, South-Western Cengage Learning, Mason, 2012.
- [4] E. Eroğlu, **Yöneylem Araştırması**, İstanbul Üniversitesi Açık ve Uzaktan Eğitim Fakültesi, İstanbul, 2019.
- [5] H. H. Sümer, **Hukuka Giriş**, Seçkin Yayıncılık, Ankara, 2018.
- [6] Y. Atar, **Türk Anayasa Hukuku**, Seçkin Yayıncılık, Ankara, 2018.
- [7] Internet: Yargıtay, Yargıtay Büyük Genel Kurul Kararı, <https://www.yargitay.gov.tr/documents/ek1-1580383414.pdf>, 25.12.2020.
- [8] E. E. Atay, **İdari Yargılama Hukuku**, Seçkin Yayıncılık, Ankara, 2020.
- [9] Internet: Danıştay, Danıştay Başkanlık Kurulu Kararı, <https://www.danistay.gov.tr/icerik/26>, 25.12.2020.
- [10] Internet: Türkiye Barolar Birliği, Türkiye Barolar Birliği, <https://www.barobirlik.org.tr/dosyalar/duyurular/cmk.pdf>, 25.12.2020.
- [11] Internet: Türkiye Barolar Birliği, Türkiye Barolar Birliği, <https://www.barobirlik.org.tr/Haberler/adli-yardim-avukat-atama-ve-takip-sistemi-14780>, 25.12.2020.
- [12] C. Haksever and B. Render, **Service Management: An Integrated Approach to Supply Chain Management and Operations**, FT Press, New Jersey, 2013.
- [13] W. L. Winston and J. B. Goldberg, **Operations Research: Applications and Algorithms**, Thomson Brooks/Cole, California, 2014.

Appendix-1

Civil Cases	
Cases Arising from Family Law	Cases Arising from Labor Law
Cases Arising from Real Estate Law	Cases Arising from Contract Law
Compensation Cases Arising from Tort	Cases Arising from Commercial (Commercial Business-Companies-Negotiable Instruments) Law
Cases Arising from Execution and Bankruptcy Law	Compensation Cases Arising from Traffic Accidents
Expropriation Cases	Cases Arising from Consumer Law

Criminal Cases	
Crimes Against Property (Theft, Looting, Damage to Property, Fraud and Fraudulent Bankruptcy)	Crimes Against Public Health (Manufacturing and Trafficking in Drugs or Stimulant Substances, Using Drugs or Stimulant Substances, and Behaving Against Measures Regarding Communicable Diseases)
Offenses Against Life and Body Immunity (Killing, Injuring, Torture and Suicide) and Insult	Crimes in the Field of Information (Intrusion, Blocking, Destroying the System, Destroying or Modifying Data and Misuse of Bank or Credit Cards)
Offenses Against Sexual Immunity (Sexual Assault, Sexual Abuse of Children and Sexual Harassment)	Crimes Against Public Peace (Praising the Crime and the Criminal, Inducing or Insulting the Public to Enmity and Enmity, and Forming an Organization to Commit a Crime)
Crimes Against Freedom (Threat, Blackmail, Deprivation of Liberty and Violation of Immunity of Housing)	Crimes Related to Economy, Industry and Trade (Tender Rigging and Usury)
Endangering Traffic Safety, Polluting the Environment, Causing Noise and Causing Zoning Pollution)	Crimes Against the Trustworthiness and Functioning of the Public Administration (Employment, Extortion, Bribery and Misconduct)

Administrative Cases	
Ecrimisil Cases	Cases Arising from Higher Education Legislation
Cases Arising from Energy and Customs Legislation	Cases Arising from Private Education Institutions Legislation
Cases Regarding Public Officials	Cases Regarding Taxes, Duties, Fees and Similar Financial Liabilities
Cases Related to Tenders	Cases Arising from the Law on the Protection of Competition and Privatization Practices
Cases Arising from Zoning and Settlement Legislation	Cases Regarding Administrative Fines

Appendix-2

	H 1	H 2	H 3	H 4	H 5	H 6	H 7	H 8	H 9	H 10	C 1	C 2	C 3	C 4	C 5	C 6	C 7	C 8	C 9	C 10	I 1	I 2	I 3	I 4	I 5	I 6	I 7	I 8	I 9	I 10
L 1	3	5	5	8	9	9	9	4	3	2	1	3	6	8	1	1	3	6	8	7	9	3	6	9	8	9	9	2	4	4
L 2	9	3	8	4	7	6	8	1	8	3	5	9	1	6	7	7	5	9	6	6	2	6	8	3	8	9	1	5	1	5
L 3	9	7	2	6	7	6	1	1	3	4	7	6	1	3	3	4	5	7	6	8	7	8	8	5	5	9	3	4	1	9
L 4	8	3	5	4	8	5	5	1	9	2	5	4	8	5	4	7	3	6	7	4	6	2	8	7	8	3	9	6	5	3
L 5	2	3	5	9	9	7	3	5	9	4	3	1	1	2	8	9	4	3	9	4	4	5	5	3	3	6	1	3	8	8
L 6	2	5	6	6	6	2	4	5	5	1	3	7	8	8	2	6	9	6	7	7	1	1	2	3	9	7	6	7	4	7
L 7	5	1	1	5	4	3	5	1	9	5	6	1	7	1	7	9	6	1	7	5	6	1	3	6	2	6	7	6	6	1
L 8	9	1	2	3	9	7	9	5	5	7	2	9	5	4	3	8	9	3	9	4	9	1	7	7	2	9	9	9	3	2
L 9	4	4	6	1	9	2	2	1	3	4	1	3	6	3	6	2	1	3	5	4	3	7	2	1	5	8	7	4	8	3
L 10	8	2	2	9	3	4	2	3	3	9	1	8	5	2	6	6	3	2	4	3	2	1	1	4	1	2	1	2	5	5

H 1	H 2	H 3	H 4	H 5	H 6	H 7	H 8	H 9	H 10	C 1	C 2	C 3	C 4	C 5	C 6	C 7	C 8	C 9	C 10	I 1	I 2	I 3	I 4	I 5	I 6	I 7	I 8	I 9	I 10
1	3	2	5	5	2	1	2	1	4	3	4	5	5	3	4	1	4	3	2	1	5	1	3	3	5	4	5	3	2

Prediction of Life Quality Index Value Rankings of Countries After the COVID-19 Pandemic with Artificial Neural Networks*

Ali UÇUCU^{1*}, Başak GÖK², Hadi GÖKÇEN³

¹OSTIM Technical University, Management Information Systems, Ankara, Turkey

²Gazi University, Management Information Systems, Ankara, Turkey

³Gazi University, Industrial Engineering, Ankara, Turkey

¹ali.ucucu@ostimteknik.edu.tr, ²basakgok@gazi.edu.tr ³hgokcen@gazi.edu.tr.

Abstract— As there are indicators that affect the performance criteria of businesses, countries also have some indicators that affect the quality of life. Indicators affecting the quality of life; which are property, pollution, health care, crime, traffic, and cost of living index. These indexes are created by analyzing data entered by citizens. Economic crisis, pandemics, natural events, etc. conditions affect these indexes. The COVID-19 pandemic affected all these indexes. In this study, the impact of the COVID-19 pandemic on the quality of life of countries was investigated. In this context, the artificial neural network was trained with the data of 5 countries in the European continent, consisting of a total of six indexes between 2012 and 2019. In the scenario where the COVID-19 pandemic never happened, the quality of life of the countries for 2020 was estimated and compared with the quality of life realized in 2020. It has been observed that the ANNs tested by charting the MAPE criterion and past period estimates gave successful results. As a result, it was determined that the COVID-19 pandemic caused a significant decrease in the quality of life of three countries but did not affect the quality of life of the remaining two countries.

Keywords— artificial neural networks, deep learning, quality of life index, prediction, time series

1. INTRODUCTION

The most basic features of Artificial Neural Networks (ANNs) are that they can make decisions based on information while producing or trying to produce solutions to events or problems and they can produce predictions about the next events by learning the events with the information at hand. These systems enable solving problems that are almost impossible to solve manually. ANNs consist of cells that are hierarchically connected to each other, these cells, called process elements, store the information they learn in connections. The main task of an ANN is to produce an output set that can correspond to the set of inputs shown to it. In order to do this, the network should be trained with examples of related events and have the ability to generalize. These systems, which work by imitating the working principles of the human brain such as learning information, being able to derive, discover and predict new information, although they remain primitive compared to the speed of the human brain, nowadays have become quite fast. These speeds can be measured in nanoseconds [1].

With the first cases of COVID-19 seen at the end of 2019, it has been observed that countries are experiencing both social and economic changes. Due to the fact that the year 2020 includes the entire pandemic period, almost all the habits of people have changed. In this study, it will be investigated how the changes experienced affect the indexes of the countries. ANN, which will be trained with index data before the year 2020, will estimate the quality of life index for the year 2020 when there is no COVID-19 pandemic. The obtained result will be compared with the index for the year 2020, which was experienced in real. The trust of citizens in their countries after the COVID-19 pandemic and how the pandemic affects the quality of life of citizens will be investigated within the scope of this study. At the end of the study, it will be presented how countries are affected by COVID-19 pandemic. In this way, it is expected that countries will be able to identify their deficiencies and take precautions against extraordinary situations. At the same time, it is expected that this study will contribute to the widespread use of ANN in the field of social sciences.

2. CONCEPTUAL FRAMEWORK

2.1 Artificial Neural Networks

Artificial neural network cells are similar to biological nervous systems. The task of the artificial neural network, just like the biological nervous system, is to produce an output as shown in Figure 1 in response to the electrical signals it uses as input. For this to be possible, the network must be trained with certain data sets. In this way, the network reaches the level that it can generalize and determines the outputs with these capabilities [2]. At the basis of artificial neural networks are three layers, these layers are the input layer, hidden layer and output layer [3].

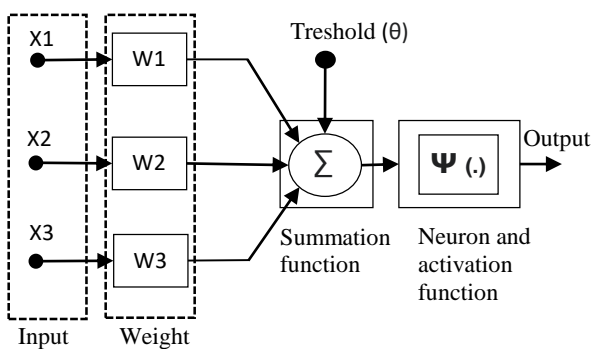


Figure 1. Artificial nerve cell model [4]

The estimation error in ANN consists of the difference between the actual value and the estimated value for a certain period [5]. The accuracy of the information produced by the established ANN is examined with error criteria. The network should be tested with the error criteria in the literature and should be within acceptable limits. It is very important for the accuracy of the network that the error values are within acceptable limits.

The MAPE criterion generally gives more effective results than other criteria, and it is sufficient to explain the change. The criteria to be considered while examining the MAPE values are as in Table 1 [6].

Table 1. MAPE value and descriptions

MAPE value	Descriptions
MAPE < %10	High accuracy
MAPE > %10 and < %20	Medium accuracy
MAPE > %20 and < %50	Low accuracy but acceptable
MAPE > %50	No predictive value

2.2. Country Indexes

Some measurements affect the quality of life of countries and these are defined as country indexes. These indexes are property price index, pollution index, health care index,

crime index, traffic index, and cost of living index. The fact that the quality of life index highly depends on the fact that the health service index is high, while other indexes are low. Country index descriptions are as follows [7]:

When creating a property index, the ratio of house prices to incomes, the ratio of the mortgage loan to the percentage of income, the affordability ratio of the mortgage loan, the ratio of housing prices to housing rents, and the annual gross rental yield are established using the information. The pollution index is formed by an eight-question questionnaire, in which the weight is given to air pollution and water pollution. While creating the traffic index, the time spent in traffic, the inefficiency rate due to the lack of public transport use, and the carbon dioxide emission rate emitted depending on the traffic duration are used. The crime index is also determined by the results obtained by a twelve-question questionnaire. While creating the cost of living index, rent index, market index, restaurant, and cafe index, and local purchasing power are used. The cost of living index is adjusted for New York, the cost of living index for New York is 100%. If the cost of living index of a city is 140, it means that the city is 40% more expensive than New York City.

The indexes mentioned above were created by processing the data entered by the citizens by software written in the Java programming language. The quality of life index is reached by evaluating these indexes with certain weights.

3. METHOD

In this study, the 2020 quality of life indexes of Germany, France, Italy, Turkey, and Greece were estimated by ANN. YSA, which has a successful result in almost every field in which it is used; It can work with incomplete information, adapt to the problems it encounters, make decisions under uncertainties and be tolerant of errors [8]. Therefore, in this study, ANN has been preferred for the creation of an estimation model.

All ANN's installed were created using the Matlab 2019B program. At the same time, while creating the ANN, a multi-layered, feedback network type was chosen as seen in Figure 2. For the input layer, 13x6 data was used as the equivalent of six indexes collected over thirteen terms, and for the output layer, 1x13 data was used as the equivalent of the quality of life index collected over thirteen terms. Trainlm function is used as a training function, Learnsgdm function is used as learning function, MSE function is used as performance function, and Tansig function is used as activation function. While determining the number of hidden neurons to be used in the hidden layer, it was decided to use 15 hidden neurons in the hidden layer by trial and error method, since it is not possible to act according to certain rules and there is no systematic

approach to find the optimal number of neurons [9]. Five separate networks have been created for five countries.

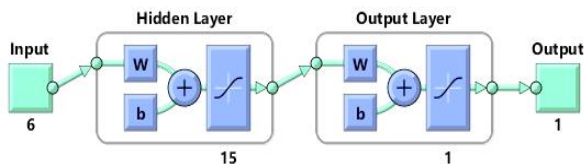


Figure 2. Artificial neural network model

70% of the data was used for training, 15% was used for verification, and the remaining 15% was used for testing.

The data used for training from these stages are aimed at learning the network, the data used for verification is used to determine when the training should be terminated, while the data used for testing checks whether the network has achieved the desired performance [10].

3.1. Training ANN with Country Indexes and Predicting the Year 2020 when COVID-19 does not exist.

3.1.1. Index Data Used

As a period was chosen between 2012, the first year for which complete data was available, and 2019 when the pandemic had not yet started. The data presented once a year from 2012 to 2014 began to be presented twice a year from 2014, so there are fourteen terms of data in each index of each country. These data are property price index,

pollution index, health service index, crime index, traffic index, cost of living index, and quality of life index, and these data were used in the study. For ANN trained with one-term transferred; Six thirteen term-old index data (13x6) was used as the input layer, and thirteen-periods quality of life index data (13x1) was used as the output layer.

The ANN data set up for the estimation was taken from Numbeo [7]. These data were used for training the network and for comparing its outputs.

3.1.2. Creation and Training of the ANN

For each country 2012, 2013, 2014, 2014-2, 2015, 2015-2, 2016, 2016-2, 2017, 2017-2, 2018, 2018-2, 2019 the six indexes of their terms are taken from Numbeo and these data are used in the input layer. Since the data are given in six-month periods after 2014, it is seen that two data are received every year since 2014. These independent variables formed the input layer of the ANN. For each country 2013, 2014, 2014-2, 2015, 2015-2, 2016, 2016-2, 2017, 2017-2, 2018, 2018-2, 2019, 2019-2 the quality of life index of their terms are taken from Numbeo and these data are used in the output layer.

4. RESULTS

Input layer data of Germany used for ANN training in Table 2 and output layer data from Germany used for ANN training in Table 3 are given.

Table 2. The input layer of Germany's artificial neural network

Years	Property prices	Pollution	Health care	Crime	Traffic	Cost of living
2012	5,21	45,14	71,96	21,20	80,77	88,55
2013	4,78	28,43	73,25	23,86	74,55	91,64
2014	5,60	30,20	75,26	27,14	132,35	87,14
2014-2	6,13	30,33	76,66	29,13	102,09	86,82
2015	6,61	28,32	75,61	28,49	98,41	76,27
2015-2	7,07	29,22	76,37	30,64	86,25	65,96
2016	7,13	29,91	75,85	32,95	87,46	65,54
2016-2	7,23	27,68	76,02	32,70	97,03	66,79
2017	7,46	27,05	76,74	32,03	95,92	65,50
2017-2	7,82	27,68	75,99	32,64	97,02	73,18
2018	8,05	27,32	75,44	36,65	102,94	74,35
2018-2	8,48	28,15	74,88	36,16	101,49	67,89
2019	9,02	28,01	74,32	34,51	98,09	67,62

Table 3. The output layer of Germany's artificial neural network

Years	Quality of Life
2013	204,84
2014	192,69
2014-2	186,61
2015	195,94
2015-2	244,31
2016	199,70
2016-2	189,87
2017	189,74
2017-2	178,92
2018	190,04
2018-2	189,13
2019	187,05
2019-2	184,30

According to the designed ANN, 2013 was estimated with the input data of 2012, and 2014 was estimated with the input data of 2013. In this way, ANN, trained with 13-term data in this way, was simulated with the data of the 2019-2 period and reached the 2020 output data, that is, the quality of life index of the year 2020, when COVID-19 was not experienced.

The regression values of the training, testing, and verification stages of the ANN established for Germany are shown in Figure 3.

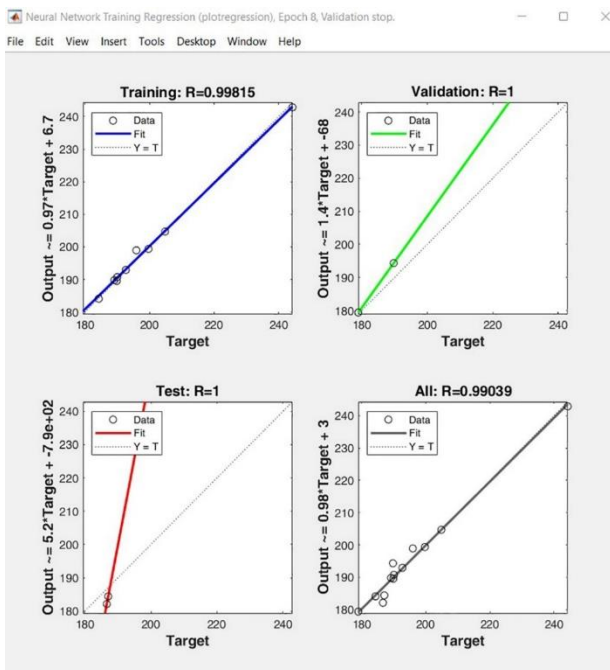


Figure 3. YSA regression results for Germany

When the trained ANN was simulated with 2019-2 input data, the quality of life of Germany in 2020 was estimated at 229,906 in the scenario where there is no COVID-19. In the presence of COVID-19, Germany's quality of life was

179.78. Figure 4 shows the retrospective estimates of the network. It is seen that all the historical forecasts are in agreement with the actual forecasts. The reason for the difference in 2020 is COVID-19.

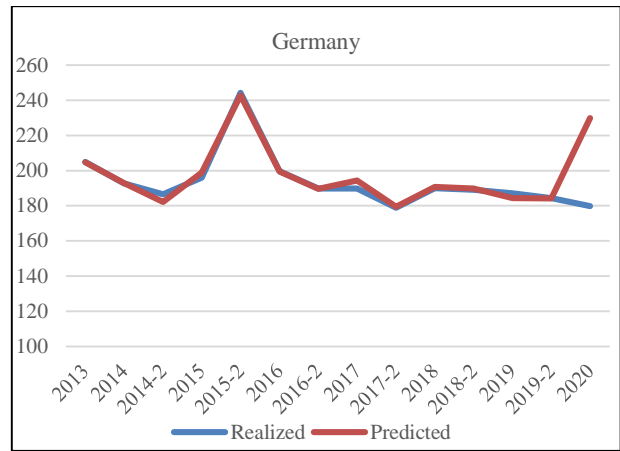


Figure 4. Predicted-realized comparison of Germany

Detailed knowledge of the trained network is explained by the example of Germany. In the continuation of the study, the outputs of the network trained for France, Italy, Turkey, and Greece will be given.

The regression values of the training, testing, and verification stages of the ANN established for France are shown in Figure 5.

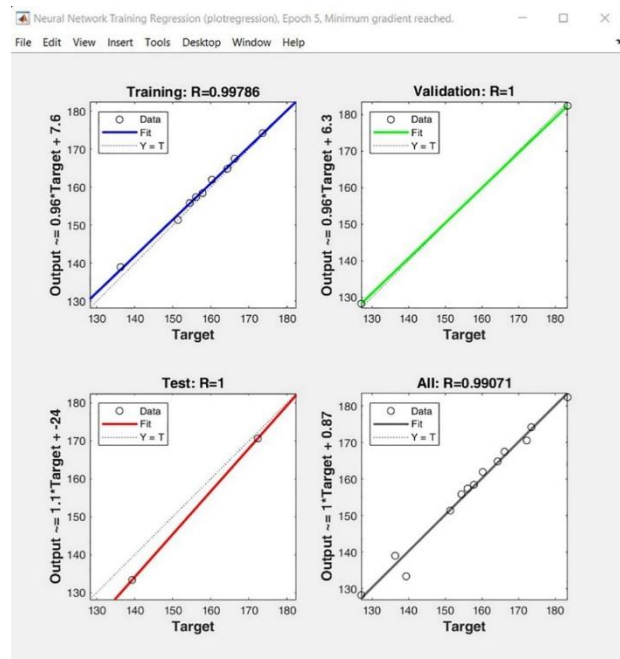


Figure 5. YSA regression results for France

As a result of the training, it was predicted that the quality of life index of France was 157.69 in the scenario where there was no pandemic. As a result of the pandemic, the quality of life index of France was published as 153.95. In Figure 6, retrospective estimates of the network are given.

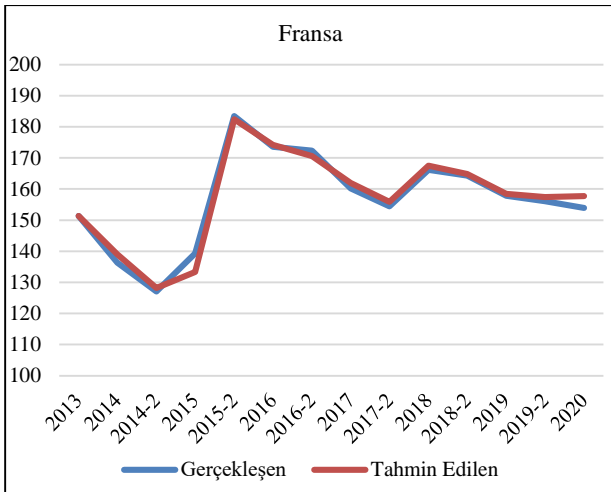


Figure 6. Predicted-realized comparison of France

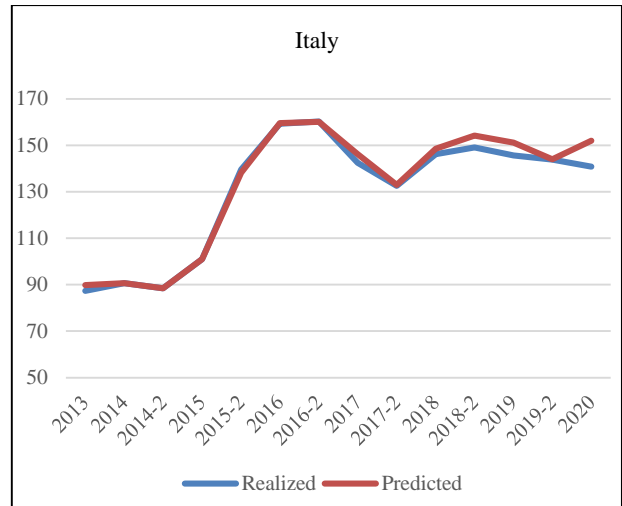


Figure 8. Predicted-realized comparison of Italy

The regression values of the training, testing, and verification stages of the ANN established for Italy are shown in Figure 7.

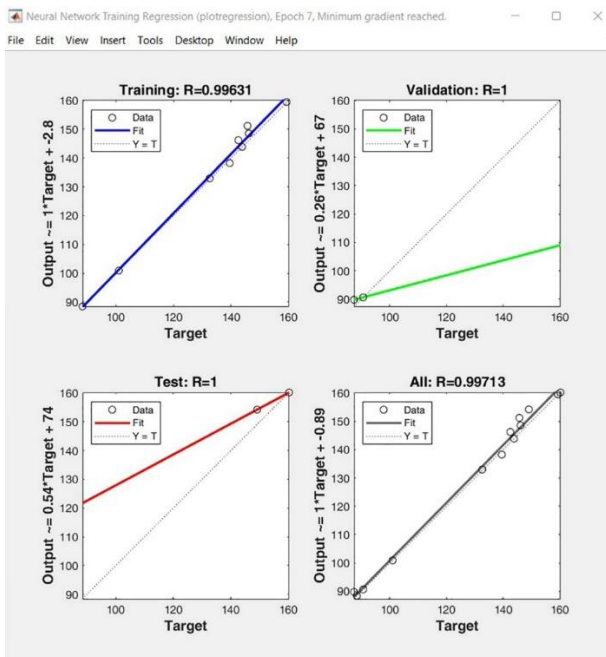


Figure 7. YSA regression results for Italy

As a result of the training, it was predicted that the quality of life index of Italy was 152.02 in the scenario where there was no pandemic. As a result of the pandemic, the quality of life index of France was published as 140.76. In Figure 8, retrospective estimates of the network are given.

The regression values of the training, testing, and verification stages of the ANN established for Turkey are shown in Figure 9.

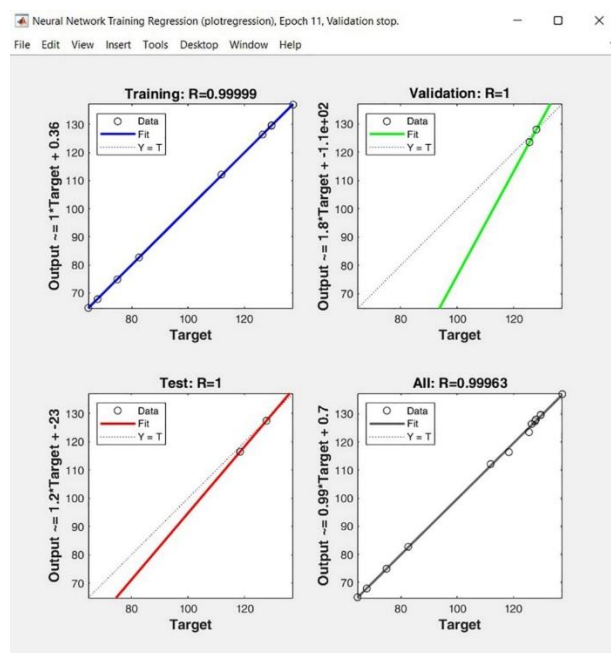


Figure 9. YSA regression results for Turkey

As a result of the training, it was predicted that the quality of life index of Turkey was 127.60 in the scenario where there was no pandemic. As a result of the pandemic, the quality of life index of France was published as 127.10 In Figure 10, retrospective estimates of the network are given.

The regression values of the training, testing, and verification stages of the ANN established for Greece are shown in Figure 11.

As a result of the training, it was predicted that the quality of life index of Greece was 134.30 in the scenario where there was no pandemic. As a result of the pandemic, the quality of life index of France was published as 133.07 In Figure 12, retrospective estimates of the network are given.

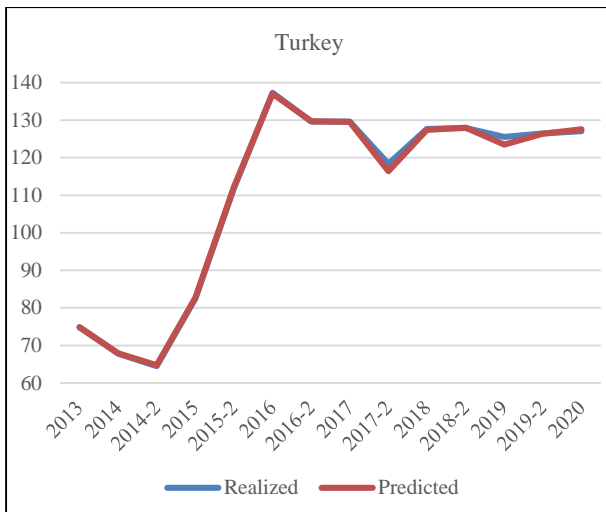


Figure 10. Predicted-realized comparison of Turkey

quality of life is Germany. Germany is followed by Italy, France, Greece and Turkey respectively.



Figure 12. Predicted-realized comparison of Greece

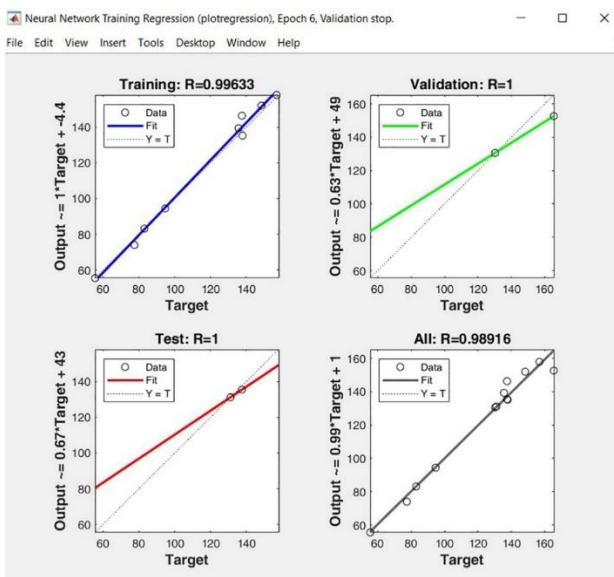


Figure 11. YSA regression results for Greece

Table 4. ANN performance analysis by country

Countries	R	R ²	RMSE	MAPE
Germany	0,99	0,9801	2,16	0,76
France	0,99	0,9801	2,1	1,03
Italy	0,997	0,994009	2,53	1,21
Turkey	0,999	0,998001	0,78	0,36
Greece	0,989	0,978121	4,75	2,21

Table 5. Country impact rates from COVID-19.

Countries	Realized	Predicted	Difference	Difference (percent)
Germany	179,78	229,906	50,12	27,88
Italy	140,76	152,024	11,26	8
France	153,95	157,69	3,74	2,42
Greece	133,07	134,302	1,23	0,92
Turkey	127,1	127,601	0,5	0,39

3.3. Analysis of the ANN

The most accepted performance criteria for ANN analysis in the literature are R2, MSE, RMSE and MAPE values [11,12]. In this study, ANNs established with R, R2, RMSE and MAPE Criteria were tested. Table 4 lists the performance analyzes for five countries.

As a result of the analysis, it was found that the outputs of the network give high accuracy.

The comparisons of the quality of life indexes of 5 countries, estimated by ANN and realized in 2020, and country rankings are shown in Table 5.

As can be seen from Table 5, by comparing the predicted and realized quality of life indices for 2020, it is seen that the country most affected by the pandemic in terms of

5. CONCLUSION

In a study conducted in 2017, the effectiveness of countries in health expenditures was investigated using data envelopment analysis. As a result of this research, Turkey and Greece were found to be effective in health expenditures, while Germany, France and England were found to be ineffective [13]. Demirci et al., in their study in 2020, stated that Turkey, Greece and Italy have efficient health systems, but Germany and France do not have an efficient health system [14]. The results obtained in this study are similar to the results of the two studies mentioned above.

Cinel stated in a 2020 study that the impact of pandemics on health is relatively limited and its economic consequences are more long-lasting and destructive [15].

Based on this discourse, the export difference of countries between 2019 and 2020 was investigated using data obtained from the OECD [16]. These export differences are calculated as a percentage. As a result of the ranking, it was seen that France was the country with the highest falling export rate. France was followed by Italy, Germany, Greece and Turkey, respectively.

It is seen that both the confidence in the health system and the economic indicators show similarities with the impact rate of the COVID-19 pandemic on the quality of life achieved as a result of this study.

Possible future scenarios can be revealed with different estimation methods, but in this study, alternative past estimation has been made, not future estimation. Just as people were not ready, states were caught unprepared for the pandemic. For this reason, the pandemic has caused both social and economic crises.

The quality of life consists of a general index containing six indices and countries are affected by these indices at different rates. As a result of this study, countries are advised to take measures to increase indices that reduce the quality of life and to be prepared for new pandemics that may occur. For other studies, researchers are advised to study countries separately under the indexes of real estate, pollution, health, crime, traffic and cost of living. In this way, the areas where countries have succeeded and failed within the scope of the fight against COVID-19 will be seen.

REFERENCES

- [1] Z. Demir, **Estimation of the brightness and coverage of aluminum paint by artificial neural networks and multiple regression methods**, Master Thesis, Sakarya University, Institute of Science, Industrial Engineering 2019.
- [2] M. F. Keskenler, E. F. Keskenler, "Geçmişten günümüze yapay sinir ağları ve tarihçesi", *Takvim-i Vekayi*, 5(2), 8-18, 2017.
- [3] K. K. Çevik, E. Dandıl, "Yapay sinir ağları için net platformunda görsel bir eğitim yazılımının geliştirilmesi". *Bilişim Teknolojileri Dergisi*, 5(1), 19-28, 2012.
- [4] F. Sönmez Çakır, **Analysis of financial time series with model hybridization**, Gebze Technical University, Institute Of Social Sciences, Business Administration 2018.
- [5] D. Saatçioğlu, N. Özçakar, "Yapay Sinir Ağları Yöntemi ile Aralıklı Talep Tahmini", *Beykoz Akademi Dergisi*, 4(1), 1-32, 2016
- [6] C.D. Lewis, **Industrial and Business Forecasting Methods**, Butterworths Publishing, Londra, U.K., 1982.
- [7] Internet: Numbeo, <http://www.numbeo.com.12.11.2021>.
- [8] B. Ataseven, "Yapay Sinir Ağları ile Öngörü Modellemesi", *Öneri*, 10 (39), ss. 101-115, 2013
- [9] M.S. Gaya, M.U. Zango, L.A. Yusuf, M. Mustapha, B. Muhammad, A. Sani, A. Tijjani, N.A. Wahab, M.T.M. Khairi, "Estimation of turbidity in water treatment plant using Hammerstein-Wiener and neural network technique" *Indonesian Journal of Electrical Engineering and Computer Science*, 5(3), 666-672, 2017.
- [10] N. Aksu, K. Uçan, "Zaman ve Konum Girdileri Kullanılarak Yapay Sinir Ağlarıyla Referans Evapotranspirasyonun Tahmin Edilmesi" *El-Cezeri Journal of Science and Engineering*, 3(2), 204-221, 2016.
- [11] V. Cho, "Comparison of three different approaches to tourist arrival forecasting", *Tourism Management*, 24, 323-330, 2003
- [12] G. Zhang, M.Y. Hu, "Neural Network Forecasting Of The British Pound/Us Dollar Exchange Rate", *Omega*, 26(4), 495-506, 1998
- [13] C. Boz, E. Önder, "OECD Ülkelerinin Sağlık Sistemi Performanslarının Değerlendirilmesi", *Sosyal Givence*, 11, 24-61, 2017
- [14] Ş. Demirci, M. Konca, G. İlgün, "Sağlık finansmanının sağlık sistemleri performansına etkisi: Avrupa Birliği üyesi ve aday ülkeler üzerinden bir değerlendirme", *Sosyoekonomi*, 28(43), 229-242, 2020
- [15] E.A. Cinel, "COVID-19'un küresel makroekonomik etkileri ve beklentiler", *Politik Ekonomik Kuram*, 4(1), 124-140, 2020
- [16] Internet: OECD, <https://data.oecd.org>, 30.11.2021.

An Information System Proposal for Determining of Material Picking Route and Picking Box Quantity in Warehouse

Volkan ÇETİN¹, Başak GÖK², Hadi GÖKÇEN³

¹ Department of Management Information Systems, Institute of Informatics, Gazi University, Ankara, Turkey
² Department of Management Information Systems, Faculty of Applied Sciences, Gazi University, Ankara, Turkey
³ Department of Industrial Engineering, Faculty of Engineering, Gazi University, Ankara, Turkey
volkan.cetin3@gazi.edu.tr, basagok@gazi.edu.tr, hgokcen@gazi.edu.tr

Abstract— The primary focus of lean thinking is to prevent waste. Overproduction, unnecessary transportation, unnecessary movement, unnecessary waiting, unnecessary processing, excess inventory and errors are defined as 7 wastes. It is aimed to prevent waste by finding the number of picking boxes along with the shortest path in the warehouse of the personnel who will pick the materials to be sold or produced in the warehouse. Within the scope of the study, an information system design was made to find the shortest path in warehouse activities and to determine the number of picking boxes and their in-box sequences. It is expected that the information system to be developed will reduce the company's losses due to unnecessary movements and provide a competitive advantage.

Keywords—Information system, warehouse management, determining of material picking route, determining of material picking box quantity

1. INTRODUCTION

Businesses focus on reducing product costs in order to make improvements on the product price in the competitive environment. Keeping the storage and material transfer movements at a minimum level reduces the product cost and this provides a competitive advantage to the enterprises.

One of the main activities of warehouse management is the preparation and delivery of raw materials, parts, semi-materials and products for production or sale. Considering all the costs related to warehouse operations, material Picking activities constitute 55%-65% of them [1]. Fast and correct material preparation activities for production line or customers not only prevent the production line or customer from waiting for materials, but also ensure efficient use of the workforce in the warehouse.

The main activities in a warehouse are the storage (stocking) of the material in suitable conditions and its

transportation (handling) from one point to another point. For the preparation of raw materials, parts, semi-materials and products for production or sale from the warehouse, which is one of the handling operations, the picking activity of materials in the warehouse stock areas comes first in the warehouse activities. This process constitutes the largest part of the storage activities. Improvements to be made regarding the picking of materials will directly affect the entire activity as a major improvement.

In this study is presented the design of an information system called Material Picking Information System, which is planned to be developed for improvement the material picking route from the warehouse and determining the number of picking boxes needed before the picking activity. With the designed information system, it is expected that the losses of the enterprise due to unnecessary movements will decrease and it will provide a competitive advantage.

2. CONCEPTUAL FRAMEWORK

2.1. Role of Information System in Plant

An information system is a system that takes data sources as inputs, processes them, and produces information products as outputs. In this case, while the information system achieves this, the question arises “What elements and activities does it include?” [2].

Successful management of information systems and technologies provides great opportunities and advantages for businesses. Shortening of the product development period and rapid introduction of the new product to the market, expanding the market, increasing the product range, decreasing the costs, adapting quickly to the changing economic conditions, and meeting the changing customer demands and requirements accurately and on time can be given as examples [3].

Information systems have been one of the main functional areas in businesses as important as accounting, finance, marketing or human resources [2].

2.2. Warehouse Management

Storage problems are generally grouped as warehouse design and warehouse operations in the literature [4]. Warehouse design includes overall structure, sizing and dimensioning, department layout, equipment selection, and operation strategy. And warehouse operation includes receiving and shipping, storage and order picking.

As an example of warehouse problems; classical discrete optimization problems are expressed as traveling salesman problem (TSP) and box packing problem (BPP). These two problems are easy to formulate, but difficult to solve. BPP is often used to decide the position of materials in the warehouse or on the pallet. The most labor-intensive activity in the warehouse is order picking. There are many studies on order picking policies in the literature [5].

As an example of order picking algorithms, Armstrong, Cook, and Saipé (1979) developed an order picking algorithm that minimizes the total picking time with a mixed integer linear programming model [6]. They used artificial orders for exact programming. They have considered field picking operations for the distribution center. There are also studies in the literature that emphasize the usefulness of the heuristic algorithm in picking material from the warehouse [7-9]. Researchers aimed to minimize the total material picking route. Ratliff and Rosenthal (1983) presented a picking route algorithm that minimizes picking time for a warehouse system with two cross aisles, front and back [10]. They developed a robust procedure for determining a picking route using a simple step structure. Tsai, Liou, and Huang (2008) consider not only the lap distance but also the early and late penalty costs in the models they propose on the picking of different material lots [11].

In 2016, Lin, Kang, Hou, and Cheng proposed a model that combines order grouping and order picking route in the same model. They developed a model for order grouping and order picking using the Particle Swarm Optimization (PSO) algorithm [12]. Erdoğan (2017) developed an MS Excel VBA Macro based on the large neighborhood search heuristic algorithm. This study helps to decide the quantity of the box type and the algorithm to put them in the designated boxes [13].

As can be seen in the literature, efforts to find the best solution still keep going and businesses continue to seek optimum solutions to gain competitive advantage.

3. METHOD

3.1. Definition of the Problem and System Expectations

In a warehouse system with parallel aisles and starting and ending points in the same location, the activities of preparing raw materials, parts, semi-finished products and finished products for production or sale, defined as stock materials, are among the daily routine tasks. In these routine activities carried out, the determination of the picking route is made based on the intuition and experience of the personnel. For this reason, there is time loss and wastage in the routine processes carried out in the warehouse. At the same time, due to the inability of the personnel to determine the exact amount and dimensions of the picking box they need at the beginning of activity, the process of taking or replacing a new picking box occurs by returning to the starting point many times. These repetitive processes increase the time and labor losses of the enterprise. In this study, in order to minimize the mentioned losses, the shortest possible routing, solving the problem of determining the rectangular box type and quantity required in packaging and the development of the simulation including the placement of the material in the box presenting the warehouse personnel including a data flow diagram has been developed an information system design before starting the material picking activity from the warehouse.

3.2. Description of Current System and Limitations

The storage system (warehouse) has a width of 26 m and a length of 16 m. There are 5 parallel aisles and 4 aisles perpendicular to these aisles and 72 defined stock locations in the warehouse. Each warehouse consists of 6 floors and the warehouse has a total height of 5 m. Warehouse corridors are 2 m wide. The width of the stock areas is 2x2 m. As for the coding of stock locations, the first subscript indicates the column number, while the second subscript represents the shelf row number, as shown in Figure 1.

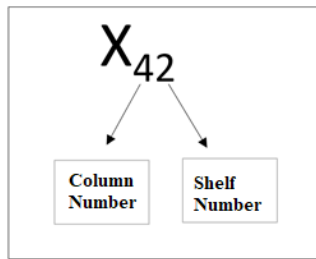


Figure 1. Stock address coding

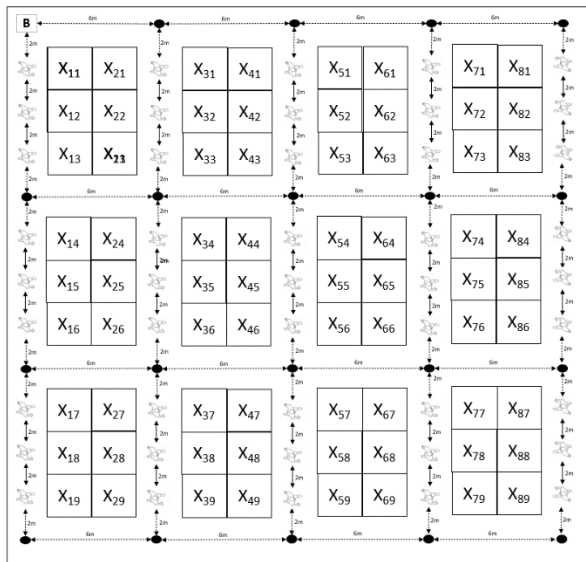


Figure 2. Warehouse map

As can be seen in Figure 2, X_{42} is located on the 2. shelf of the 4th column.

In the warehouse system, the shortest material picking route, number of picking boxes and in box placement

simulation are planned to work according to 20 kinds of materials to be picking from stock areas. The reason for this is to prevent the prolongation of the solution time based on the MS Excel Solver infrastructure to be used in the solution. It is thought that this variety and solution time can be increased in different software infrastructures. In the infrastructure of the system to be created, there will be 72 types of materials at each stock address.

The personnel who will start picking materials from the starting point (B) will return to point (B) after the picking.

By running the information system, the required number of boxes, the contents of the box and how it will be placed will be presented to the warehouse personnel with simulation.

The materials and quantities in the picking list will be made into a connected lot and the picking will be carried out. In other words, it is requested that all quantities of a material to be picked are contained in a single box.

Picking box sizes are standard and 3 types. They are respectively in width, length and height 0.40m. x 0.60m. x 0.28m., 0.60m. x 0.80m. x 0.32m. and 0.61m. x 0.81m. x 0.52m. Materials that do not fit in any picking box are transported on pallets. The number of picking boxes are 5 for each size and 15 in total. Models of Picking boxes are given in Figure 3.

Data Flow Diagrams (DFD), were used in the development of the information system.

DFD is a very popular tool used in structural system development and provides a graphical representation of the system. With DFD, it is possible to describe any complex system using 4 symbols [14].

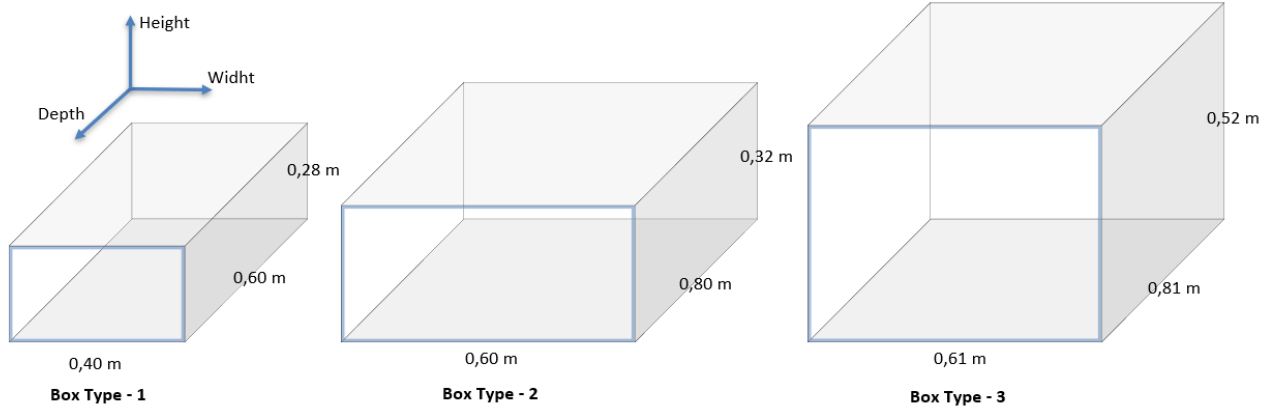


Figure 3. Picking boxes types (Container types)

3.3. Creating Context Diagram

Considering the limitations and expectations of the system, the Context Diagram was created as in Figure 4. Inputs of Material Picking Information System; transport box types (container types) information, warehouse address and distance between addresses, material stock address information, material size information, sequence information according to pieces, the required transport box and in-box layout calculation request, 3D in-box layout request and report are requests.

Outputs of Material Picking Information System; It consists of the order of picking the materials from the stock places, the number of boxes, the 3D in-box layout list and other reports that give the shortest path. Warehouse

supervisors and warehouse operators constitute the external units. The warehouse supervisor is the administrative officer of the warehouse activities and is in the managerial position. Planning and organizing the activities during the day are among the primary responsibilities of the warehouse supervisor. Warehouse operators are employees who actually carry out the activities with the plan they received from the warehouse supervisor.

3.4. Creating 0-Level Data Flow Diagram

0-Level Data Flow Diagram was created as in Figure 5. There are 9 processes and 9 datastores in the system. 2.0 Process sublevel diagram given in Figure 6.

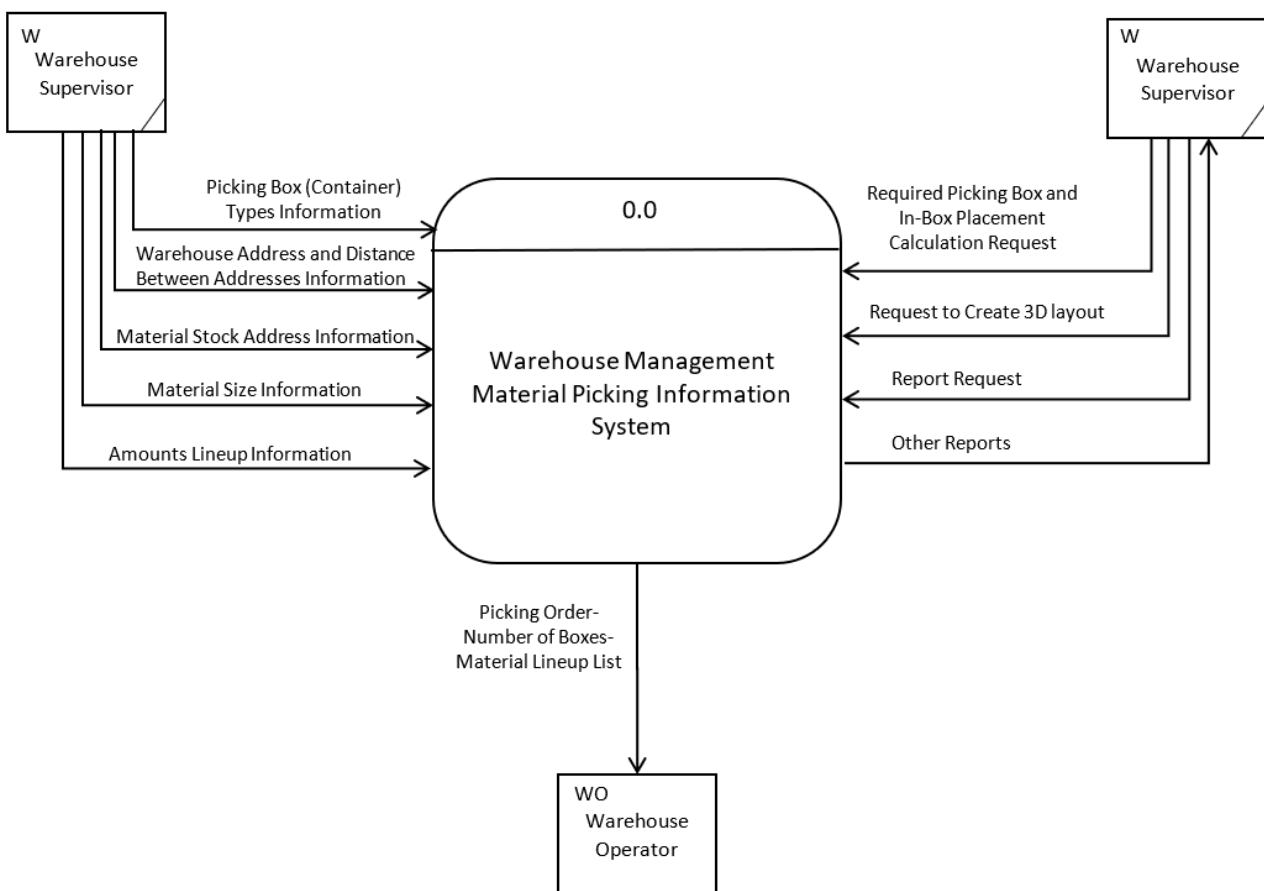


Figure 4. Context diagram

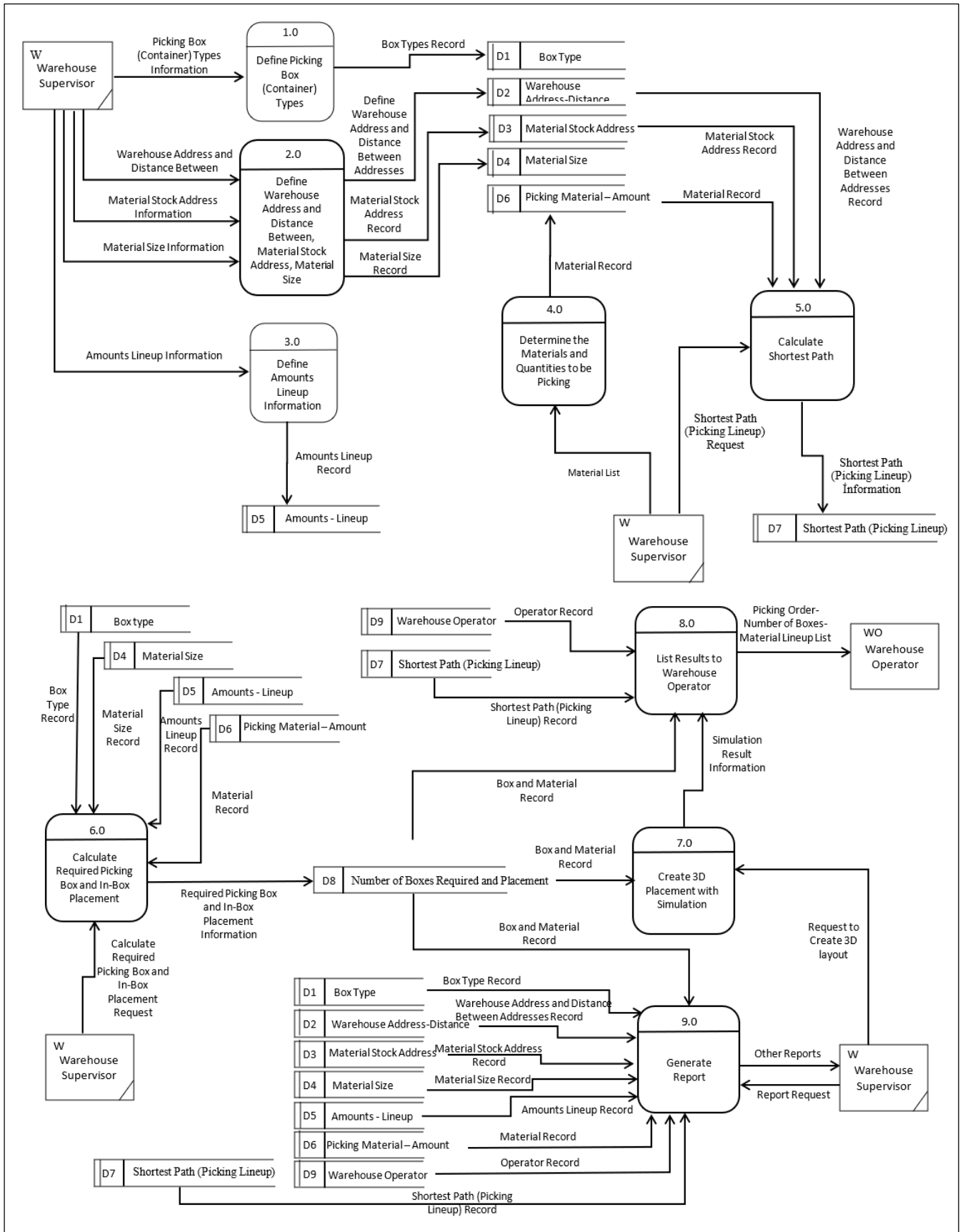


Figure 5. 0-Level data flow diagram

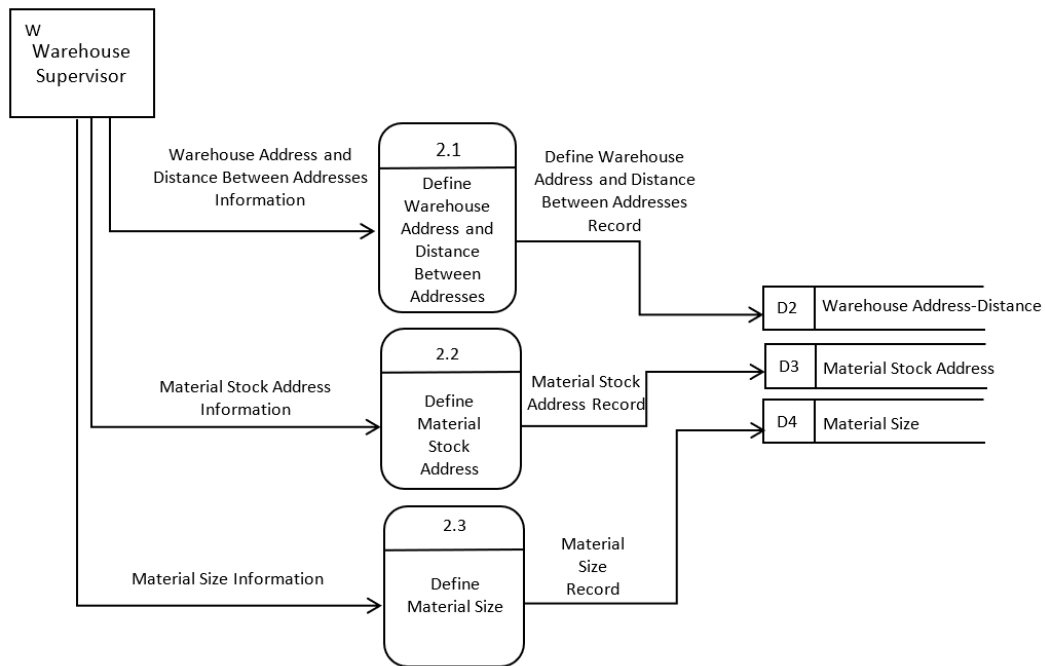


Figure 6.2.0 Process sublevel diagram

3.5. Creating System Data Store

In order to provide information entry to the 6 datastores in the system, an interface was designed and the system defined the transport box type (container type) information, material stock address information, material size information, warehouse address-distance information, box sequence information and warehouse operator information.

3.6. Some Information System Results

The materials and quantities in Table 1 were chosen as examples for the calculation of the number of picking boxes and the arrangement within the box.

When the Material Picking Information System is run to find the shortest path in the material picking route, the material picking lineup result in Figure 7. is obtained.

Table 1. Materials and picking amounts list

Material	Picking Amount
STK-882-C1-DE5944	8
STK-721-B4-GY2652	12
STK-500-A4-AB3817	9
STK-323-B1-GY8424	10
STK-967-C5-RT7031	10
STK-100-B2-DE3240	2
STK-240-B9-AN4800	10
STK-208-A7-AC5790	10
STK-684-A3-FT1640	6
STK-232-C5-AC2271	2

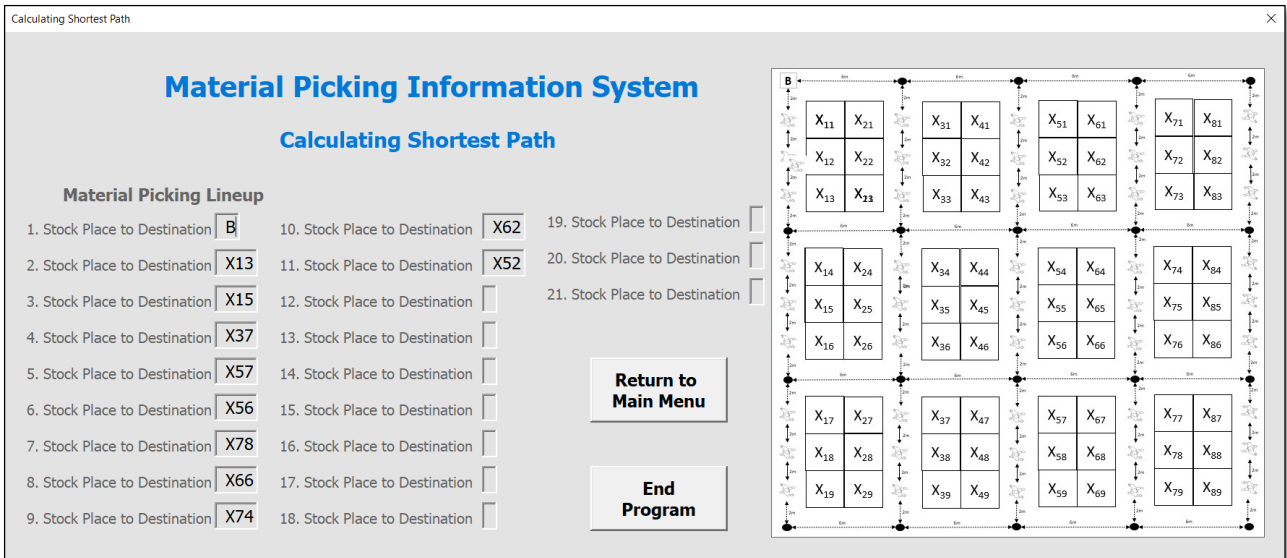


Figure 7. Example shortest path screen output

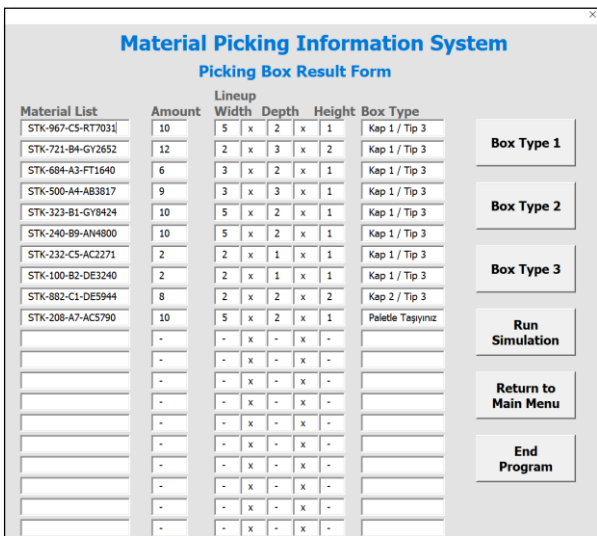


Figure 8. Screen output of number and type of picking boxes required

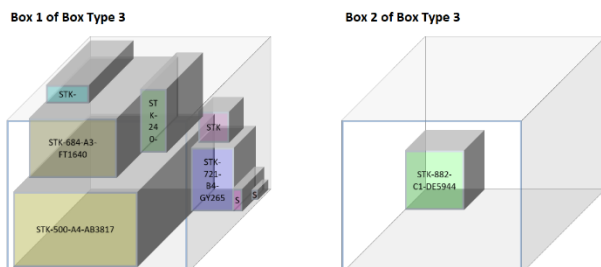


Figure 9. In-box placement screen output

Considering the amount of materials, the Material Picking Information System was run to find the number of Picking boxes needed and the result in Figure 8. was found.

In-box layout image as a result of the 3D supported simulation created by the Material Picking Information System for the sample collection list is given in Figure 9.

4. CONCLUSION

Many businesses are focused on their costs because they cannot increase their product prices in order to maintain the economy in their business. It can be aimed to reduce costs by reducing the losses in the activities of preparing raw materials, parts, semi-materials or products for production or sale in warehouses. In this study, an information system analysis and design, which is planned to be developed, was created in order to establish the shortest material picking route and to determine the number of materials picking boxes. As the final of the ongoing work, it is planned to complete the information system and present a two-stage result to the users. This study is aimed to increase productivity by reaching an optimum or near-optimal result through the Material Picking Information System, which was designed with a systematic approach to material picking operations that were previously carried out based on employee experience and intuition.

REFERENCES

[1] C. Theys, O. Braysy, W. Dullaert and B. Raa, "Using a TSP Heuristic for Routing Order Pickers in Warehouses", *European Journal of Operational Research*, 200(3), 755–763, 2010.

- [2] M. Şahin, **Management Information Systems**, Anadolu University, Eskişehir, Turkey, 2006.
- [3] A. Schultheis, Sumner and Robert, **Management Information System: The Manager's View**, Irwin/McGraw Hill, 4th ed., 1998.
- [4] J. Gu, M. Goetschalckx and L. F. McGinnis, "Research on Warehouse Operation: A Comprehensive Review", *European Journal of Operational Research*, 177, 1–21, 2007.
- [5] F. Yener, H. R. Yazgan, "Optimal Warehouse Design: Literature Review and Case Study Application", *Computers & Industrial Engineering*, 129, 1-13, 2019.
- [6] R. D. Cook and A. L. Saipé, "Optimal Batching in a Semi-Automated Order Picking System", *Journal of the Operational Research Society*, 30(8), 711–720, 1979.
- [7] D. E. A. Elsaye, "Algorithms for Optimal Material Handling in Automatic Warehousing Systems", *International Journal of Production Research*, 19(5), 525–535, 1981.
- [8] G. R. Stern and E. A. Elsayed, "Computerized Algorithms for Order Processing in Automated Warehousing Systems", *International Journal of Production Research*, 21(4), 1983.
- [9] E. A. Elsayed and O. Unal, "Order Batching Algorithms and Travel-Time Estimation for Automated Storage/Retrieval Systems", *International Journal of Production Research*, 27(/), 1097-1114, 1989.
- [10] H. D. Ratliff and A. S. Rosenthal, "Order-picking in a Rectangular Warehouse: A Solvable Case of the Traveling Salesman Problem", *Operations Research*, 31(3), 507–521, 1983.
- [11] C. Y. Tsai, J. J. H. Liou and T. M. Huang, "Using a Multiple-GA Method to Solve the Batch Picking Problem: Considering Travel Distance and Order Due Time", *International Journal of Production Research*, 46(22), 6533–6555, 2008.
- [12] C. C. Lin, J. R. Kang, C. Y. Cheng, "Joint Order Batching and Picker Manhattan Routing Problem", *Computers and Industrial Engineering*, 95, 164–174, 2016.
- [13] Internet: G. Erdoğan, CLP Spreadsheet Solver, Released on 07.09.2017, <https://researchportal.bath.ac.uk/en/publications/clp-spreadsheet-solver>, 31.05.2021.
- [14] H. Gökçen, **Information Management Systems: Analysis and Design**, Afşar Matbaacılık, Ankara, 2011.

Decision Support Systems in Engineering: A Bibliometric Analysis and Systematic Review

Musab Talha Akpınar¹, Mehmet Atak², Uzeyir Fidan³

¹Ankara Yıldırım Beyazıt University (Management Information Systems, Ankara Yıldırım Beyazıt University, Turkey)

²Gazi University (Industry Engineering, Gazi University, Ankara, Turkey)

³Uzeyir Fidan (Computer Programming, Usak University, Uşak, Turkey)

takpinar@ybu.edu.tr, matak@gazi.edu.tr, uzeyir.fidan@usak.edu.tr

Abstract— This paper uses a bibliometric approach to examine the growth of and changes in the Decision Support Systems (DSS) field over the past forty years. One of the best bibliographic databases which is Web of Science (WOS) provide valuable information on academic disciplines as it contains both the articles published and the articles cited. This article aims to characterize scientific production regarding the development of Decision Support Systems based on multicriteria methods for the multiple sectors. Using the bibliometrics technique, department segments applying multicriteria methods were identified, and an examination was conducted of the multicriteria methods employed, the countries where these studies were carried out, the most cited authors, the number of papers published per year, and the most frequent keywords. The scientific literature is reviewed in a systematic way using a comparative analysis of existing tools. The results indicate that there are relatively many developments in determined DSS categories. However, even in these categories some issues and gaps still remain, so further improvements are needed. The analysis indicates a relatively low level of technology readiness of tools and concepts presented in academic literature. This signifies a need to move beyond the conceptual stages toward demonstration and validation in realistic, operating environments. Findings indicated that the Analytic Hierarchy Process was the multicriteria method employed most often.

Keywords— decision support system (dss), bibliometric analysis, multicriteria methods, systematic review

1. INTRODUCTION

The concept of bibliometrics was first introduced by Alan Pritchard [1] in "Statistical Bibliography or Bibliometrics?" used in his article. It is defined as the application of mathematical and statistical methods to books and other communication tools. The word "bibliography", which dates back to the 1920s, was first used by Wyndham Hulme [2]. Bibliometric is the process of converting publications published in a certain field into a data set, performing analyzes with various mathematical and statistical methods, and obtaining information such as authors, subjects, years, and countries based on publication information [3].

Bibliometric analysis, which aims to determine the profile of studies in a science or discipline, shows the systematic development of scientific knowledge production [4]. It is an extremely important type of research in terms of determining the focus and deficiencies of the knowledge production process [5]. As "Decision Support Systems Studies" increased, it became difficult to monitor the development of the discipline through scientific publications. In this respect, Web of Science (WoS) serves

as the most comprehensive database for researchers. Scientific mapping method, which is one of the bibliometric analysis methods, is expressed as the creation and processing of data sets of publication information obtained from various databases, the creation of networks and their interpretation by an expert [6][7]. Scientific Mapping Analysis, which is applied according to the data obtained from bibliographic sources, allows to see the change in the conceptual structure of a particular scientific field over time.

On the other hand, for more than 35 years, information system researchers and technicians have constructed and studied Decision Support Systems (DSS). DSS advancements started in the late 1960s with the creation of model-oriented DSS, followed by theoretical developments in the 1970s and the deployment of financial planning systems and Group DSS in the early and mid 1980s. DSS have been employed from the beginnings of Executive Information Systems, OLAP and Business Intelligence, and eventually, in the mid-1990s, with the installation of Web-based DSS.

MIS was first focused on providing managers with structured, monthly reports in the early 1960s. The majority of the data came from accounting and transaction systems. In the late 1960s, a new sort of information system, model-oriented DSS or management decision systems, became practicable. According to Peter Keen and Charles Stabell, the concept of decision support evolved from theoretical studies of organizational decision-making conducted at the Carnegie Institute of Technology in the late 1950s and early 1960s, as well as technical work on interactive computer systems conducted primarily at the Massachusetts Institute of Technology in the 1960s [9].

Around 1970, management journals began to publish studies on managerial decision-making systems, strategic planning systems, and decision support systems, according to Sprague and Watson [111]. Scott Morton and colleagues, for example, released a number of decision support studies in 1968. Ferguson and Jones published an article in *Management Science* in 1969 on a computer-aided decision system. Michael S. Scott Morton's seminal book *Management Decision Systems: Computer-Based Support for Decision Making* was released in 1971. Scientific mapping analysis offers a method especially for researchers who conduct bibliometric studies on a specific scientific field [8]. Such a study will allow data on the DSS research to be included in as many samples as possible.

Bibliometric is to analyze this discipline with mathematical and statistical methods by using the data of publications belonging to a discipline. The data included in the analysis, on the other hand, consist of information such as the authors, citations, sources, subjects, countries or institutions of their production, and publication years [10]. Bibliometric analysis, on the other hand, provides the opportunity to reveal the general view of a particular scientific discipline by examining such data from scientific publications [3]. In this context, the use of bibliometric analysis is seen as a necessity in classifying and organizing data, making it simpler and more understandable today, when the volume of data has reached its maximum level and becomes more complex. In other words, scientific mapping analysis allows the monitoring of the conceptual structure and development of a scientific field and the limitation of research areas [11, 12].

Bibliometric analysis is one of the most important elements to examine existing research in a particular field of research [13]. These studies are meta-analysis and structured literature review [14] and mapping methods that show how sources are interconnected [15]. The clustered results will be represented by mapping [16]. Today, classical literature studies are criticized for taking shape (subjective / qualitative) according to the researcher's view, contrary to the bibliometric method [17]. The bibliometric method is; It uses a systematic, transparent and repeatable (objective / quantitative) approach for definition, evaluation and monitoring of the research [13]. In this paper, the bibliometric method consists of methods such as

citation analysis, co-author and common word analysis shown on Table 1.

Table 1. Used analyses in scope of the paper

Analysis	Definition	Aim	Source
Citation Analysis	Citation analysis is a mathematical and statistical analysis method that calculates the frequency of citations by counting the titles, authors and sources.	The most cited authors in this specified field are considered the most important authors, and the most cited works are considered the most valuable works.	[18], [19]
Co-Author Analysis	Co-author analysis examines the social networks created by scientists collaborating on scientific research.	Co-authorship in research creates social and professional networks of scientists in the academic community and enlarges the field of knowledge in disciplines.	[20], [21]
Common Word Analysis	Keyword analysis is an analysis technique that tries to understand the conceptual structure of the field by establishing relationships between words in scientific studies.	This analysis focuses on the words commonly used in scientific articles, suggesting that they will form a conceptual structure. These words are searched in document titles, keywords, abstracts or words in full texts.	[22]

In this study, scientific mapping will be done using the determined methods of bibliometric analysis. The work flow process of the study is shown in the Figure 1. Firstly, a systematic and reproducible approach is applied to gather and filter the data sample. Subsequently, bibliometrics is applied to investigate the collected data sample. The utilization of this method results in a global overview of DSS-related tools and concepts. In this type of analysis, quantitative data regarding scientific production, such as the number of documents, authors' contributions, and occurrences of keywords are determined. The obtained information allows for identifying various collaboration networks in order to indicate authors, countries, as well as institutions significantly involved in the analyzed domain.

The aim of this study is to reveal the bibliometric map of the studies on "Decision Support System" + "Decision Support Systems" in the field of engineering studies. The research questions of this paper are given below.

- RQ1. What is the scientific map of the "Decision Support Systems" articles?
- RQ2. How is the distribution of articles by years?
- RQ 3. What is the distribution of articles according to journals?
- RQ 4. How is the distribution of keywords used in the articles?

- RQ 5. Who are the center, peripheral or intermediary researchers in the studies on " Decision Support Systems " in the field of Industry Engineering?
- RQ 6 How did the " Decision Support Systems" concept spread in the engineering field of research?
- RQ 7. What is the structure of the scientific community in the " Decision Support Systems " field?

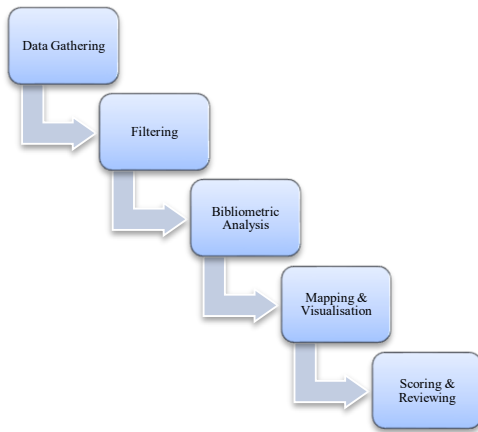


Figure 1. Work flow of the paper

2. USAGE

The purpose of bibliometric data is to review and quantify the impact of a scientific field. The field of bibliometric studies publication patterns by using quantitative analysis and statistics [11]. In this paper, searching the correct bibliometric information involved multiple search options such as using keywords like "Decision Support Systems" in order to retrieve the most relevant information. In this section, the stages in the analysis of the study will be discussed step by step. These headings are data acquisition, data preparation, analysis, viewing and interpretation.

2.1. Data Acquisition

Web-based online bibliographic databases as ISI, Web of Science, Scopus, CiteSeer, Google Scholar or NLM and MEDLINE or others are common data sources for bibliometric research [23]. There are three major bibliometric databases investigated as follows: Google Scholar, Scopus and the Web of Science [39]. It is due to multiple factors, including its scholarly citations, The Web of Science, an online database of bibliographic information produced by Thomson Reuters draws its real value from the scholarly citation index at its core [40]. The tools and content of the Web of Science are trusted by a vast amount of institutions worldwide. The Web of Science's content and tools are trusted by nearly 7000 of the world's leading scholarly institutions responsible for scientific policy making. [41] and for its ability to refine search options within its own database, this allowed us, when using keywords to search for relevant information for this paper, to narrow the search results with a multitude of options such as documentation type, language, times cited, research domains, countries, authors and so on. All these

capabilities made the Web of Science database an extremely effective tool in searching for relevant information quickly in a format of our choosing. The articles needed for this analysis and their bibliography are available on the web of knowledge (<http://apps.webofknowledge.com>) site. The analysis file obtained with the help of the web of knowledge website was examined by frequency analysis and spelling mistakes, coding errors and spelling differences were corrected.

2.2. Data Preparation

Bibliometric analysis and scientific mapping analysis cannot be directly applied to data obtained from bibliographic databases. A pre-processing stage is required to improve the quality of the data and determine the time frames for analysis by removing error-containing data such as duplication or misspelling [8].

In this research, books, research notes, etc. Instead of studies, only articles in academic journals are targeted. The reason for this is that the articles published in these journals have proved the accuracy of the information they carry because they are subjected to a peer review process [19]. All studies determined within the scope of the research from January 1980 to October 2021 were used as data. The titles, keywords and abstracts of the articles in the selected journals were scanned with the keyword "Decision Support System" + "Decision Support Systems" 5752 articles were identified (Searches title, abstract, author keywords, and Keywords Plus.) The obtained sources were listed and checked under 54 different titles, from title to publication years, from keywords to references. Spelling mistakes, coding errors were corrected and the data was made ready for analysis.

2.3. Analysis

The clean data obtained after the data preparation phase are scientific publications in the web of science (wos) database and include the concept of "decision support system" + "decision support systems" "at least once in their title, summary and keywords. The bibliometric analysis of the study was done with "vosviewer" software. As a result, 5752 articles were included in the analysis and all these articles were published in 788 different journals. 498 (8.65%) of these articles were published in expert systems with applications and 228 (3.96%) were published in environmental modelling software. These journals are followed by computers industrial engineering, artificial intelligence in medicine, international journal of production research, IEEE access, computer methods and programs in biomedicine, water resources management. The journals in which the articles were published are shown figure 2.



Figure 2. Journal in which articles were published

When we look at the affiliations, 75 (1.3%) of the articles were published from Indian Institute of Technology System Iit Florida. This is followed by State University System of Florida with 72 articles, followed by Hong Kong Polytechnic University, Centre National De La Recherche Scientifique Cnrs, Consiglio Nazionale Delle Ricerche Cnr, University of Texas System. The distribution of publications by affiliations is also shown in the Figure 3.

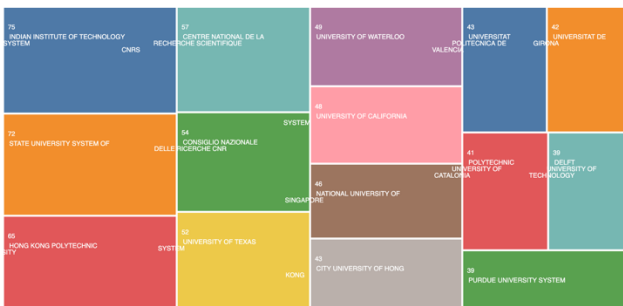


Figure 3. Distribution of publications by affiliations

The first of the 5.752 articles accessed is the paper named "A Decision Support System for Managing the Transportation Needs of a Large Corporation" written by Bezael [43]. From that study, the number of studies carried out in the field increased year by year and finally 478 studies were completed in 2020. It is shown that Figure 4, the article numbers have increased since 2002.

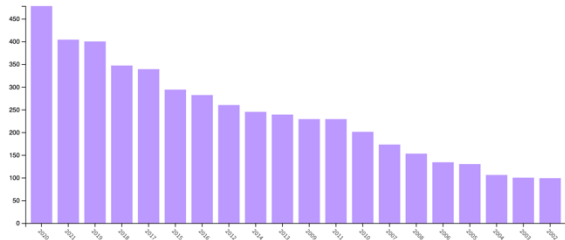


Figure 4. Distribution of publications by years

Until 2021, among 5.752 articles contributed by 16.058 different researcher, the author with the most articles is "Poch M." with 27 articles. After him "Huang GH" with 25 studies and, Comas J and Hipel KW is following with 20 studies. The list of authors who have more than 15 publications in the field is shown in the Table 2.

Table 1. Authors who have more than 20 publications

Numbers	Author Name	Number of Articles
1	Poch M	27
2	Huang GH	22
3	Comas J	20
4	Hipel KW	20
5	Labadie JW	17
6	Chan CW	16
7	Lee S	16
8	Tiwari MK	16
9	Chan FTS	15

3. VISUALIZATION - INTERPRETATION

In this study, network analysis method was preferred to demonstrate the relationship between articles/authors/citations/keywords. The VOSviewer program, which can show the strength of the relationships in the network and automatically identify the clusters in the field, has been selected. 5.754 articles included in the analysis have 182.376 citations totally. Before starting the analysis, the threshold value was determined as 20. At this point, the threshold value indicates the minimum number of times a reference is repeated. Thus, references that were repeated at least 100 times were taken into account in this study.

3.1. Citation Analysis

5754 publications are included in the bibliography of 182.376 articles and the number of citations made to these studies is 111.540. These citations showed a steady increase in parallel with the increase in the studies presented, and 14,333 citations were given in the studies conducted in 2020. The citations are shown in Figure 5.

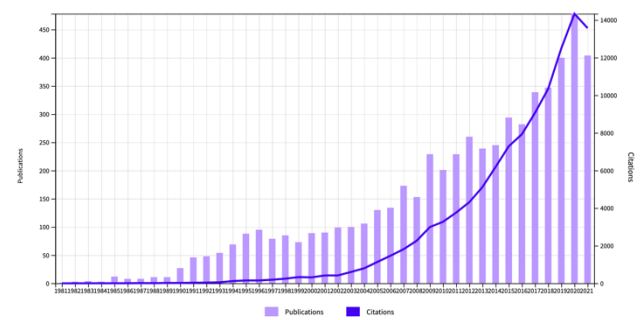


Figure 5. Number of citations

The 10 most cited articles among the 5754 articles within the scope of the study are shown in the Table 3.

Table 3. Most cited articles

No	Titles	References	Year	Times Cited
1	Optimal operation of multireservoir systems: State-of-the-art review	[24]	2004	972
2	APSIM - Evolution towards a new generation of agricultural systems simulation	[25]	2014	700
3	A decision support system for supplier selection using an	[26]	1998	670

	integrated analytic hierarchy process and linear programming			
4	Methods used for the development of neural networks for the prediction of water resource variables in river systems: Current status and future directions	[27]	2010	489
5	Drought modeling - A review	[28]	2011	482
6	Supply chain modeling: past, present and future	[29]	2002	469
7	Meal simulation model of the glucose-insulin system	[30]	2007	452
8	Selecting among five common modelling approaches for integrated environmental assessment and management	[31]	2013	437
9	An overview of time-based and condition-based maintenance in industrial application	[32]	2012	376
10	WEAP21 - A demand-, priority-, and preference-driven water planning model Part 1: Model characteristics	[33]	2005	376

When the 182,376 citations obtained from the articles included in the analysis were examined, it was understood that these citations were made to 5,752 different academic studies. This reveals the total number of studies carried out in this field, in areas that work jointly in this field and in other areas that support this field. In addition, among these DSS studies (5752) in which the references in the bibliography are used, the most cited work is Labadie [24] with 972 citations. This study is followed by Holzworth [25] with 700 citations, and Ghodsypour [26] with 670 citations. The general view of these studies is shown in Figure 6.

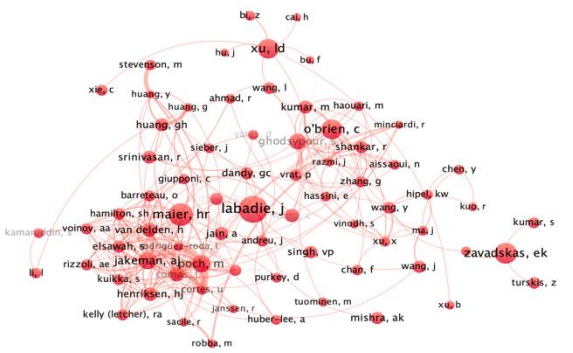


Figure 6. Link of most cited studies

3.2. Co-Author Analysis

Co-author analysis examines the social networks created by scientists collaborating on scientific research [21]. Co-authorship in research creates networks in the academic community and enlarges the field of knowledge in disciplines [34]. These networks provide a carefully documented record of scientists' social and professional networks [35]. Thus, the degree to which the social identity of this discipline is shared is revealed [36]. In addition, since the bibliographic data includes information about the authors' institutional relations and geographic locations, co-author analysis can examine the issues of cooperation at the level of institutions and countries [13].

Figure 7 shows that 43 matches of 16,058 authors with min 2 articles and min 100 citations.

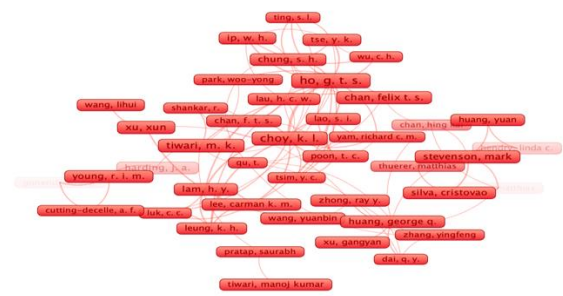


Figure 7. Co-author analysis

3.3. Common Word Analysis

Common word analysis introduced by Callon in 1983 [38] is an analysis technique that tries to understand the conceptual structure of the field by establishing relationships between words in scientific studies. This analysis focuses on the words commonly used in scientific articles, suggesting that they will form a conceptual structure. These words are searched in document titles, keywords, abstracts or words in full texts [22]. While other types of bibliometric analysis try to create a structure by indirectly linking the studies, this analysis method directly uses the cat of the document to establish a meaningful structure [13]. The output of common word analysis is a network of themes and their relationships representing the conceptual structure of a field [37]. The 5752 articles in the data set contain a total of 19,020 different keywords. Cluster maps of 559 words, whose keywords are counted at least 10 times, are shown Figure 8.

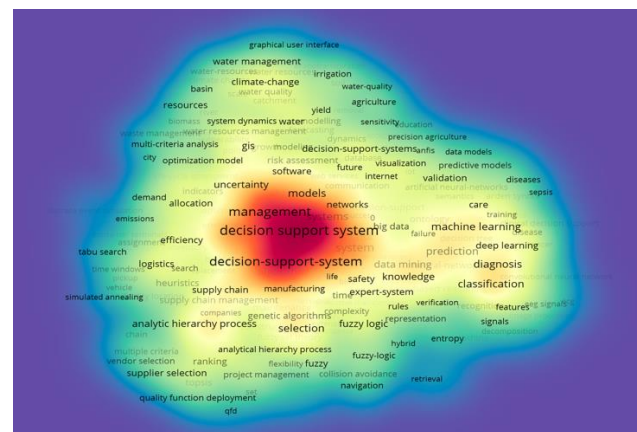


Figure 8. Common word analysis

3.4. Conclusion

In this article, it is aimed to investigate the appearance and structure of academic articles (scanned in web-based online bibliographic database "Web of Science") containing the term "Decision Support System" + "Decision Support Systems". In order to reach this goal, citation, joint citation, and keyword analysis were applied via social network analysis. The findings of this article are

run in the comments section and the overview of these DSS articles over the last 40 years is shown. This research has a few limitations. First of all, the research was conducted from academic journals scanned by "Web of Science". Examining its status in other SSCI and international journals in this field will be able to provide more extensive information on the field. Secondly, only the terms "Decision Support System" + "Decision Support Systems" were chosen as key words in this study. Increasing and diversifying the terms related to this field will help him participate in the analysis in other studies in Engineering field. Moreover, this study can reply without filtering research areas. Finally, this analysis was designed to examine the citations analysis, co-author analysis and common word analysis and bibliographies of the publications in the field defined by the term "Decision Support System". However, conducting a bibliometric matching analysis in which the bibliography of these studies can be defined and the relationship structures between them will be helpful to understand this area.

REFERENCES

- [1] Pritchard A. Statistical bibliography or bibliometrics. *Journal of Documentation* 1969;25: 358-59.
- [2] HULME, E. W. Statistical bibliography in relation to the growth of modern civilization. London, 1923.
- [3] Martinez, M. Angeles, Cobo, Manuel J., Herrera, Manuel ve Herrera Viedma, Enrique (2015), "Analyzing the Scientific Evolution of Social Work Using Science Mapping", *Research on Social Work Practice*, 25(2): 257-277.
- [4] Ruhanen, Lisa, et al. "Trends and patterns in sustainable tourism research: A 25-year bibliometric analysis." *Journal of Sustainable Tourism* 23.4 (2015): 517-535.
- [5] van Raan, A. F. J. (2005), "Measuring science" in H. F. Moed, W. Glanzel, U. Schmoch (Eds.), *Handbook of quantitative science and technology research*, Springer, Netherlands, 19-50.
- [6] Cobo MJ, López-Herrera AG, Herrera-Viedma E, et al. Science mapping software tools: review, analysis, and cooperative study among tools. *Journal of the American Society for Information Science and Technology* 2011;62: 1382-402.
- [7] Noyons E. Bibliometric mapping of science in a sciencepolicy context. *Scientometrics* 2001; 50: 83-98.
- [8] Cobo, Manuel J., et al. "SciMAT: A new science mapping analysis software tool." *Journal of the American Society for Information Science and Technology* 63.8 (2012): 1609-1630.
- [9] Keen, P. G. W. and M. S. Scott Morton, *Decision Support Systems: An Organizational Perspective*. Reading, MA: Addison-Wesley, Inc., 1978.
- [10] Small, H. (1999), "Visualizing Science by Citation Mapping", *Journal of the American Society for Information Science*, 50(9): 799-813.
- [11] Cobo, Manuel J., et al. "Science mapping software tools: Review, analysis, and cooperative study among tools." *Journal of the American Society for information Science and Technology* 62.7 (2011): 1382-1402.
- [12] Noyons, E., H. Moed, and A. Van Raan. "Integrating research performance analysis and science mapping." *Scientometrics* 46.3 (1999): 591-604.
- [13] Zupic, Ivan, and Tomaž Čater. "Bibliometric methods in management and organization." *Organizational Research Methods* 18.3 (2015): 429-472.
- [14] Schmidt, Dieter, Mona Pache, and Martin Schumacher. "The congenital unilateral retinocephalic vascular malformation syndrome (Bonnet-Dechaume-Blanc syndrome or Wyburn-Mason syndrome): review of the literature." *Survey of ophthalmology* 53.3 (2008): 227-249.
- [15] van Leeuwen, Thed N., and Clara Calero Medina. "Redefining the field of economics: Improving field normalization for the application of bibliometric techniques in the field of economics." *Research Evaluation* 21.1 (2012): 61-70.
- [16] Boyack, Kevin W., and Richard Klavans. "Creation of a highly detailed, dynamic, global model and map of science." *Journal of the Association for Information Science and Technology* 65.4 (2014): 670-685.
- [17] Tranfield, D., D. Denyer, and P. Smart. "Undertaking systematic review: developing an evidence-based approach for management research." *Academy of Management Conference*, Denver, CO. 2002.
- [18] Diodato, V. I. "Dictionary of bibliometrics. Environmental scanning in libraries" (issue theme). *Minnesota Libraries, Pb(Winter 1987-1988)*, 376-399. New York: Haworth, 1994
- [19] Ramos-Rodríguez, Antonio-Rafael, and José Ruíz-Navarro. "Changes in the intellectual structure of strategic management research: A bibliometric study of the Strategic Management Journal, 1980–2000." *Strategic management journal* 25.10 (2004): 981-1004.
- [20] Newman, Mark EJ. "Who is the best-connected scientist? A study of scientific coauthorship networks." *Phys. Rev. E* 64.016131 (2001).
- [21] Acedo, Francisco Jose, et al. "Co-authorship in management and organizational studies: An empirical and network analysis." *Journal of Management Studies* 43.5 (2006): 957-983.
- [22] Wu, Ying, Xing Jin, and Yunzhen Xue. "Evaluation of research topic evolution in psychiatry using co-word analysis." *Medicine* 96.25 (2017).
- [23] Cobo, Manuel J., et al. "Science mapping software tools: Review, analysis, and cooperative study among tools." *Journal of the American Society for information Science and Technology* 62.7 (2011): 1382-1402.
- [24] Labadie, J. W. (2004). Optimal operation of multireservoir systems: State-of-the-art review. *Journal of water resources planning and management*, 130(2), 93-111.
- [25] Holzworth, D. P., Huth, N. I., deVoil, P. G., Zurcher, E. J., Herrmann, N. I., McLean, G., ... & Keating, B. A. (2014). APSIM– evolution towards a new generation of agricultural systems simulation. *Environmental Modelling & Software*, 62, 327-350.
- [26] Ghodsypour, S. H., & O'Brien, C. (1998). A decision support system for supplier selection using an integrated analytic hierarchy process and linear programming. *International journal of production economics*, 56, 199-212.

- [27] Maier, H. R., Jain, A., Dandy, G. C., & Sudheer, K. P. (2010). Methods used for the development of neural networks for the prediction of water resource variables in river systems: Current status and future directions. *Environmental modelling & software*, 25(8), 891-909.
- [28] Mishra, A. K., & Singh, V. P. (2011). Drought modeling—A review. *Journal of Hydrology*, 403(1-2), 157-175.
- [29] Min, H., & Zhou, G. (2002). Supply chain modeling: past, present and future. *Computers & industrial engineering*, 43(1-2), 231-249.
- [30] Dalla Man, C., Rizza, R. A., & Cobelli, C. (2007). Meal simulation model of the glucose-insulin system. *IEEE Transactions on biomedical engineering*, 54(10), 1740-1749.
- [31] Kelly, R. A., Jakeman, A. J., Barreteau, O., Borsuk, M. E., ElSawah, S., Hamilton, S. H., ... & Voinov, A. A. (2013). Selecting among five common modelling approaches for integrated environmental assessment and management. *Environmental modelling & software*, 47, 159-181.
- [32] Ahmad, R., & Kamaruddin, S. (2012). An overview of time-based and condition-based maintenance in industrial application. *Computers & industrial engineering*, 63(1), 135-149.
- [33] Yates, D., Sieber, J., Purkey, D., & Huber-Lee, A. (2005). WEAP21—A demand-, priority-, and preference-driven water planning model: part 1: model characteristics. *Water International*, 30(4), 487-500.
- [34] Newman, Mark EJ. "Scientific collaboration networks. I. Network construction and fundamental results." *Physical review E* 64.1 (2001): 016131.
- [35] Newman, Mark EJ. "Coauthorship networks and patterns of scientific collaboration." *Proceedings of the national academy of sciences* 101.suppl 1 (2004): 5200-5205.
- [36] Durand, Rodolphe, Robert M. Grant, and Tammy L. Madsen. "The expanding domain of strategic management research and the quest for integration." *Strategic Management Journal* 38.1 (2017): 4-16.
- [37] Börner, Katy, Chaomei Chen, and Kevin W. Boyack. "Visualizing knowledge domains." *Annual review of information science and technology* 37.1 (2003): 179-255.
- [38] Callon, Michel, et al. "From translations to problematic networks: An introduction to co-word analysis." *Social science information* 22.2 (1983)
- [39] A. W. Harzing and S. Alakangas, "Google Scholar, Scopus and the Web of Science: a longitudinal and cross-disciplinary comparison," *Scientometrics*, vol. 106, no. 2, pp. 787–804, 2016.
- [40] M. E. Mcveigh, "Citation Indexes and the Web of Science," *Encycl. Libr. Inf. Sci. Third Ed.*, no. August 2013, pp. 37–41, 2009.
- [41] T. Reuters, "WEB OF SCIENCE, Citation Report, Thomson Reuters," Thomson Reuters, 2014.
- [42] Sprague, R. H., Jr. and H. J. Watson, "Bit by Bit: Toward Decision Support Systems", *California Management Review*, vol. XXII, no. 1, Fall 1979, pp. 60-68.
- [43] Gavish, B. (1981). A decision support system for managing the transportation needs of a large corporation. *AIIE Transactions*, 13(1)

Deep Learning Based Forecasting Of Cancellation, Delay And Orientation On Flights

Anil AYAYDIN¹, M. Ali AKCAYOL²

¹Computer Science, Informatics Institutes, Gazi University, Ankara, Turkey
²Computer Engineering, Faculty of engineering, Gazi University, Ankara, Turkey
anil.ayaydin1@gazi.edu.tr, akcayol@gazi.edu.tr

Abstract— In this study, three different methods from the machine learning and deep learning have been implemented for forecasting of cancellation, delay and diversion on flights, which constitute a major problem in the aviation industry. Deep recurrent neural network (DRNN), long-short term memory (LSTM), and random forest (RF) methods have been extensively tested and compared using a real data set covering 368 airports across the world with respect to the success rate of forecasting of cancellation, delay and diversion on flights. Experimental results show that the recall value has the highest success rate when LSTM is used as 96.50%.

Keywords— forecasting, deep learning, machine learning, aviation.

1. INTRODUCTION

According to the statistical data published by the United States Bureau of Transportation, 14.69% of the flights that have taken place so far in 2020 have not been on time due to delays or cancellations. Of these flights, 0.19% was diversion (the plane landed at a different airport than the planned airport) [1]. Delays in scheduled flights cause serious financial losses for airline companies, which also leads to a decrease in passenger satisfaction. Based on this, many countries in the world have made deterrent legal arrangements in order to compensate their citizens for the damages caused by these delays that are likely to be experienced, and they impose monetary penal sanctions on the relevant airline companies. However, although these sanctions are applied, there are still frequent delays and cancellations in flights. Delay codes for departures within the scope of commercial passenger flights have been standardized by International Air Transport Association (IATA) and the main causes of flight delays are unsuitable weather conditions, air traffic density, airport capacity (number of passengers, runways and bridges, etc.), technical and mechanical problems, national aviation systems and security [2].

In this study, three different methods from the deep learning and the machine learning have been implemented for forecasting of cancellation, delay and diversion on flights, which constitute a major problem in the aviation industry. In order to train the model, unlike similar studies which have not just a few airports, operations were carried out on the data set covering 368 airports.

Three different models were presented in the study and independent result data were obtained by using different input data on these models. It is aimed that the results obtained within the scope of the study will also be a source for other studies related to the aviation sector to be made with deep learning architectures [3].

2. LITERATURE REVIEW

There are many studies in the literature on the forecasting of aviation flights. Flight cancellation, delay and diversion are widely recognized as a critical performance indicator of flights in the commercial aviation industry. Approximately 87.50% of academic studies on estimating flight delays took place in the 2000s and were published between 2007 and 2017. Most of these studies have been carried out with the help of Machine Learning algorithms and Data Mining methods [3]. Researchers evaluate flight delays from different perspectives. These can be listed as optimization of airport planning, airport capacity increase, flight cancellation, facility location and flight change. In the study by Kang and Hansen [4], the effects of on-time and early arrival data on scheduled block time (SBT) adjustment decisions were investigated using data from 5 major United States-based airlines. In the study, changes in SBT were modeled for flights that took place two years in a row. Flight data from January 2008 to April 2014 were obtained from the Federal Aviation Administration (FAA) and Aviation System Performance Metrics (ASPM) database. In the methods used section, a mixed logit model is used to capture the heterogeneous preferences of each airway and the possible correlation between alternative

flights. On-time and early arrival features are generated for each alternative arrangement. Mixed logit models are estimated based on these features and the selected block time setting.

In the study of Kenan, Jebali and Diabat [5], the integrated flight planning, fleet allocation and the problem of the aircraft routing was formulated. Delays have been considered to make the model more realistic. The maintenance route was considered indirectly by ensuring that each aircraft's route ends at the same destination from which it started. In the main problem, the solution is reached with the help of a column-based formulation containing a large number of variables. The dataset includes 228 destinations and 45 different destinations. The flights were carried out with 59 aircraft with 5 different aircraft types. In this study, a two-stage stochastic programming model is developed for the problem of integrated flight planning, fleet allocation and aircraft routing under uncertainty. The importance of this model lies in the possibilities that can extend beyond it. This model has proven to be of high complexity and therefore a commercial solver such as CPLEX cannot resolve samples of the complexity. In this study, three column generation-based approaches were developed, and all three proved superior to CPLEX by resolving large samples in less than 4 hours to an optimality gap of less than 1%. Then, the effect of some of these obtained parameters on the airline company was also examined by performing a sensitivity analysis.

Yazdi, Dutta, and Steven's study [6] examined the links between the application of baggage charges and delayed flights in the airline industry. As the dataset, a panel dataset of 46 quarters was collected from the Timely Performance Database from the US Department of Transportation's Bureau of Transportation Statistics. The data set continues from the third quarter of 2003 to the fourth quarter of 2014. The On-Time Performance database provides information about non-stop domestic flights of major air carriers, including actual departure time, estimated departure time, actual arrival time and estimated arrival time. It also reports minute flight-level delays divided into five categories of causes: Carrier Delay, Air Delay, Security delay, Late Aircraft Delay, and National Air System Delay. In the study, it has been investigated how Baggage Fees (BF) affect the delays with the help of formulations created by arranging them according to the flight time. Eleven carriers, ten of which apply baggage fees, were studied. The results show that, on average, RF implementation results in improved on-time performance as judged by direct flights and indirectly through ticket prices and market demand. The results also show that developments are affected by the hub-airports on the route and the classification of passengers as leisure or business. The analysis shows that the first phase in implementation will lead to more flight delays, but full implementation really improves and still improves late flights.

Kim et al. [7] attempted to predict departure and arrival delays using flight order and weather data from the

proposed Recurrent Neural Networks, National Oceanic and Atmospheric Administration. The accuracy of the delayed flight forecast was measured at 91.81% for McCarran International Airport and 71.34% for Sky Harbor International Airport due to the difference in data volume.

Choi et al. [8] focused on the relationship between flight delays and weather. Weather data was collected from the National Oceanic and Atmospheric Administration. As a result, Random Forest algorithm, which is a proposed ensemble learning method, predicted the arrival delay with 80.36% accuracy.

Belcastro et al. [8] estimated flight delays due to bad weather conditions using Random Forest via MapReduce. The following results were obtained using weather data from the National Oceanic and Atmospheric Administration. At a 15-minute delay threshold, 74.2% accuracy and 71.8% recall, at a 60-minute threshold, 85.8% accuracy and 86.9% recall values were obtained.

Thiagarajan et al. [9] made a flight delay estimation using the weather data at the locations of the departure and arrival airports, obtained via the World Weather Online API service, with the help of Gradient Boosting, a machine learning technique. They achieved 94.35% accuracy in arrival delays and 86.48% accuracy in departure delays.

Prasad et al. [10] achieved a 78% success rate in classification with Decision Tree and 77% in classification with Regression within the scope of flight delay estimation using the same data source as Thiagarajan.

Yu et al. [11], using the departure and arrival flight delay data of PEK airport between January 2017 and March 2018, the accuracy rate was 93% with DBN-SVR estimation method, 87% with k-NN, 87% with Support Vector Machine and finally It was calculated as 82% with Logistic Regression. Manna et al. [12] adopted the statistical approach method and using the Gradient Boosted Decision Tree classifier, on the flight delay dataset of the 70 busiest airports belonging to the US Department of Transportation in the April-October 2013 period, approximately 92% accuracy in arrival delays and approximately 94% accuracy in departure delays. obtained rate.

In this study, a supervised learning model has been developed by using deep learning architectures to predict cancellation, delay or diversion situations in flights, which are an important problem in the aviation industry. The data of the flights that took place within the scope of 368 airports were used. At this point, a more comprehensive dataset than other studies in the literature has been studied. Thus, the high diversity of data on the basis of airports and data lines that can repeat each other have been prevented.

The dataset used has been normalized by digitization method by removing the erroneous data. DRNN, LSTM

and RF methods were applied to the model, respectively. At the stage of teaching the model to the system, unlike similar studies in the literature, operations were carried out on the data set covering 368 airports, not just a few airports.

The other literature studies were listed in below.

Table 1. The other literature studies

Source	Target	Method	Parameters
[13]	Taxi departure time estimation	Queuing model	Airline, terminal, air, destination, queue size
[14]	Delay prediction	Probability model, reinforcement leaning	Seasonal trend, delay data
[15]	Taxi departure time estimation	Reinforcement learning	Wheel departure time, wheel landing time, seasonal average number of taxi entries and exits
[16]	Delay classification	Probability model	Type of flight, number of passengers, number of delayed flights, visibility, wind speed
[17]	Delay prediction	Queuing model	Planned flight time, delay time
[18]	Delay prediction	Adaptive network	Destination, arrival time, arrival delay, scheduled arrival time
[19]	Delay prediction	Random Forest	Hour, day, month, delay status, delay day type, previous day type, airport delays, departure-destination places

3. DEEP LEARNING

Artificial intelligence encompasses the method of machine learning, in which machines can learn by experience and acquire skills without human. Deep learning, on the other hand, is defined as a subset of machine learning where artificial neural networks (ANNs) and algorithms inspired by human intelligence learn from data. Similar to learning process of people, the deep learning algorithm have a better result with a little improvement for each time to improve the outcome. Any problem that requires thought can be realize by deep learning. It has been determined that 2.5 quintillion bytes of data are produced every day [20]. Because of requiring more data to improve the learn, data generation has led to an increase the capabilities of the deep learning models in recent years. In addition, deep learning algorithms take advantage of the more powerful data processing capacity available today, along with the development of ANN technology [21].

Artificial intelligence (AI) algorithms need for deep learning without a large initial investment. Even when using a diverse, unstructured and interconnected dataset, deep learning enables machines to solve successfully complex problems. The deeper learning algorithms are trained, the better they perform [22].

Convolutional neural network (CNN), which is a forward-looking neural network from machine learning techniques, is a type of multilayer perceptron. It was first introduced in the early 1980s by LeCun [23] with the LeNet-5 architecture. This architecture consists of an input layer, several convolutional layers, pooling and output layers. According to the convolutional layers level, they play a role in the extraction of features by performing operations on the inputs received from the previous layers. While the first convolutional layer provides the lowest feature, it provides higher-level feature extraction as the convolutional layers are added, that is, the higher the level. The pooling layer is; it is used to simplify the outputs produced with convolutional layers. The output layer of the architecture; it can be connected to fully connected layers or to layers, such as the sigmoid layer. Data in the input layer; Since it can be multimedia data, such as sound, image, video, the researchers preferred it in many signal processing fields in recent years because of its high performance [24]. A strategy to reduce the margin of error is implemented by using a back-propagation algorithm that adjusts the learning weights to be updated with margins of error throughout the training process of CNN architectures [25].

Autoencoders are also one of the popular models in the field of deep learning. As the name suggests, it aims to automatically learn to convert any data into code. It consists of two parts: Encoder and Decoder. These two parts are trained together as if they were a single model during the training phase. After the training is over, these models are separated and used to compress the data and decompress the compressed data. For example, if the data is to be moved from one place to another, less data can be transported by placing an Encoder on the sender and a Decoder on the receiver.

Since recursive neural networks (RNNs) models the behavior of dynamically changing systems through their hidden layers and are defined as a type of ANN in which the output from the previous step is fed as input to the next step, it stands out as the method that works best with the data set used in the study and has been preferred in practice for this reason. In the next sub-heading, recursive neural networks, which is a deep learning architecture and determined as a method in this project, will be mentioned.

4. ANALYSIS OF AVIATION FLIGHT DATA AND FUTURE PREDICTION

4.1. Dataset

In the dataset, there are domestic flights within the borders of our country between the years 2010-2020 and international origin flights between the flight points. Within the scope of the data set, data with the number of 62597 rows were obtained after the data cleaning phase, which was carried out with the help of a console application developed separately with C#. Based on the data set obtained in this study, flights were classified according to cancellation or delay/forwarding status.

Table 2. Features of the flight dataset

Feature	Value
Class	Anonym
Category	Aviation and flight data
Subcategory	Flight Dataset
Data Owner	General Directorate of State Airports Authority
Description	Aviation Flight Dataset which has 62597 row data
Format	CSV
Keywords	Flight, Airport, Cancelled, Delayed, Diversion
Contact	https://www.dhmi.gov.tr
Row Count	62597
Update Time	May 2020
Size	4.6 MB

4.2. Forecasting of Aviation Flight Data with Deep Learning

4.2.1. Deep Recurrent Neural Network Method

In the scope of the study, DRNN has been used as a supervised learning model. The architecture used works iteratively as mentioned in the previous titles, and the output data of the previous step affects the output of the next step. Since the data set used in this study is compatible with the model in question, DRNN architecture has been the preferred method in analysis. Since single-layer artificial neural network (ANN) models will be insufficient for complex models, multi-layer feedback ANN, that is, the architecture where the outputs of the layers are fed back to the previous layers, is preferred. In the model built as a result of this, there are multiple hidden layers that form the basis of deep learning architectures in addition to the input and output units. There is no analytical method for determining the number of layers and neurons in the hidden layers, and the methods suggested by Karsoliya were used as the upper limit [26].

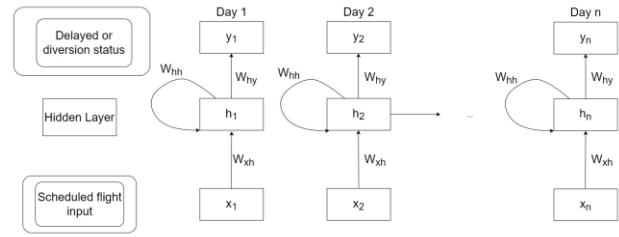


Figure 1. The DRNN model

In the diagram above W_{hh} corresponds to the weight matrix for the transition between the previous step and the hidden layers on the current step, W_{xh} , the weight matrix between the input and hidden layers, W_{hy} , the weight matrix between the hidden layer and the output. There are a number of hidden layers between the x_t planned flight input and y_t output layers. Since the result of the previous step is iterative to affect the next one, it is a suitable model for the aviation flight data set used in this study.

4.2.2. Long Short-Term Memory Model

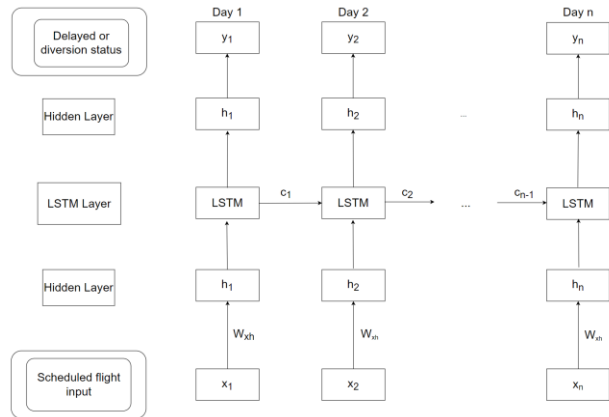


Figure 2. LSTM model

Here W_{hh} , W_{xh} and W_{hy} in the LSTM model. As n are several hidden layers and LSTM layers between the planned flight input and output layers. c_t refers to the memory value of the current step and c_{t-1} corresponds to the memory value of the previous step. Since the memory data of the previous step is kept in the LSTM layer, this model developed to be implemented on the aviation flight data set increases the performance.

4.2.3. Random Forest Model

As described in Figure 3, in the flight prediction model with RF, the best prediction result is obtained by comparing the planned flight inputs from the flight training data set with the previous one, taking the better

value by voting and similarly progressing to the value through this comparison with the value of n .

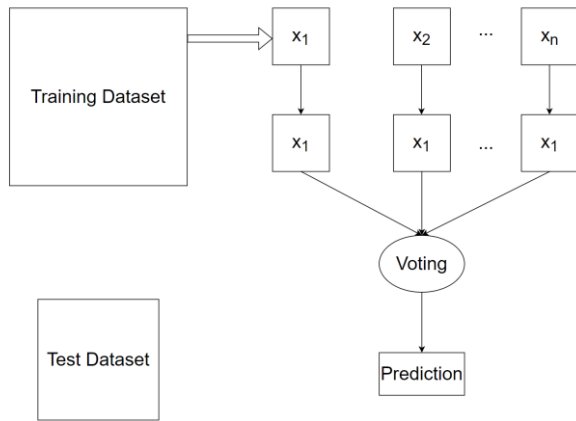


Figure 3. RF model

Although there are different approaches in the literature as a training method, the DRNN model was compiled using Jupyter Notebook to determine the most appropriate and effective training function, and it was observed that the LSTM has better results compared to other tests, and this model was also preferred in experimental studies.

In the model built with the RF algorithm, the date, time, departure and arrival points were determined and estimation was made with the help of a method developed. In the DRNN model, in addition to making the estimation process using the same parameters, the system was trained at the determined iteration value by creating a multiple hidden layer artificial neural network model and the correct classification values were measured. In the LSTM model, unlike the DRNN model, the results of the previous step are stored with the help of a memory cell and the results are obtained. In the light of these data, the DRNN architecture has stepped forward numerically. However, no distinctive difference could be detected in terms of accuracy rate between DRNN and LSTM recursive deep learning architectures.

Although the RF machine learning algorithm gives better results in terms of correct classification, deep learning models are the focus of the study. In this context, the LSTM model has a higher accuracy value compared to the DRNN model and will set an example in terms of being used in future studies concerning the aviation industry.

5. THE EXPERIMENTAL RESULTS

Approximately 75% of the data set used consists of delayed/redirected flights, while 25% consists of canceled flights. Accordingly, the success rate in classification was approximately 75% when the DRNN architecture, which was modeled as cancellation or delay/redirect by applying cross validation with 80% training and 20% test data, was trained. Dropout values have been added to each hidden

layer to reduce the memorization that can occur with each iterative step.

Since it can store the information of the hidden layers in the memory in the long term, a success rate of approximately 79% was achieved in classification by applying the LSTM, which outperforms the DRNN.

Before the application of deep learning architectures, which is the main subject of the study, with RF, a machine learning algorithm in which classification processes are carried out by teaching the features to the architecture, the classification was made and the result data were obtained. According to this, RF was more successful than the LSTM method with an accurate classification rate of about 82%, although it was not a big difference.

Among the models applied based on deep learning, the LSTM is the architecture with the highest numerical value in terms of correct classification and performance percentage. As a result of the experiments on all models, the results in Table 2 were obtained.

Table 3. Values calculated by means of different models according to the parameters divided into terms to fit in the column and in accordance with its notation.

Scores	RF	DRNN	LSTM
Accuracy Value (%)	82.21	78.41	76.96
Recall (%)	96.21	94.07	96.50
Precision (%)	96.18	80.36	77.96
F-score (%)	96.20	86.68	86.30
ROC-AUC	90.11	78.35	71.68

6. CONCLUSIONS

In this study, it is aimed to ensure the feasibility of operations such as classification and estimation by creating a system in which deep learning architectures are applied by making use of the flight data obtained, within the scope of reducing the negative effects of cancellation, delay and diversion events in the aviation ecosystem, where our country has become one of the few centers in the world in recent years.

DRNN, LSTM and RF have been applied for the forecasting of cancellation, delay and diversion on flights. A real data set covering 368 airports across the world have been used for training and test of the models. The experimental results show that the recall value has the highest success rate when LSTM is used as 96.50%.

ACKNOWLEDGEMENT

We would like to express our gratitude and respect to the General Directorate of State Airports Authority and its esteemed managers, who are the authority of the airport management sector in Turkey, while creating the flight data set within the scope of this study for allowing the use

of the public flight data and providing financial support for participation to the conference.

REFERENCES

- [1] Internet: Bureau of Transportation Official, Bureau of Transportation Statistics, <https://www.bts.gov/>, 15.11.2021
- [2] L. Belcastro, F. Marozzo, D. Talia, and P. Trunfio, "Using Scalable Data Mining for Predicting Flight Delays", *ACM Trans. Intel. syst. Technol.*, 1, 5-20, 2016.
- [3] A. Sternberg, J. D. A Soares, D. Carvalho, E. Ogasawara, "A Review on Flight Delay Prediction.", *arXiv: Computers and Society*, 2017.
- [4] L. Kang, M. Hansen, "Behavioral analysis of airline scheduled block time adjustment", *Transp. pic. Part E: Log. Transp. Rev.* 103, 56–68, 2017.
- [5] N. Kenan, A. Jebali, A. Diabat, "The integrated aircraft routing problem with optional flights and delay considerations", *Transp. pic. Part E: Log. Transp. Rev.*, 118, 355–375, 2018.
- [6] A. A. Yazdi, P. Dutta, A. B. Steven. "Airline baggage fees and flight delays: a floor wax and dessert topping?", *Transp. pic. Part E: Log. Transp. Rev.*, 104, 83–96, 2017.
- [7] YJ. Kim, S. Choi, S. Briceno, D. Mavris, "A deep learning approach to flight delay prediction.", **2016 IEEE/AIAA 35th Digital Avionics Systems Conference (DASC)**, Sacramento, CA, USA, 1-6, 2016.
- [8] L. Belcastro, F. Marozzo, D. Talia, P. Trunfio, "Using scalable data mining for predicting flight delays.", *ACM Trans Intel SystTechnol.*, 8, 1–20, 2016.
- [9] B. Thiagarajan, L. Srinivasan, A. V. Sharma, D. Sreekanthan, V. Vijayaraghavan, "A machine learning approach for prediction of on time performance of flights.", **In: Proceeding of 2016 IEEE/AIAA 35th digital avionics systems conference (DASC)**, Sacramento, CA, USA, 6, 2016.
- [10] U. S. Prasad, P. A. Chauhan, L. S. Asha., "Data mining & predictive analysis on airlines performance.", **Int J Pure Appl Math.2018;118:1–12. AIAA 35th digital avionics systems conference (DASC)**, Sacramento, CA, USA, 6, 1–6, 2016.
- [11] B. Yu, Z. Guo, S. Asian, H. Wang, G. Chen, "Flight delay prediction for commercial air transport: A deep learning approach", *Transportation Research Part E*, 125, 203 – 221, 2019.
- [12] S. Manna, S. Biswas, R. Kundu, S. Rakshit, P. Gupta, S. Barman, "A Statistical Approach to Predict Flight Delay Using Gradient Boosted Decision Tree", **International Conference on Computational Intelligence in Data Science**, Chennai, India, 1-5, 2017.
- [13] H. Idris, J. Clarke, R. Bhuva, L. Kang, "Queuing Model for Taxi-Out Time Estimation", *ATC Quarterly*, 10(1), 1-22, 2002.
- [14] Y. Tu, M. Ball, W. Jank, "Estimating Flight Departure Delay Distributions – A Statistical Approach with Long-Term Trend and Short-Term Pattern", *Journal of the American Statistical Association*, 103, 112-125, 2008
- [15] P. Balakrishna, R. Ganesan, L. Sherry, "Accuracy of Reinforcement Learning Algorithms for Predicting Aircraft Taxi-Out Times: A Case-Study of Tampa Bay Departures", *Transportation Research Part C: Emerging Technologies*, 18, 950-962, 2012.
- [16] R. Wesonga, F. Nabugoomu, P. Jehopio, "Parameterized Framework for the Analysis of Probabilities of Aircraft Delay at an Airport", *Journal of Air Transport Management*, 23, 1-4, 2012.
- [17] M. Arikan, V. Deshpande, M. Sohoni, "Building Reliable Air-Travel Infrastructure Using Stochastic Models of Airline Networks", *Operations Research*, 61 (1), 45-64, 2013.
- [18] S. Khanmohammadi, C. Chou, H. Lewis, D. Elias, (2014), "A Systems Approach for Scheduling Aircraft Landings in JFK Airport", **IEEE International Conference on Fuzzy Systems (IEEE 2014)**, Beijing, China, 11, 1578-1585, 2014.
- [19] J. Rebollo, H. Balakrishnan, "Characterization and Prediction of Air Traffic Delays,", *Transportation Research Part C: Emerging Technologies*, 44, 231-241, 2014.
- [20] Internet: B. Marr, What Is Deep Learning AI? A Simple Guide With 8 Practical Examples, <https://www.forbes.com/sites/bernardmarr/2018/10/01/what-is-deep-learning-ai-a-simple-guide-with-8-practical-examples/#3be4ab4c8d4b>, 15.11.2021.
- [21] K. Funahashi, Y. Nakamura, "Approximation of dynamical systems by continuous time recurrent neural networks", *Neural networks*, 6, 801–806, 1993.
- [22] Internet: A. Amidi, Recurrent Neural Networks Cheatsheet, <https://stanford.edu/~shervine/teaching/cs-230/cheatsheet-recurrent-neural-networks>, 15.11.2021.
- [23] Y. Lecun, B.E. Boser, J.S. Denker, D. Henderson, R.E. Howard, W.E. Hubbard, LD. Jackel, "Handwritten digit recognition with a backpropagation network", *Advances in Neural Information Processing Systems*, 2, 396 –404, 1990
- [24] M. Valenti, S. Squartini, A. Diment, G. Parascandolo, T. Virtanen, "A convolutional neural network approach for acoustic scene classification", **International Joint Conference on Neural Networks (IJCNN)**, Anchorage, Alaska, 2017.
- [25] H. Eghbal-Zadeh, B. Lehner, M. Dorfer, G. Widmer, "A hybrid approach using binaural I-vectors and deep convolutional neural networks", **Workshop on Detection and Classification of Acoustic Scenes and Events (DCASE)**, Budapest, Hungary, 2016.
- [26] S. Karsoliya, "Approximating Number of Hidden layer neurons in Multiple Hidden Layer BPNN Architecture", *International Journal of Engineering Trends and Technology*, 6, 714–717, 2012.

Comparison of AutoML Libraries in order to Use in Anomaly Detection Processes for Conveyor Line Used in Autonomous Production System*

Cihan BAYRAKTAR^{1*}, Hadi GÖKÇEN²

¹ Eskipazar Vocational School, Karabuk University, Karabuk, Turkey

² Engineering Faculty, Gazi University, Ankara, Turkey
cihanbayraktar@karabuk.edu.tr, hgokcen@gazi.edu.tr

Abstract-- The digitalization processes in production, which are realized with the increasing use of information technologies being more integrated into production systems, provided the emergence of the Industry 4.0 revolution. One of the important tasks that emerged for smart factories, whose name is frequently mentioned with the Industry 4.0 revolution, is anomaly detection processes. In this study, a simulation of the transport unit was developed for the popcorn production systems and the anomaly detection processes over the streaming data were analyzed using AutoML libraries with the data obtained. As a result of the tests, the learning model obtained by the Auto-Sklearn library and created with the hyper-parameters of the Random Forest algorithm was the model that gave the best result with an accuracy rate of 91.3%.

Keywords—industry 4.0, smart factories, anomaly detection, machine learning

1. INTRODUCTION

Industry 4.0 has been invented as an initiative by Germany Federal Government together with universities and private companies in 2011. The aim of this program, which emerged as a strategic program, is to increase the productivity, activity and efficiency of the industry and to create advanced production systems. This structure contains bringing together a number of technologies known to contribute to the product life cycle under one roof and integrating them into a common structure. Industry 4.0 allows the formation of flexible lines in an advanced production or intelligent production structure. In this way, it is used as a system in which production processes are adjusted automatically in line with changing conditions [1].

Industry 4.0 has had a great importance in the global market, as faster and higher efficiency production processes are realized with its technologies. As a result of the processes that took place from the very beginning of the industrial revolutions, technical developments have emerged to a large extent. In general, Industry 4.0 is expressed as a combination of applications working under one roof. Some of them are listed as big data, autonomous robotics, industrial internet of things, simulation, artificial intelligence, cyber security, cyber physical systems, cloud systems and smart factories [2].

Smart Factories concept that is one of the building block of Industry 4.0 is defined from different angles by academicians and applicators. However, there is still no consistent and shared definition on this subject. According to many of the definitions, smart factories are a concept that is located at the center of cyber-physical systems, which includes all communication systems, information infrastructure and control processes in order to create smart production systems. Smart factory systems aim to create cyber-physical systems based on higher-level production to obtain physical assets with the production system and self-adaptive production processes, which are formed by the extremely flexible behaviour of autonomous and smart machines and products [3].

In these times when the internet of things, artificial intelligence and cloud computing systems are fully integrated, one of the most important tasks for smart factories is to detect anomalies that may occur instantly in production lines. In the past, this task was tried to be done by the inspectors who constantly observed the system on the production line. But now it has begun to take its place as automatic systems that are checking the data read by sensors instantly and detecting whether there are anomalies in the smart factories. In this way, it is planned to prevent serious damage to the devices used in the production lines and to prevent possible economic losses due to errors. In the future, making the anomaly detection processes

completely autonomous and the structures that enable the necessary systems to be triggered independently of people are among the objectives to be achieved in the process [4].

Within the scope of the study, a simulation was created for a conveyor system belonging to the transport unit of a popcorn production system and was developed an application that is detecting anomalies by analyzing stream data what are created during work the conveyor. In the second part of the study, some previous studies on the subject are included. In the third part, the application is explained and the results are analyzed.

2. RELATED WORKS

Hranisavljevic et al. In their study in 2016, they have proposed an anomaly detection algorithm with automatic learning feature that can be used in hybrid production systems. They have benefited from the combination of deep learning techniques and timed automata systems to create the detection model from the observations applied in the system. They have explained that the algorithm they tested on various datasets, including two real systems, gave promising results [5].

In a study conducted by Pahl and Aubet in 2018, a resource-efficient approach that learns the communication and working structure of in-service services in the IoT system and keeps itself constantly updated have been proposed. They have stated that this proposed approach can detect anomaly in line with the model learned by analyzing the data flowing in the process communication between nodes. Thanks to the study, it has been concluded that the security levels of IoT systems can reach higher levels [6].

In a study on deep learning structures for anomaly detection, researchers have aimed to be better understood the anomaly detection systems created based on deep learning. In the study, anomaly detection techniques applied before deep learning structures have been explained in the first place. Afterwards, the techniques used for the high-tech deep learning-based anomaly detection systems applied today to overcome the problems of the traditional algorithms used before have been discussed [7].

In the study by H. Wu et al., an in-network solution proposal suitable for microservices has been developed to decompose the acoustic data used in anomaly detection. As a result of the study, it was determined that the time spent for data separation could be reduced by 43.75% and it was stated that the application defines how it can be used better for critical tasks [4].

In another study on anomaly detection, an anomaly detection system was developed by taking the sounds produced by machines working in a production system as acoustic signals and converting them into spectrogram images and using the convolutional neural network algorithm on these images. It has been stated that the model tested in a real factory environment is 99.44% accurate,

thus helping to reduce injuries by speeding up the detection of machine malfunctions [8].

3. IMPLEMENTATION

The system is built on simulating the transportation of the required raw corn product to the interim storage unit in 5000 gr batches in line with the work order. On the Ubuntu operating system, containers were created with the docker virtualization system and the simulations of the devices were written with the python programming language and transferred to the containers. In this way, it is aimed to produce data by working as if each container is a part of the production system. The created simulation have been planned to consist of Enterprise Resource Planning (ERP) device, transport unit and storage unit as seen in Figure 1 of the work flow chart.

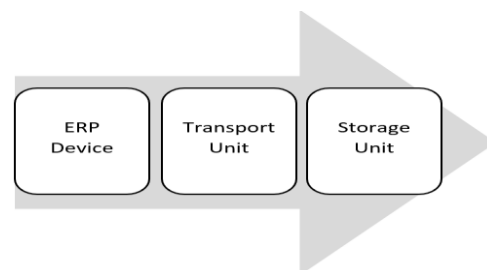


Figure 3. Flow diagram of the simulation

For the creation and operation of the system, a computer with an Intel Core i7 9750H model 2.6 Ghz 6 core CPU, 6 Gb Ram capacity and Ubuntu 20.0 operating system have been used. The simulation created consists of the ERP device used in the creation of the work order to start the production processes, and the Transport unit that starts working after the work order and transfers the raw corn products.

After the system was installed, the data were recorded repeatedly for a certain period of time, and tests were carried out with Auto-Keras, Auto-Sklearn and PyCaret AutoML libraries on these data, and the most suitable algorithm and hyperparameters were tried to determine.

3.1. ERP Device

The ERP unit was established for the management of the production processes of the system and the creation of work orders. The system has 3 different types of production capacity. These varieties are determined as small size package (100g), medium size package (250g) and large size package (500g). When the ERP application is run, the amount of raw corn available and how many packages of popcorn are available in stock are displayed on the main screen. In addition, the system carries out production with 3 different speed options. The desired speed option can be selected according to the time requirement of the production process to be applied, and accordingly the realization and completion of the process is ensured.

The ERP module has two different usage options. In the first option, it allows entering the sales information in the package type selected directly from the stock and deducting the number of products sold from the stock. In this structure, if the product quantity drops below 100 at the end of the sale, the production system is automatically activated and 250 units of the relevant package are produced. In the second application, by entering information by the user, it is ensured that the desired product is produced for the desired amount of stock.

With ERP Module, If the production order is entered or after the customer order is entered, if the stock amount falls below the determined limit, the system starts the production process with the movement of the Transport unit.

3.2. Transport Unit

The transport unit consists of a silo filled with raw corn, an aspirator for regulating the air flow during transport, and a conveyor system that will take the filled silo to the storage unit and return the empty silo.

In line with the information coming from the ERP, the Transport unit will start to transport the total amount of raw corn required to the Storage unit by obtaining the package size and quantity information to be produced. The following equation (1) is used to calculate the required total amount of raw corn.

$$Raw_{Product} = \frac{(Processed_{Product} * 100)}{92} + (Package_{Number} * Factor) \quad (1)$$

In the above equation;

- **Raw_{Product}**: Represents the required raw corn weight,
- **Processed_{Product}**: Represents the total weight of processed corn that should be obtained at the end of production,
- **Package_{Number}**: Represents the total number of packages that need to be produced,
- **Factor**: Represents the value (Factor=1 for Small Size; Factor=2 for Medium Size; Factor=3 for Large Size) which will allow the amount of spare raw corn to be added to the total weight, depending on the package size and number of packages to be produced.

After the amount of raw corn required for production is calculated, the Transport unit will start to transport the raw corn to the Storage unit, with an average of 5000g (4850g - 5150g) each time. The transport unit carries out each trip in 400 steps and in 30, 35 or 40 seconds depending on the preferred speed option. The average time between steps is 0.079 sec to 0.108 sec. During working time at each step, Unix timestamp, silo occupancy information, information showing that the aspirator and conveyor are working and stopped, the distance taken by the conveyor during transportation, the power and voltage information consumed, the amount of raw corn transported instantly,

the amount of raw corn transported to that moment and the remaining corn to be transported produces information about the amount. This information is also recorded in the database and what is needed is transmitted to other relevant machines. Table 1 shows the definitions of the data set created in order to use the anomaly detection system specific to the transport unit.

Table 6. Attribute definitions of transport unit data set

Attribute	Definitions
TimeID	System Operating Speed Information (Numeric)
SiloFull	Silo Full Occupancy Information (0 – 1)
SiloMinFull	Silo Minimum Occupancy Information (0 – 1)
Asprator	Working Status Information of the Aspirator (0 – 1)
Conveyor	Working Status Information of the Conveyor (0 – 1)
Distance	Conveyor Distance Information (Numeric)
Power	Conveyor Instant Power Consumption Information (Numeric)
Voltage	Conveyor Instant Used Voltage Information (Numeric)
Active	Amount of Instantly Transported Raw Corn (Numeric)
Class	Anomaly Status Information (0 – 1 – 2 – 3 – 4)

Also, an example piece from the Transport Unit dataset is presented in Figure 2 to illustrate the data structure.

```
1.17,1,1,0,1,3749,425,223,5003,0
1.17,1,1,0,1,3719,423,229,5003,4
1.17,1,1,0,1,3682,453,228,5003,2
1.17,1,1,0,1,3652,530,221,5003,4
1.17,1,1,0,1,3622,504,247,5003,3
1.17,1,1,0,1,3592,432,225,5003,0
1.17,1,1,0,1,3562,456,242,5003,3
1.17,1,1,0,1,3532,379,220,5003,4
```

Figure 4. An example piece from the transport unit dataset

Distance information from the variables shown in Table 1 is the value obtained by moving the conveyor 30 units in each step. During the transport of the full silo to the storage unit, the distance value is calculated by increasing; During the return of the empty silo, the calculation process proceeds in decreasing order. The voltage value, the instantaneous measured value between 220V and 230V coming from the network is shown. Conveyor Power Equations shown in equation (2) and (3) were used to calculate the power consumed by the conveyor.

$$P = \frac{A * m * v}{1000 * \eta} \quad (2)$$

$$A = \left(\frac{2}{D} * \left(\mu * \left(\frac{d}{2} \right) + f \right) + \sin \alpha \right) * g \quad (3)$$

The meanings of the variables specified in the above equations used for the conveyor power calculation process are as follows [9]:

- P: Power consumed by the conveyor (W)
- m: Total weight on belt (belt + load) (kg)
- v: Belt Speed (m/sn)

- η : Efficiency
- D : Support drums outer diameter (mm)
- d : Support rollers hub diameter (mm)
- f : Rolling resistance
- α : Horizontal angle of conveyor
- g : Gravity acceleration (m/sn^2)
- μ : Frictional resistance coefficient

3.3. Theoretical Concepts

Within the scope of the study, 3 different AutoML libraries were used and comparison processes were made. Explanations about the AutoML libraries used are given in the subtitles.

3.3.1. Automatic Machine Learning (AutoML)

Preparing machine learning algorithms for use and making the necessary adjustments is a process that requires expertise and takes time. The reason for the development of AutoML systems is to simplify these processes and enable learning models to be created more quickly and easily. AutoML is preferred for people who are not competent in data analysis processes to use machine learning structures efficiently, and for experts to perform faster data analysis, especially by automating the selection of hyper-parameters [10]. One of the reasons for creating AutoML structures is to enable experts to automatically create learning models without the need for statistical and machine learning knowledge and experience [11].

3.3.2. Auto-Keras

Auto-Keras, like other AutoML libraries, is an open source library developed to enable people who do not have sufficient knowledge of machine learning technologies to easily perform data analysis. However, Auto-Keras focuses mostly on deep learning models compared to other libraries [12]. The purposes of use of the Auto-Keras library are stated as running the automatic search functions in order to determine the most suitable architecture to obtain the learning model and automatically selecting the necessary hyperparameters for the deep learning algorithms to be used in the analysis process. Auto-Keras, which was developed to simplify machine learning processes, makes use of neural architecture search (NAS) algorithms for this [13].

3.3.3. Auto-Sklearn

Auto-Sklearn uses meta-learning and Sequential Model-Based Algorithm Configurator (SMAC) methods to create the most appropriate learning model for a training data set. After proposing a reasonable learning model on the data set uploaded by the user, it initiates a SMAC search with the meta-learning structure so that the model can be evaluated. In this way, the best performing model is determined by testing the suggested models with iterative SMAC searches, and the best model is returned to the user using the Bayesian Approach [14].

Auto-Sklearn is a library built on Scikit Learn. Like all other AutoML libraries, it aims to make machine learning processes faster and easier by automatically selecting the algorithm that creates the best learning model and the necessary hyper-parameters. Auto-Sklearn has pipeline editing and uses the Bayesian approach to optimize it. In this way, the necessary parameter balancing can be done through the Bayesian approach in the processes of making hyper parameter organizations [13]. Auto-Sklearn makes comparisons using 15 machine learning algorithms, 14 preprocessing methods and 4 data preprocessing techniques to create learning models [15].

3.3.4. PyCaret

PyCaret, like all other AutoML libraries, is an open source AutoML library that has been developed for the purpose of successfully performing machine learning processes in a short time, without experts, and uses less code than others [16]. PyCaret produces models over 18 different machine learning classification algorithms in order to create learning models and presents their comparative tables to the user. In general terms, it can emphasize the preferred models in accordance with various criteria determined by the user. In addition, it shows the user the most important features obtained in the tables it produces [17].

3.4. Analysis Informations for Transport Unit

By operating the transport unit, both normal process data and anomaly data were created. In the studies so far in anomaly assessments, scenarios regarding the amount of faulty raw corn transported for the transport unit and the instant movement distance of the conveyor used in the unit, the mains voltage value and unbalanced power consumption have been used. Anomaly data about the amount of raw corn transported instantaneously are produced using the Gaussian function by increasing the standard deviation value obtained from the error-free data set by 40%. Anomaly data related to conveyor movement distance are data obtained by instantaneous speed error. Voltage anomaly data are erroneous data that may occur in the voltage data coming from the network. It is calculated randomly by increasing the standard deviation by 40%, by using the Gaussian function from the mean voltage data. Anomaly data, which occurs as a result of unbalanced power usage of the conveyor, is calculated by the unbalanced consumption of power as a result of a random error up or down in the current drawn by the conveyor. Each anomaly value in the data set is coded separately and necessary arrangements have been made to predict the type of anomaly that occurs with learning models. The Transport Unit is a responsible unit for carrying the raw corn to the storage unit with 5 kg silos upon the issuance of the work order. The data set used in the process of creating the necessary learning model for the anomaly detection processes in the Transport Unit consists of 39600 lines. The types of anomalies that may occur during the operation of the Transport Unit and their distribution in the data set are shown in Table 2.

Table 7. Anomaly informations of transport unit

Anomaly Code	Attribute Definition	Distribution of the Anomaly Data
0	Normal Data	10400
1	Instant Transport Amount Anomaly	9620
2	Conveyor Distance Information Anomaly	7632
3	Conveyor Voltage Information Anomaly	7254
4	Conveyor Used Power Information Anomaly	4694

The detailed results obtained as a result of the tests performed on 3 AutoML libraries in the created data set are shown in Table 3.

Since the class values in the data set used for the tests did not show an equal distribution, Precision, Recall and F-Score values have been also measured in order to interpret the test results properly. Thus, it has been aimed to obtain healthier results at the end of the tests.

Table 8. Transport unit data set detailed test results

		Normal (0)	Transported Product Anomaly (1)	Conveyor Distance Anomaly (2)	Conveyor Voltage Anomaly (3)	Conveyor Power Anomaly (4)	Model Accuracy
Auto-Keras	Precision	0,734	0,727	0,888	0,945	0,713	0,794
	Recall	0,783	0,818	0,731	0,961	0,621	
	F-Score	0,758	0,770	0,802	0,953	0,664	
Auto-Sklearn	Precision	0,891	1,000	0,899	0,953	0,729	0,913
	Recall	0,989	1,000	0,744	0,964	0,757	
	F-Score	0,938	1,000	0,814	0,958	0,743	
PyCaret	Precision	0,891	1,000	0,912	0,944	0,734	0,912
	Recall	0,989	1,000	0,725	0,967	0,782	
	F-Score	0,938	1,000	0,808	0,955	0,757	

According to the results shown in Table 3, the two best results have been obtained by the Auto-Sklearn and PyCaret AutoML libraries. The algorithm that gave the best results within the Auto-Sklearn library has been the Random Forest (RF) algorithm with the hyperparameters shown in Figure 3. In the PyCaret library, the algorithm that gave the best results has been the Light Gradient Boosting Machine (LGBM) algorithm with the hyperparameters shown in Figure 4.

```
RandomForestClassifier(bootstrap=False,
criterion="gini",
max_depth=None,
max_features=0.23832696118792362,
max_leaf_nodes=None,
min_impurity_decrease=0.0,
min_sample_leaf=1,
min_sample_split=16,
min_weight_fraction_leaf=0.0)
```

Figure 5. Hyper parameter values given for rf algorithm in auto-sklearn library

```
LGBMClassifier(boosting_type="gbdp",
class_weight=None, colsample_bytree=1.0,
importance_type "split", learning_rate=0.1,
max_depth=-1, min_child_samples=20,
min_child_weight=0.001, min_split_gain=0.0,
n_estimators=100, n_job=-1, num_levels=31,
objective=None, random_state=4570,
reg_alpha=0.0, reg_lambda=0.0, silent=True,
subsample=1.0, subsample_for_bin=200000,
subsample_freq=0)
```

Figure 6. Hyper parameter values given for lgbm algorithm in pycaret library

After recording the learning models, tests have been carried out about the instantaneous anomaly detection on the stream data for the RO and LGBM algorithms, which gave the best results on the static data set for the Transport Unit. 11000 rows of data have been used for each test. As a result, a decision can be made by taking the average of the results of all tests performed. The averages of the tests performed for instant anomaly detection on the streaming data are given in Table 4.

Table 9. Transport unit, averages of tests performed on stream data.

		Normal (0)	Transported Product Anomaly (1)	Conveyor Distance Anomaly (2)	Conveyor Voltage Anomaly (3)	Conveyor Power Anomaly (4)	Accuracy
LGBM	Precision	0,975	1,000	0,711	0,970	0,683	0,947
	Recall	0,980	0,950	0,703	0,949	0,777	
	F-Score	0,977	0,974	0,707	0,959	0,727	
RO	Precision	0,975	1,000	0,653	0,977	0,695	0,945
	Recall	0,978	0,950	0,720	0,956	0,751	
	F-Score	0,977	0,974	0,685	0,967	0,722	

As can be seen in Table 4, there is no significant difference between the two algorithms. According to the results here,

although the LGBM algorithm seems to be slightly better, it has been determined that the generated simulation codes

do not show good results after they are converted from the notebook file to a python compatible .py file and transferred to the relevant container. For this reason, it has been decided to use the RO algorithm obtained from the Auto-Sklearn library in the Anomaly Detection Device of the Transport Unit.

4. CONCLUSION

As a result of the applied studies, it has been determined that the system, which was established as a simulation of the transport unit designed for the popcorn production system, can be used for anomaly detection by instantly evaluating the data streaming during the operation. It has been observed that the learning models obtained in the experiments on the system detected the correct anomaly type at the level of 91.3% on static data and at the level of 94.7% on stream data. In addition, an anomaly tracking screen has been created within the established system to support the auditor, and if the same type of anomaly is detected at least 20 times in the last 10 seconds, it is taken into consideration and displayed on the anomaly tracking screen.

In future studies, it is planned to design and implement the anomaly detection process to cover an entire production system, and to continue with instantaneous anomalies displayed on the screen together with the address information.

REFERENCES

- [1] A. G. Frank, L. S. Dalenogare, and N. F. Ayala, "Industry 4.0 technologies: Implementation patterns in manufacturing companies," *International Journal of Production Economics*, 210, 15–26, 2019.
- [2] R. Kumar, P. Gupta, S. Singh, and D. Jain, "Implementation of Industry 4.0 Practices in Indian Organization: A Case Study", **Advances in Industrial and Production Engineering. Lecture Notes in Mechanical Engineering**, Editor: Phanden R.K., Mathiyazhagan K., Kumar R., Paulo Davim J. Springer, Singapore, 35-44, 2021.
- [3] Z. Shi, Y. Xie, W. Xue, Y. Chen, L. Fu, and X. Xu, "Smart factory in Industry 4.0," *Systems Research and Behavioral Science*, 37(4), 607–617, 2020.
- [4] H. Wu, Y. Shen, X. Xiao, A. Hecker, and F. H. Fitzek, "In-Network processing of acoustic data for anomaly detection in smart factory", **2021 IEEE Global Communications Conference: IoT and Sensor Networks**, Madrid, Spain, 7-11 December, 2021.
- [5] N. Hranisavljevic, O. Niggemann, and A. Maier, "A Novel Anomaly Detection Algorithm for Hybrid Production Systems based on Deep Learning and Time Series Analysis", **The 27th International Workshop on Principles of Diagnosis: DX-2016**, Denver, Colorado, USA, 1-17, 4-7 October, 2016.
- [6] M.-O. Pahl and F.-X. Aubet, "All Eyes on You: Distributed Multi-Dimensional IoT Microservice Anomaly Detection", **14th International Conference on Network and Service Management**, Rome, Italy, 72-80, 5-9 November 2018.
- [7] R. Wang, K. Nie, T. Wang, Y. Yang, and B. Long, "Deep Learning for Anomaly Detection," *Proceedings of the 13th International Conference on Web Search and Data Mining*, Houston, TX, USA, 894-896, 3-7 February 2020.
- [8] C.-B. Seo, G. Lee, Y. Lee, and S.-H. Seo, "Echo-Guard: Acoustic-Based Anomaly Detection System for Smart Manufacturing Environments," *LNCIS*, 13009, 64–75, 2021.
- [9] Yılmaz Redüktör, **Mekanik Uygulamalar**, Uygulama Örnekleri, 2011.
- [10] P. Gijssbers, E. LeDell, J. Thomas, S. Poirier, B. Bischl, and J. Vanschoren, "An Open Source AutoML Benchmark", **6th ICML Workshop on Automated Machine Learning**, California, USA, 1-8, 11-13 June 2019.
- [11] M. A. Zöller and M. F. Huber, "Benchmark and Survey of Automated Machine Learning Frameworks," *Journal of Artificial Intelligence Research*, 70, 409–472, 2021.
- [12] H. Jin, Q. Song, and X. Hu, "Auto-Keras: An Efficient Neural Architecture Search System" **The 25th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining**, New York, USA, 1946–1956, 4–8 August 2019.
- [13] Ş. Özdemir and S. ÖRSLÜ, "Makine Öğrenmesinde Yeni Bir Bakış Açısı : Otomatik Makine Öğrenmesi (AutoML)", *Journal of Information Systems and Management Research*, 1(1), 23-30, 2019.
- [14] F. Fabris and A. A. Freitas, "Analysing the Overfit of the Auto-sklearn Automated Machine Learning Tool", **Analysing the Overfit of the Auto-sklearn Automated Machine Learning Tool**, Siena, Italy, 508-520, 10-13 September 2019.
- [15] M. Feurer, A. Klein, K. Eggensperger, J. T. Springenberg, M. Blum, and F. Hutter, "Auto-sklearn: Efficient and Robust Automated Machine Learning", **Automated Machine Learning**, Editor: F. Hutter, L. Kotthoff, J. Vanschoren, The Springer Series on Challenges in Machine Learning, Cham, Switzerland, 113–134, 2019.
- [16] A. Munjal, R. Khandia, and B. Gautam, "A Machine Learning Approach For Selection Of Polycystic Ovarian Syndrome (Pcos) Attributes And Comparing Different Classifier Performance With The Help Of Weka and Pycaret", *International Journal Of Scientific Research*, 9(12), 59–63, 2020.
- [17] U. Gain and V. Hotti, "Low-code AutoML-augmented Data Pipeline – A Review and Experiments," *Journal of Physics: Conference Series*, 1828(1), 1–8, 2021.

Sentiment, Stance, and Emotion Analysis on Twitter for COVID-19 Vaccination: A Brief Survey*

Doğan Küçük¹*, Nursal Arıcı²

¹Department of Computer Engineering, Gazi University, Ankara, Turkey

²Faculty of Applied Science, Gazi University, Ankara, Turkey
dogan.kucuk@gazi.edu.tr, nursal@gazi.edu.tr

Abstract—Social media analysis is used for various purposes ranging from trend analysis to recommender systems and public health monitoring. In this paper, we briefly review social media analysis studies conducted to determine people's sentiments, stances, and emotions towards COVID-19 vaccination. COVID-19 is an ongoing pandemic, and there exist debates on Twitter about COVID-19 vaccines. It is practical and promising to extract the opinions of Twitter users and consider them during health-related policy-making. Hence, there are several recent studies on this vital topic, and our paper aims to present a brief survey of the related significant studies. The findings of the reviewed papers and the current survey paper can contribute to health-related policy-making processes.

Keywords— social media analysis, Twitter, COVID-19, NLP, sentiment analysis, stance detection, emotion recognition

1. INTRODUCTION

COVID-19 is a global pandemic that is still effective all over the world since its starting in 2019. Due to the widespread use of social media in many countries, people also tend to share their opinions and feelings about the pandemic and its various aspects on social media. For instance, people like to publish social media posts about quarantines, vaccines, and face masks applied in several countries to prevent the spread of the disease. People also like to express their disease-specific symptoms and emotions on social media [1, 2].

COVID-19 vaccination is one of the controversial topics about which debates are observed on Internet forums, as news comments, and on social media sites. Both the proponents and opponents of vaccination are freely sharing their ideas and arguments. This textual data is a valuable source for public health professionals and decision-makers. The ideas and arguments of individuals can be observed and taken into consideration during related policy-making processes. Several related papers claim that these posts can be analyzed to understand vaccine hesitancy and vaccine confidence better. The vaccine hesitancy of the people can be decreased by considering people's concerns about the side effects of the vaccines. However, due to the large number of posts published instantly, automated social media analysis approaches must obtain useful information from these posts.

Three prevalent social media analysis techniques are sentiment analysis, stance detection, and emotion recognition. The current paper aims to review those papers in which sentiment, stance, and emotion analysis are applied to tweets on COVID-19 vaccination. Since it is a very recent topic, all papers considered have been published in 2020 and 2021, and almost all are published in 2021. Yet, due to the topic's significance, the number of related papers is increasing, and there is a need for a survey of these related papers.

To the best of our knowledge, this is the first review paper prepared on this topic of sentiment, stance, and emotion analysis on tweets about COVID-19 vaccination. The rest of this paper is organized as follows: the three social media techniques are described in Section 2. The actual papers on social media analysis techniques performed on tweets about COVID-19 vaccination are reviewed in Section 3. Section 4 presents future research topics, and Section 5 concludes the paper with a summary.

2. SOCIAL MEDIA ANALYSIS TECHNIQUES

Sentiment analysis, stance detection, and emotion recognition are related and vital problems of social media analysis and natural language processing (NLP). They can be considered text classification tasks, and

more recently, they are mainly applied to social media posts.

The rest of this section briefly describes these three problems before reviewing the studies that use these techniques to analyze social media posts for COVID-19 vaccines.

2.1. Sentiment Analysis

In the related literature, *sentiment analysis* is commonly defined as the determination of the polarity in a given text into one of these classes: positive, negative, and neither [3]. Sometimes, only two sentiment classes (positive and negative) can be used, while in some other studies, a more significant number of sentiment classes (such as highly positive, positive, etc.) can be used. There is a large body of related studies on sentiment analysis, and there are subproblems of sentiment analysis such as aspect-based sentiment analysis. In aspect-based sentiment analysis, the sentiments towards an aspect or a set of aspects of an entity are usually explored. There are rule-based, lexicon-based, machine learning-based, and deep learning [4] based approaches to sentiment analysis in the related literature. Sentiment analysis is also known as opinion mining, and however, sentiment analysis and opinion mining are also used to mean emotion recognition (defined in Section 2.3 below) in some studies.

2.2. Stance Detection

In stance detection (or stance prediction), there is a piece of text and a target as input and the position (stance) of the text author towards the given target is expected as output [5]. Most commonly, there are three stance classes (favor, against, and neither), however, in some studies, only two basic classes (favor and against) are utilized, and in some other studies, a neutral class label is added to the existing stance classes [5]. Instead of favor and against classes of stance detection, pro and anti may also be used in some related studies. There are subproblems of stance detection, such as multi-target stance detection and cross-target stance detection. In multi-target stance detection, the input does not include a single item as a target, but instead, it consists of a set of targets. The training dataset is provided for a target in cross-target stance detection, but the test set is available for a related but different target [5]. Some studies jointly perform stance detection and sentiment analysis where annotated datasets are also available [6]. The most common approaches to stance detection are either machine learning or deep learning-based.

2.3 Emotion Recognition

Emotion recognition (or emotion extraction) in textual content aims to determine the emotion in the given text in various granularity levels [7]. That is, the number of emotion class labels changes in different studies. In

some studies, less number of labels are used (joy, sadness, fear, disgust) while in other studies more labels are used (joy, sadness, fear, disgust, love, trust, surprise, anger, anxiety, anticipation). It should again be noted that some studies use the term sentiment (or opinion) analysis to mean emotion recognition. Emotion recognition is also commonly applied to other input genres like speech or videos. In the related literature, lexicon-based, machine learning-based, and deep learning-based methods are all used for emotion recognition.

The interrelationships and common classes of these three research problems are presented in Figure 1.

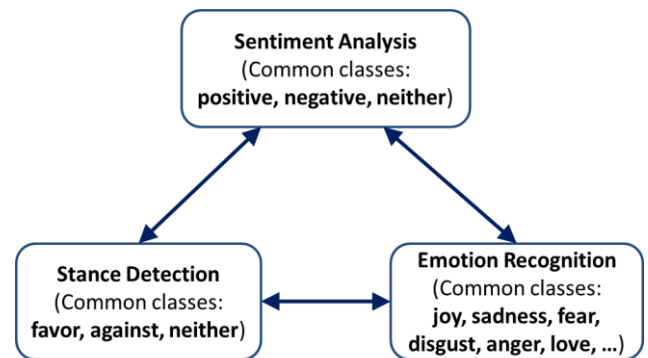


Figure 1. Interrelationships and classes of sentiment analysis, stance detection, and emotion recognition

3. SURVEY OF PAPERS ON SOCIAL MEDIA ANALYSIS FOR COVID-19 VACCINATION

As pointed out in Section 1, there are several recent works on using social media analysis to determine public attitudes and opinions about COVID-19 vaccines. In [8], it is emphasized that artificial intelligence, and particularly NLP techniques, can be used for social media analysis for COVID-19 vaccines. It is concluded that by analyzing the sentiments in social media and linked Web platforms, public concerns and barriers against the vaccines can be determined, and this will help develop strategies to increase the uptake of COVID-19 vaccines [8].

In [9], 300,000 social media posts about COVID-19 vaccines from the US and UK are extracted, and deep learning-based sentiment analysis (with three classes: positive, negative, and neutral) is applied to these posts. The study's findings show that positive sentiment is higher (over 55%) for posts from both countries [9].

A related study was performed between February and March of 2021 on social media posts of US citizens to determine their sentiments about COVID-19 vaccines [10]. They have performed a spatiotemporal analysis of public sentiment towards the vaccines using publicly available sentiment analysis tools for this purpose [10].

The study in [11] presents sentiment analysis results using deep learning methods on tweets from six countries: Pakistan, India, Norway, Sweden, Canada, and the US. Correlation percentages for sentiment classes among neighboring countries are also presented [11].

In [12], emotion analysis with emotion labels (anger, fear, joy, sadness) is performed on Reddit comments to determine the public emotions towards COVID-19 vaccines in Canada. Random forest regression is used for automatic emotion analysis. It is pointed out that local (city-level) comments should be analyzed to better grasp people's local concerns and opinions towards the vaccine [12].

In a recent study on Australian Twitter users, both sentiment analysis (with two classes: positive, negative) and emotion recognition (with eight classes: anger, fear, anticipation, trust, surprise, sadness, joy, disgust) are performed using publicly available NLP tools [13]. It is concluded in the study that the level of positive sentiments across Twitter users may be insufficient for herd immunity, and such analysis results should be considered by the related governmental organizations [13].

In [14], both sentiment analysis and emotion recognition are performed on geotagged tweets from the US. The authors observe an increased positive and decreased negative trend in sentiments, and for emotion recognition; trust and anticipation along with fear, sadness, and anger are observed [9].

Another study [15] performs sentiment analysis on tweets related to COVID-19 vaccine types using the kNN machine learning algorithm. The sentiment classes are positive, negative, and neutral, and percentages of these classes are provided for each of the COVID-19 vaccine types [15].

In [16], tweets about COVID-19 vaccination posted from five countries are analyzed for emotions (hesitation, rage, sorrow, anticipation, faith, and contentment). The countries considered are India, the US, Brazil, the UK, and Australia, and it is reported that vaccine hesitancy varies in different geographies [16].

Emotion recognition is emphasized in [17] in order to address related issues like vaccine hesitancy and vaccine confidence. It is pointed out that suitable communication strategies should be used to identify the barriers and facilitators of COVID-19 vaccination [17].

Topic modeling and sentiment analysis are performed on tweets about COVID-19 vaccines using machine learning (logistic regression, random forest, SVM) and transfer learning methods in [18]. It is pointed out that these results can be used by education programs and other interventions to increase the acceptance of

COVID-19 vaccines [18]. In a subsequent study [19], the sentiment analysis results on the same tweet dataset of 2,678,372 tweets are presented as follows using a publicly available sentiment analyzer, called VADER [20]: 42.8% of the tweets about COVID-19 vaccines have positive sentiment while 30.3% of them have negative sentiment [19].

In [21]; topic modeling, sentiment analysis, and emotion recognition are performed on tweets about COVID-19 vaccination. Based on the reported results, the sentiments of Twitter users are increasingly positive, and the most frequent emotions are trust, anticipation, fear, and sadness [21]. In [22], Twitter conversations about COVID-19 vaccines are analyzed and mixed emotions about vaccine safety and its side effects are observed.

In [23], a total of 7,948,886 tweets about COVID-19 vaccines are analyzed for topic modeling and emotion recognition, using methods including deep learning-based ones like BERT [24]. Among the five emotions considered, it is found that fear is the most dominant emotion, and it is followed by joy [23].

A sentiment analysis study on tweets from India about COVID-19 vaccines is presented in [25]. It is concluded that 36% of the tweets have positive sentiment, 17% have negative sentiment, and 47% have a neutral sentiment. It is also presented that the tweets with positive sentiments increase as the number of COVID-19 cases increases [25].

Another topic modeling, sentiment analysis, and emotion recognition study on tweets about COVID-19 vaccines is performed in [26]. The overall sentiment is positive and while fear is the most frequent emotion in the earlier period, trust emotion has increased over time [26].

In [27]; TextBlob and VADER, sentiment analysis tools, are used on tweets about COVID-19 vaccines where three sentiment classes are considered: positive, negative, and neutral. It is found that the sentiments of the tweets are primarily positive [27].

A study performing emotion recognition on 4,765 popular tweets in English and Italian about COVID-19 vaccines is described in [28]. Knowledge graphs of semantic and syntactic associations are used in the study, and it is found that for English tweets, the emotions are polarized between trust/anticipation and anger/sadness [28].

Stance detection towards COVID-19 vaccination at the early days of vaccine announcements is performed in [29]. Based on stance analysis on collected tweets, it is found that most tweets have a neutral stance, while tweets with favor stance are higher than tweets with against stance [29].

Another stance detection study is reported in [30] regarding COVID-19 vaccines on Twitter, where the stance class labels are pro-vaccine, anti-vaccine, and neutral. About 1,000 tweets are examined, and it is recommended that the findings of this study can be used in national campaigns to increase vaccine application [30].

A summary of the reviewed papers with respect to the research problem(s) that are considered is presented in Table 1.

Table 1. A summary of the reviewed papers

Reference	Sentiment Analysis	Stance Detection	Emotion Recognition
[9]	✓	✗	✗
[10]	✓	✗	✗
[11]	✓	✗	✗
[12]	✗	✗	✓
[13]	✓	✗	✓
[14]	✓	✗	✓
[15]	✓	✗	✗
[16]	✗	✗	✓
[17]	✗	✗	✓
[18]	✓	✗	✗
[19]	✓	✗	✗
[21]	✓	✗	✓
[22]	✗	✗	✓
[23]	✗	✗	✓
[25]	✓	✗	✗
[26]	✓	✗	✓
[27]	✓	✗	✗
[28]	✗	✗	✓
[29]	✗	✓	✗
[30]	✗	✓	✗

4. FUTURE RESEARCH DIRECTIONS

Most of the studies reviewed in the previous section emphasize that social media analysis for COVID-19 vaccination is essential for related policy-makers to better understand the public concerns and opinions about the vaccination and related process. Several studies are conducted to determine vaccine hesitancy in different communities of the population, such as [31]. Based on the findings of the reviewed papers, the following future research studies can be conducted:

1. The sizes of the related publicly available datasets (of social media posts) should be increased in order

to increase the performance of the learning systems. Additionally, the number of datasets in languages other than English should also be increased. Creating annotated social media posts in various languages (multilingual datasets) is also very important.

2. The reviewed papers demonstrate that related studies are performed on related tweets from several countries, including India, the US, the UK, Canada, and Australia. Similar studies can also be applied to tweets from other countries, and the results can be compared with previous work.
3. The number of studies that jointly use different social media analysis tasks (like the three tasks considered in the current paper) at the same time can be increased. Because each task produces a distinct and complementary output, the final output will be more beneficial for practical purposes.

5. CONCLUSION

Social media analysis is essential for health informatics, in addition to other diverse application areas. Because of the recent COVID-19 pandemic, social media posts have become crucial information sources of public opinions and ideas about the pandemic. This paper presents a brief review of sentiment, stance, and emotion extraction studies on social media posts about COVID-19 vaccination. Although the COVID-19 pandemic has started in 2019, there are several social media analysis studies about COVID-19 and its various aspects (like the lockdowns, masks, and vaccines). This survey paper is significant because the findings of the reviewed recent papers can be beneficial to the decision and policy-making procedures related to COVID-19 vaccination. Thereby, the applications of vaccination can be increased to better cope with this global pandemic.

REFERENCES

- [1] A. Al-Laith, M. Alenezi, "Monitoring people's emotions and symptoms from Arabic tweets during the COVID-19 pandemic". *Information*, 12(2), 86, 2021.
- [2] D. Küçük, N. Arıcı, E. E. Küçük, "Sosyal medyada otomatik halk sağlığı takibi: Güncel bir derleme (Automatic public health monitoring on social media: A recent survey)". *Niğde Ömer Halisdemir Üniversitesi Mühendislik Bilimleri Dergisi*, 10(2), 2021.
- [3] W. Medhat, A. Hassan, H. Korashy, "Sentiment analysis algorithms and applications: A survey". *Ain Shams Engineering Journal*, 5(4), 1093-1113, 2014.
- [4] D. Küçük, N. Arıcı, "Doğal dil işlemede derin öğrenme uygulamaları üzerine bir literatür çalışması (A literature study on deep learning applications in natural language processing)". *Uluslararası Yönetim Bilişim Sistemleri ve Bilgisayar Bilimleri Dergisi*, 2(2), 76-86, 2018.
- [5] D. Küçük, F. Can, "Stance detection: A survey". *ACM Computing*

- Surveys (CSUR), 53(1), 1-37, 2020.
- [6] S. M. Mohammad, P. Sobhani, S. Kiritchenko, "Stance and sentiment in tweets". *ACM Transactions on Internet Technology (TOIT)*, 17(3), 1-23, 2017.
- [7] B. Gained, V. Syal, S. Padgalwar, "Emotion detection and analysis on social media", arXiv preprint arXiv:1901.08458, 2019.
- [8] A. Hussain, A. Sheikh, "Opportunities for artificial intelligence-enabled social media analysis of public attitudes toward Covid-19 vaccines", *NEJM Catalyst Innovations in Care Delivery*, 2(1), 2021.
- [9] A. Hussain, A. Tahir, Z. Hussain, Z. Sheikh, M. Gogate, K. Dashtipour et al., "Artificial intelligence-enabled analysis of public attitudes on Facebook and Twitter toward COVID-19 vaccines in the United Kingdom and the United States: Observational study", *Journal of Medical Internet Research*, 23(4), e26627, 2021.
- [10] G. G. Ali, M. M. Rahman, A. Hossain, S. Rahman, K. C. Paul, J. C. Thill, J. Samuel, "Public perceptions about COVID-19 vaccines: Policy implications from us spatiotemporal sentiment analytics", SSRN, 3849138, 2021.
- [11] R. Batra, A. S. Imran, Z. Kastrati, A. Ghafoor, S. M. Daudpota, S. Shaikh, "Evaluating polarity trend amidst the coronavirus crisis in peoples' attitudes toward the vaccination drive", *Sustainability*, 13(10), 5344, 2021.
- [12] C. Yan, M. Law, S. Nguyen, J. Cheung, J. Kong, "Comparing public sentiment toward COVID-19 vaccines across Canadian cities: Analysis of comments on Reddit", *Journal of Medical Internet Research*, 23(9), e32685, 2021.
- [13] S. W. H. Kwok, S. K. Vadde, G. Wang, "Tweet topics and sentiments relating to COVID-19 vaccination among Australian Twitter users: Machine learning analysis", *Journal of Medical Internet Research*, 23(5), e26953, 2021.
- [14] T. Hu, S. Wang, W. Luo, Y. Yan, M. Zhang, X. Huang et al., "Revealing public opinion towards COVID-19 vaccines using Twitter data in the United States: a spatiotemporal perspective", medRxiv, 2021.
- [15] F. M. J. M. Shamrat, S. Chakraborty, M. Imran, J. N. Muna, M. M. Billah, P. Das, M. O. Rahman, "Sentiment analysis on Twitter tweets about COVID-19 vaccines using NLP and supervised KNN classification algorithm", *Indonesian Journal of Electrical Engineering and Computer Science*, 23(1), 463-470, 2021.
- [16] H. Chopra, A. Vashishtha, R. Pal, A. Tyagi, T. Sethi, "Mining trends of COVID-19 vaccine beliefs on Twitter with lexical embeddings", arXiv preprint arXiv:2104.01131, 2021.
- [17] W. Y. S. Chou, A. Budenz, "Considering emotion in COVID-19 vaccine communication: addressing vaccine hesitancy and fostering vaccine confidence", *Health Communication*, 35(14), 1718-1722, 2020.
- [18] S. Liu, J. Li, J. Liu, "Leveraging transfer learning to analyze opinions, attitudes, and behavioral intentions toward COVID-19 vaccines: Social media content and temporal analysis", *Journal of Medical Internet Research*, 23(8), e30251, 2021.
- [19] S. Liu, J. Liu, "Public attitudes toward COVID-19 vaccines on English-language Twitter: A sentiment analysis", *Vaccine*, 39(39), 5499-5505, 2021.
- [20] Internet: C. Hutto, E. Gilbert, VADER Sentiment Analyzer, <https://github.com/cjhutto/vaderSentiment>, 12.18.2021.
- [21] J. C. Lyu, E. Le Han, G. K. Luli, "COVID-19 vaccine-related discussion on Twitter: topic modeling and sentiment analysis", *Journal of Medical Internet Research*, 23(6), e24435, 2021.
- [22] E. Malova, "Understanding online conversations about COVID-19 vaccine on Twitter: vaccine hesitancy amid the public health crisis", *Communication Research Reports*, 1-11, 2021.
- [23] M. Monselise, C. H. Chang, G. Ferreira, R. Yang, C. C. Yang. "Topics and sentiments of public concerns regarding COVID-19 vaccines: Social media trend analysis", *Journal of Medical Internet Research*, 23(10), e30765, 2021.
- [24] J. Devlin, M. W. Chang, K. Lee, K. Toutanova, "BERT: Pre-training of deep bidirectional transformers for language understanding", arXiv preprint arXiv:1810.04805, 2018.
- [25] Praveen, S. V., Ittamalla, R., & Deepak, G. (2021). "Analyzing the attitude of Indian citizens towards COVID-19 vaccine—A text analytics study", *Diabetes & Metabolic Syndrome: Clinical Research & Reviews*, 15(2), 595-599.
- [26] S. N. Saleh, S. A. McDonald, M. A. Basit, S. Kumar, R. J. Arasaratnam, T. M. Perl et al., "Public perception of COVID-19 vaccines through analysis of Twitter content and users", medRxiv, 2021.
- [27] N. S. Sattar, S. Arifuzzaman, "COVID-19 vaccination awareness and aftermath: Public sentiment analysis on Twitter data and vaccinated population prediction in the USA", *Applied Sciences*, 11(13), 6128, 2021.
- [28] M. Stella, M. S. Vitevitch, F. Botta, "Cognitive networks identify the content of English and Italian popular posts about COVID-19 vaccines: Anticipation, logistics, conspiracy and loss of trust", arXiv preprint arXiv:2103.15909, 2021.
- [29] L. A. Cotfas, C. Delcea, I. Roxin, C. Ioanăș, D. S. Gherai, F. Tajariol, "The longest month: Analyzing COVID-19 vaccination opinions dynamics from tweets in the month following the first vaccine announcement", *IEEE Access*, 9, 33203-33223, 2021.
- [30] D. Scannell, L. Desens, M. Guadagno, Y. Tra, E. Acker, K. Sheridan et al., "COVID-19 vaccine discourse on Twitter: A content analysis of persuasion techniques, sentiment and mis/disinformation", *Journal of Health Communication*, 26(7), 443-459, 2021.
- [31] L. J. Lim, A. J. Lim, K. K. Fong, C. G. Lee "Sentiments regarding COVID-19 vaccination among graduate students in Singapore", *Vaccines*, 9(10), 1141, 2021.

Accessibility Evaluation of e-Nabız Web Site*

Mehmet Fevzi Akdeniz^{1*}, Pinar Onay Durdu²

¹ Department of Computer Engineering, Kocaeli University, Kocaeli, Turkey

² Department of Computer Engineering, Kocaeli University, Kocaeli, Turkey
m.fevziakdeniz@gmail.com, pinar.onaydurdu@kocaeli.edu.tr

Abstract— It is important that websites are accessible so that everyone can benefit equally from the offered web services. In this study, the accessibility evaluation of the e-Nabız Personal Health System website was carried out using automated evaluation tools, since it is a public website that has many users today and its use has been increased even more during the current pandemic. According to the evaluation results, various accessibility problems have been determined on the e-Nabız website such as there are some links that are not meaningful and consistent with the content they are related to, and there are no descriptive labels and directions for data-entry fields. Solutions are provided for the determined problems.

Keywords— accessibility evaluation, website accessibility, WCAG 2.0, automated evaluation, e-Nabız, e-government

1. INTRODUCTION

Nowadays, Internet is one of the fundamental methods for accessing information. Websites for various domains, such as public services, education, research, commerce, health, business, and social media have many diverse users. Especially most public services are now offered through e-government websites with the advancement of technology. Therefore, it has become much more important for these websites to serve users from all segments of society as much as possible [1]. e-Government websites should be accessible to all users, regardless of users' abilities and disabilities [2]. Website accessibility is the availability of web content to all individuals, regardless of any barriers or environmental restrictions [3]. On the other hand, according to the World Health Organization data, 15% of the world population live with a disability and this rate is expected to increase in the following years [4].

Various studies are carried out to make websites accessible to all users. The Web Content Accessibility Guidelines (WCAG) has been developed by the World Wide Web Consortium (W3C) to standardize the accessibility of websites and to ensure equal access and opportunity for all users [3]. In addition, some countries require accessibility with laws and regulations such as ISO 9241-210 [5] or Section 508 [6]. The Public Websites Guideline (KAMİS) Project [2] was implemented by the Tubitak Bilgem Software Technologies Research Institute (YTE) and the Ministry of Development to guide the public websites in Turkey to comply with accessibility standards and to raise national awareness about accessibility.

There are various methods to evaluate the accessibility of websites. These methods are conformance tests with automated evaluation tools, user tests or expert-

based/manual evaluations [7]. Automated evaluation tools are tools that evaluate the accessibility of websites within the framework of accessibility guidelines. Expert-based evaluation is the method in which an expert evaluates a website interface or similar system against a checklist manually. These methods do not require the involvement of end users in the evaluation process except for user tests in which evaluations are conducted with website users while interacting with the system.

Automated evaluation tools are frequently preferred by researchers as they allow evaluating many web pages quickly and easily. However measuring the accessibility of a website with only one automated evaluation tool does not provide reliable results [8]. In addition, accessible rich internet applications (ARIA), which have become widespread with the development of technology, can negatively affect the accessibility evaluations of automated evaluation tools [9]. Therefore, using more than one evaluation tool or using this method together with other evaluation methods in evaluation will provide more reliable results [10].

Within the scope of the study, the accessibility of the e-Nabız Personal Health System website, whose number of users has reached 10 million today [11] and whose use has increased even more during the Covid19 pandemic experienced as of 2020, was examined. e-Nabız, as a health system that includes all kinds of detailed information about the users' examinations, has a very large user base. The study was conducted by using more than one automated evaluation tool to increase the accuracy of the results obtained. According to the results of the evaluation, accessibility problems on the e-Nabız website were

identified and solutions were provided. It was also aimed that the findings would guide developers to increase the accessibility of e-government websites.

2. BACKGROUND OF THE STUDY

2.1. Accessibility and Accessibility Evaluation Guidelines

Various definitions have been made in the literature for web accessibility. According to the definition of the Web Accessibility Initiative (WAI) [12], web accessibility means the design and development of websites, tools and technologies so that people with disabilities can use them. In other words, accessibility is the removal of all technical barriers to effective interaction [13]. Another definition for web accessibility is that technology can be used by people with disabilities as much as those without disabilities [6]. However, the concept of web accessibility should not be considered only for users with disabilities.

The WAI, which was founded by the World Wide Web Consortium (W3C) in 1997, has started to develop standards to increase web accessibility for all users to receive equal service from websites [14]. The W3C first developed the first version of the Web Content Accessibility Guidelines (WCAG) 1.0 for this purpose in 1999 [15]. Later, to adapt to the developing technology, this guide was updated in 2008 with the name WCAG 2.0 [16]. The WCAG 2.0 guide was accepted as an internationally valid document (ISO/IEC 40500:2012) on web accessibility by the International Organization for Standardization (ISO) in 2012 [17]. In 2018, the WCAG 2.1 guide was published with the update of the existing WCAG 2.0 guideline to include new technologies that have become widespread with mobile devices [18].

WCAG 2.0 is organized by four main principles which are perceivable, operable, understandable, and robust. There are 12 guidelines defined within the framework of these principles, and 61 success criteria defined to measure the accessibility of each of these guidelines. These criteria are divided into three levels of conformance which are A (25 criteria), AA (13 criteria), and AAA (23 criteria). Level A means that website conforms to guideline items at minimum level while Level AAA satisfies all the Level A, Level AA.

There are also different organizations or studies carried out in different countries to improve web accessibility [2], [6], [19], [20].

2.2. Accessibility Evaluation Methods

Web accessibility evaluation is a necessary process to check whether a website meets referenced accessibility standards [21]. It can be defined as the evaluation of how well the website under consideration can be used by disabled users [22]. Web accessibility evaluation methods can be grouped under three main categories [21]. These are as follows;

Automated evaluation is the evaluation of the accessibility of websites by the use of automated tools that check the compliance of websites within the framework of guidelines [23]. Automated evaluation tools are highly preferred by researchers as they reduce the cost considerably and provide rapid results. However, the W3C/WAI evaluated that “Web accessibility evaluation tools cannot determine the accessibility of websites, they can only help to do so” [24]. In addition, when some studies [8], [25] are taken into account, it is seen that using only automated tools in accessibility assessments is not sufficient to obtain reliable results.

Expert evaluation is the assessment of a website by an expert for compliance with accessibility guidelines [26]. While this approach is more accurate than automated evaluation in detecting more critical accessibility violations, it requires more time and effort [27].

User tests are carried out with a group of users while using the website to determine the problems they come up with [27]. This method is the most reliable accessibility evaluation method since it contains the findings obtained from the real experiences of users. However, it is time-consuming and costly to perform user tests [21].

3. RELATED WORKS

There are various studies in the literature on the accessibility of public websites. Some examples are summarized in Table 1.

Table 1. Studies on website accessibility

Study	Evaluation method	Automated evaluation tools used	Guideline
[23]	Automated evaluation	- WaaT	WCAG 2.0
[28]		- TAW	WCAG 2.0
[29]		- Achecker - TAW - eXaminator - WAVE - Total Validator - Amp - EvalAccess - HERA - Cynthia Says - WaaT - MAGENTA - SortSite	WCAG 2.0 WCAG 1.0 Section 508
[30]		- The EvalAccess 2.0	WCAG 1.0
[31]		- Achecker	WCAG 2.0
[26]		- Achecker - WAVE - Total validator	WCAG 2.0
[10]		- WAVE - Koal1y	WCAG 2.0
[32]	User test	-	Unspecified
[7]	Expert evaluation	-	WCAG 1.0
[33]		-	WCAG 2.0
[34]		-	WCAG 2.0

In these studies, it is seen that different evaluation methods, especially automated evaluation tools, are used. While some of the studies [23], [28], [30] used a single automated evaluation tool, some [10], [29] used more than one automated evaluation tool. In studies conducted with

automated evaluation tools, some common problems such as graphic elements that cannot be read by screen readers, inability to navigate the site due to empty links, lack of text alternatives for graphic elements, and dynamic elements that do not work properly have been identified [28], [29]. In addition, it has been emphasized that different automated evaluation tools provide different results on the same site, so it is important to increase the reliability of the results by using more than one evaluation tool [29]. In studies in which more than one evaluation method was used [10], [26] the findings obtained with automated evaluation tools were supported by expert evaluation [7], [33], [34]. In the study conducted with user tests [32], the participants were given tasks according to their difficulty levels and they were asked to perform these tasks through a screen reader program.

4. METHODOLOGY

Within the scope of the study, the accessibility of the e-Nabız Personal Health System website was evaluated. The importance of the e-Nabız website has become even more important during the Covid-19 pandemic. Automated evaluation tools were used for the accessibility evaluation of this website.

Automated evaluation is one of the web accessibility evaluation methods that do not include end users [8]. There are many automated evaluation tools available, either paid or free [35]. These tools are different in terms of their cost, target, domain of use, breadth of accessibility reports, compliance guideline support, and supported compliance levels [23]. In this study, automated evaluation tools are determined based on the usability, the conformance with WCAG 2.0 A and AA compliance levels and reporting support criteria of the tools. Thus, AChecker (AC), Cynthia Says (CS), TAW and WAVE tools were included in the research. Considering that automated evaluation tools have different advantages and disadvantages [36] and using more than one automated evaluation tool for more reliable results [29], [36], more than one automated evaluation tool was preferred to be used.

Another important issue is to which page or pages of the website the evaluation will be applied. In various studies [23], [28], [29], the accessibility evaluation was carried out only with the home pages of the websites. In these studies, it is stated that the accessibility problems experienced by the users on the home page, which is the first contact with the website, will make it difficult to access the other pages of the website. Therefore, the accessibility evaluation of the e-Nabız website was carried out on the home page of the site, which is shown in Figure 1. Evaluation was carried out in November 2021.

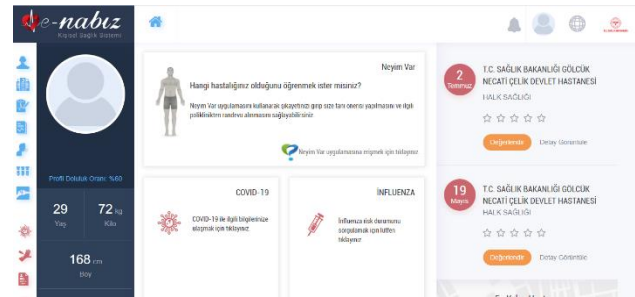


Figure 1. e-Nabız website home page

5. RESULTS

A comparison of the evaluation results according to the compliance level and accessibility principle is presented in Table 2.

Table 2. Violations by accessibility principle

	AC	CS	TAW	WAVE
1. Perceivable (340)	130	49	94	67
2. Operable (78)	10	16	9	43
3. Understandable (109)	41	15	33	20
4. Robust (83)	1	5	73	4
Total level A errors (513)	127	66	209	111
Total level AA errors (97)	55	19	0	23
TOTAL (610)	182	85	209	134

The number of A-level errors in the WCAG 2.0 accessibility guideline was determined as 127 by the AC automated evaluation tool, 66 by CS, 209 by TAW, and 111 by WAVE. Most of the level A errors in the perceivable principle are in criteria 1.1.1 (non-text content) and criteria 1.3.1. (info and relationships). Criteria 1.1.1 specifies that information conveyed by non-text content should be accessible using a text alternative [37]. Criteria 1.3.1 aims to ensure that information and relationships implied by visual or auditory formatting are preserved when the presentation format changes [38]. A-level errors in the operable principle are centered around criteria 2.4.4 (link purpose (in context)). This criteria states that the name of a document or web application presented on a web page should be sufficient to explain the purpose of the page [37]. The most A-level errors in the understandable principle were in criteria 3.3.2 (labels or instructions), which states that the user entering data into a form should be guided through tags or instructions [37]. The errors in the robust principle are all A-level, with the most criteria pertaining to 4.1.2 (name, role, value). This criteria aims to ensure that Assistive Technology (AT) collects, activates, adjusts, and keeps up-to-date on the status of user interface controls in the content [37]. AT is any item, piece of equipment, software program, or product system used to increase, maintain, or improve the functional abilities of persons with disabilities [39].

Regarding AA level errors, while the AC auto-evaluation tool detects 55 errors, this number is 19 in CS and 23 in WAVE. TAW found no AA level errors. Level AA errors

in the perceivable principle are mostly related to criteria 1.4.4 (resize text) and 2.4.6 (headings and labels). Criterion 1.4.4 specifies that text can be resized without assistive technology up to 200 percent without loss of content or functionality [37]. Criteria 2.4.6 states that it should help users understand what information is found on web pages and how that information is organized.

Since there are many different criteria violated in the evaluation results and it is difficult to include all of them in the study, the "most violated criteria" items are reported. Criteria with a total of 25 or more violations detected by each automated assessment tool were defined as the most violated criteria. The most violated accessibility criteria in the evaluation results and the number of occurrences is presented in Table 3.

Table 3. The most violated evaluation criteria and the number of detections

Criteria	Explanation	AC	CS	TAW	WAVE
1.1.1	Non-text content	9	15	34	40
1.3.1	Info and relationships	67	10	60	27
1.4.4	Resize text	46	2	0	0
2.4.4	Link purpose (in context)	9	4	9	16
2.4.6	Heading and labels	1	1	0	23
3.3.2	Labels or instructions	41	3	33	20
4.1.2	Name, role, value	0	4	44	4

6. CONCLUSIONS AND RECOMMENDATIONS

In this paper, it was aimed to evaluate the accessibility of the e- Nabız Personal Health website. The evaluation was carried out on the main page of the site, which is the first interaction of the users with the site. WCAG 2.0 was used as an accessibility evaluation guide. More than one automated evaluation tool was used in the study, but it is difficult to reach a definitive conclusion as each of these automated evaluation tools produces different results. However, it was also determined that some important criteria were violated on the site, which would make it difficult for disabled users to interact with the website.

Lack of text alternatives, no descriptive labels and redirection for data entry fields, and tables without column headers or form fields that are not specified as mandatory are some of the most common accessibility violations detected in evaluation. These violations have also been frequently reported in previous studies on the accessibility assessment of public websites in Turkey [8], [28], [29]. Failure to resize the text on the page, not understanding the purpose of the link, clear, meaningful, and consistent titles and tags, and finally, the presence of elements in the source code that are not assigned names, roles and values are other commonly perceived violations in the evaluation. The problems that the violations may cause to the users and the solution recommendations are presented in Table 4.

Table 4. Problems and solution recommendations due to accessibility violations [37], [40]

Violation	Problem	Solution
No text alternatives	- People who have difficulty in perceiving visual content have problems using the site - Inability of assistive technology to read text aloud and present it visually	- Using tag elements to define form controls - Using the LONGDESC attribute to indicate a longer summary if an image cannot be fully identified
Tables without column headers or form fields that are not required or not	- Failure to protect information and relationships when the presentation format changes	- Using tags to associate form tags with their elements - Using color and text to indicate required fields - Using TS with the scope attribute to specify the table structure
Unable to resize text	- Inability to directly read text-based or visually-created text by people with mild visual impairment without the use of assistive technologies	- Specifying font sizes in relative units such as em, percentages, or using named font sizes - Using a page layout that expands, contracts, and reflows as the browser viewport expands or contracts - Use real text instead of text images whenever possible
Not understanding the purpose of the link	- The user has difficulty understanding the purpose of the link and thus cannot decide whether or not to follow the link	- Writing the link text to explain the target of the link - Using the same text in all links to the same destination - Increasing the descriptiveness of the link by adding a descriptive title attribute
Headings and tags are not clear, meaningful and consistent	- The user cannot understand what information is contained in the web pages	- Each title should explain its own part of the text
Not including descriptive labels and redirects for data entry fields	- Confusion on the page where the user will enter data	- Giving examples of expected values and format - Using an access link to a reference page that shows the correct format - Adding instruction text at the beginning of the form
Finding elements in the source code that are not assigned a name, role, or value	- Assistive Technologies cannot obtain information about the status of UI controls in content	- Replacing the alt text with it when an image is scripted - All input items have a name attribute

Accessibility evaluation of e-Nabız website was conducted by automated tools in the current study. Although more than one automated tool was used, using more than one evaluation method is also emphasized to get more reliable results in the literature [10]. Therefore, conducting more enhanced evaluation study with implementing other evaluation methods such as user study or Barrier walkthrough is planned as a future work.

REFERENCES

- [1] A. Doğan, **Türkiye’deki devlet kurumlarına ait web sitelerinin web erişilebilirliği durumu ve bölgeler arasındaki farklılıkların incelenmesi**, Master’s Thesis, Bursa Uludağ University, Institute of Education Sciences, 2019.
- [2] Internet: Erişilebilirlik, <https://dijitalakademi.bilgem.tubitak.gov.tr/kamis/erisilebilirlik/erisilebilirlik>, 06.11.2021.
- [3] Internet: Accessibility, <https://www.w3.org/standards/webdesign/accessibility>, 06.11.2021.
- [4] World Health Organization, **Dünya Engellilik Raporu: Yönetici Özeti**, 2011.
- [5] Internet: ISO, ISO 9241-210: Ergonomics of human–system interaction - Human-centred design for interactive systems, <https://www.iso.org/standard/77520.html>, 06.11.2021.
- [6] Internet: What is Section 508?, <https://www.epa.gov/accessibility/what-section-508>, 06.11.2021.
- [7] Internet: G. Brajnik, Barrier Walkthrough, <https://users.dimi.uniud.it/~giorgio.brajnik/projects/bw/bw.html>, 06.11.2021.
- [8] S. Kurt, "Accessibility of Turkish university Web sites", *Universal Access in the Information Society*, 16(2), 505-515, 2017.
- [9] I.A. Doush, F. Alkhateeb, E.A. Maghayreh, and M.A. Al-Betar, "The design of RIA accessibility evaluation tool", *Advances in Engineering Software*, 57, 1-7, 2013.
- [10] P. Acosta-Vargas, S. Luján-Mora, T. Acosta, ve L. Salvador-Ullauri, "Toward a combined method for evaluation of web accessibility", **International Conference on Information Technology & Systems**, Libertad City, Ecuador, 602–613, January 10 – 12, 2018.
- [11] Internet: 10 Milyon Kişi e-Nabız Kullanıyor, <https://sbsgm.saglik.gov.tr/TR,52960/10-milyon-kisi-e-nabiz-kullaniyor.html>, 06.11.2021.
- [12] Internet: Introduction to Web Accessibility, <https://www.w3.org/WAI/fundamentals/accessibility-intro>, 06.11.2021.
- [13] Y. Yesilada, G. Brajnik, M. Vigo, ve S. Harper, "Understanding web accessibility and its drivers", **Proceedings of the International Cross-Disciplinary Conference on Web Accessibility**, Lyon France, 1-9, April 16 - 17, 2012.
- [14] Internet: Making the Web Accessible, <https://www.w3.org/WAI>, 06.11.2021.
- [15] Internet: Web Content Accessibility Guidelines 1.0, <https://www.w3.org/TR/WAI-WEBCONTENT>, 06.11.2021.
- [16] Internet: Web Content Accessibility Guidelines (WCAG) 2.0, <https://www.w3.org/TR/WCAG20>, 06.11.2021.
- [17] Internet: ISO/IEC 40500:2012 Information technology — W3C Web Content Accessibility Guidelines (WCAG) 2.0, <https://www.iso.org/standard/58625.html>, 06.11.2021.
- [18] Internet: W3C Issues Improved Accessibility Guidance for Websites and Applications, <https://www.w3.org/2018/06/pressrelease-wcag21>, 06.11.2021.
- [19] Internet: Europa Web Guide, <https://wikis.ec.europa.eu/display/WEBGUIDE/12.+Accessibility>, 06.11.2021.
- [20] Internet: ISO 9241-20:2008 Ergonomics of human-system interaction — Part 20: Accessibility guidelines for information/communication technology (ICT) equipment and services, <https://www.iso.org/standard/40727.html>, 01.11.2021.
- [21] J. Abascal, M. Arrue, and X. Valencia, "Tools for Web Accessibility Evaluation", **Web Accessibility**, Human–Computer Interaction Series, Editor: Y. Yesilada and S. Harper, Springer, London, 479-503, 2019.
- [22] S. Abou-Zahra, "Web Accessibility Evaluation", **Web Accessibility**, Human–Computer Interaction Series, Editor: Y. Yesilada and S. Harper, Springer, London, 79–106, 2008.
- [23] F. Kamoun ve M. B. Almourad, "Accessibility as an integral factor in e-government web site evaluation: The case of Dubai e-government", *Information Technology & People*, 27(2), 208–228, 2014.
- [24] Internet: Selecting Web Accessibility Evaluation Tools, <https://www.w3.org/WAI/eval/selectingtools>, 07.11.2021.
- [25] G. Brajnik, "Beyond conformance: the role of accessibility evaluation methods", **International Conference on Web Information Systems Engineering**, Auckland, New Zealand, 63-80, September 1-4, 2008.
- [26] H. S. Al-Khalifa, I. Baazeem, ve R. Alamer, "Revisiting the accessibility of Saudi Arabia government websites", *Universal Access in the Information Society*, 16(4), 1027–1039, 2017.
- [27] J. Grantham, E. Grantham, and D. Powers, "Website accessibility: An australian view", **Proceedings of the Thirteenth Australasian User Interface Conference**, Melbourne, Australia, 21–28, 31 January – 3 February 2012.
- [28] Y. Akgül, "Web content accessibility of municipal web sites in Turkey", **European Conference on Digital Government**, Portsmouth, UK, 1–8, 18-19 June, 2015.
- [29] Y. Akgül ve K. Vatansever, "Web accessibility evaluation of government websites for people with disabilities in Turkey", *Journal of Advanced Management Science*, 4(3), 201–210, 2016.
- [30] R. Ismailova, "Web site accessibility, usability and security: a survey of government web sites in Kyrgyz Republic", *Universal Access in the Information Society*, 16(1), 257–264, 2017.
- [31] Y. Akgül, "Accessibility, usability, quality performance, and readability evaluation of university websites of Turkey: a comparative study of state and private universities", *Universal Access in the Information Society*, 20(1), 157–170, 2021.
- [32] Çağiltay, K., & Kubuş, O., "E-devlet siteleri görme engelliler için erişilebilir mi?" **TBD Bilişim Kurultayı**, Ankara, 120–125, January 2006.
- [33] H. Braga, L. S. Pereira, S. B. L. Ferreira, ve D. S. Da Silveira, "Applying the barrier walkthrough method: Going beyond the

- automatic evaluation of accessibility”, *Procedia Computer Science*, 27, 471–480, 2014.
- [34] H. Serefoglu ve T. Henkoglu, “Web accessibility evaluation of university websites in turkey in terms of visually and hearing impaired users”, *Journal of Higher Education and Science*, 9(1), 111-122, 2019.
- [35] Internet: Web Accessibility Evaluation Tools List, <https://www.w3.org/WAI/ER/tools>, 08.11.2021.
- [36] M. Padure ve C. Pribeanu, “Comparing six free accessibility evaluation tools”, *Informatica Economica*, 24(1), 15–25, 2020.
- [37] Internet: Understanding WCAG 2.0, <https://www.w3.org/TR/UNDERSTANDING-WCAG20/Overview.html>, 10.11.2021.
- [38] Internet: Free WCAG 2.0 and Section 508 Web Accessibility Scans, <http://www.cynthiasays.com>, 10.11.2021.
- [39] Internet: What is AT?, <https://www.atia.org/home/at-resources/what-is-at>, 12.11.2021.
- [40] Internet: WCAG 2.1 A/AA Principles and Checkpoints, <https://www.boia.org/wcag-2.1-a/aa-principles-and-checkpoints>, 12.11.2021.

Mask Based Image Encryption Using Chaotic Logistic Map*

Hasan Hüseyin Sayan¹, Ali UZUN^{2*}

¹Faculty of Technology, Gazi University, Ankara, Turkey

²Institute of Informatics, Gazi University, Ankara, Turkey

hsayan@gazi.edu.tr, aunuz8424@gmail.com.tr

Abstract— Recently, image encryption has become one of the most important issues due to the flow of digital images all over the world via the transmission medium. This paper examines the competence of a method for image encryption that uses a chaotic logistic map. In this paper, the proposed algorithm uses the secret key randomly generated for initial parameter of a chaotic logistic map. Then mask image is produced using these parameters. Afterwards, a plain image is encrypted using mask image and Exclusive-OR function. The C# tools used to compare the performance of the proposed algorithm between the original and encrypted images by computing key space analysis, histogram analysis, information entropy analysis and correlation analysis. Hence, the proposed method improves the encryption efficiency compared with classic logistic map.

Keywords—chaotic logistic map, image encryption, chaotic cryptography, image decryption, chaotic cryptanalysis

1. INTRODUCTION

In recent years, people pay more attention to information security with the rapid development of communication technology and the comprehensive use of the Internet and mobile networks. Cryptology has been the fundamental reference source for data protection[1]. Although there are many encryption methods for cryptological applications, chaos based ones have been widely used due to its unique features such as the high sensitivity to initial conditions, randomness and complexity[2]. There are two primary reasons why chaotic system is used in cryptographic applications. Firstly, the output of any chaotic system is sensitive to the system's initial condition. Secondly, the chaotic output can be used to provide pseudo random numbers. According to the encryption key, the system's initial condition is set, then the generated pseudo random numbers are used to cipher the input data in chaos based encryption applications [3]. Chaos based encryption is based on the behavior of changing systems that are sensitive to initial situations, so the results based on the initial situations may change in a way that the behavior cannot be predicted [4].

There are some subtle similarities between chaos and cryptography, so chaos is used to design potential secure and effective schemes in all types of applications in cryptography [5]. Because of these similarities, many researchers and developers are now focusing on the chaotic cryptography. The chaos based image encryption may be

divided into two sections; the first is the chaotic map and the second is the encryption algorithm [6]. The proposed algorithm uses to increase the security chaotic logistic map and Exclusive-OR (EXOR) encryption algorithm. The foundations of chaos based encryption date back to 1989. Since that day, many studies have been proposed to improve its features [7]. Recently, researchers have developed new cryptographic applications for secure image encryption.

2. THEORETICAL INFRASTRUCTURE

Details of chaotic logistic map, image encryption and mask based image encryption are given respectively in this section. Firstly, important points of the chaotic logistic map are mentioned and the formula of logistic map is represented. Secondly, basic information used image encryption is given. Finally, basic structure of mask based image encryption algorithm is given and the recent studies in this field are mentioned.

2.1. Chaotic Logistic Map

Since the last decade, image encryption has improved significantly. Researchers go great lengths to make image encryption algorithms more secure specially by using chaotic systems. Chaotic systems can be analog or discrete. Logistic map is one of the most well known discrete forms

[8]. The logistic map is used in the chaotic cryptosystems because it is quite sensitive to the initial value [9]. The logistic map is calculated by the formula X_{n+1} in Formula 1. The logistic map stands for as:

$$X_{n+1} = \mu \cdot X_n \cdot (1 - X_n) \tag{1}$$

X_n represents the chaotic sequence and its value is between 0 and 1. The parameter μ can be any number in the range from 0 to 4. When this parameter represents between 3.5699 and 4.0, logistic chaotic sequence occurs.

2.2. Image Encryption

An image can be represented by a matrix of M * N pixels, with M rows and N columns. Each pixel, which is a matrix element, consists of 8 bits (0-255 values) of red, green and blue components. The encryption of large data in color images cannot be securely provided with classical encryption methods [10]. The use of chaos encryption methods in applications is increasing gradually because of the secure results of the chaos encryption methods for the secure encryption of this large data.

Traditional methods that encrypt small data such as text are not preferred for image encryption because text data and image are different from each other in terms of sizes [11]. Image sizes are quite larger than text, and encrypting an image requires much more time. Researchers working to develop image encryption algorithms have turned to fast and secure encryption methods based on chaos instead of traditional methods in order to perform the encryption process rapidly. Recently, the widespread use of chaos in the field of encryption and the fact that this situation gives better results than other encryption systems increase the interest in chaos applications. [12]. Chaos based image encryption algorithms use some operation with the help of sequences produced by the chaotic system to encrypt an image because the main aim is to make the plain image turn into a random noise.

2.3. Mask Based Image Encryption

Traditional encryption algorithms such as Data Encryption Standard (DES) or the Advanced Encryption Standard (AES) were originally designed to encrypt text data [13]. Hence image encryption algorithms encrypt block by block or byte by byte following a certain algorithm. Of course, these methods have some advantages and disadvantages. Huge computation complexity, storage capability and small keyspace can be an example of defects of such methods.

3. PROPOSED IMAGE ENCRYPTION SCHEME

In this paper, mask based image encryption using chaotic logistic map encryption algorithm has been developed. The proposed algorithm is like symmetric stream encryptions

such as RC4 encryption algorithm because when searching for image encryption algorithms in the literature, this type of algorithm has been preferred in order to eliminate the deficiencies in image encryption algorithms. This algorithm consists of three main sections. These are the process of creating initial parameters, creating a mask image and encryption. Firstly, a mask image is created using chaotic logistic map. This mask image has same size with the image to be encrypted. For example, a mask image is created in M * N sizes to encrypt an image in M * N sizes. The flow chart of this process is depicted in flowchart. Using key string and encrypted image, when the step 2 and step 3 are repeated, the decrypted image is obtained.

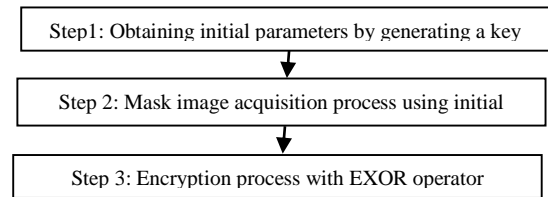


Figure 1. Encryption process steps

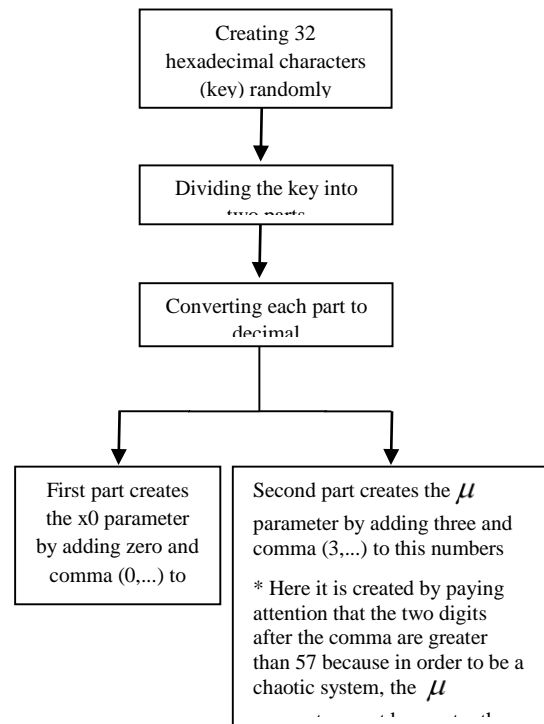


Figure 2. Details of step 1

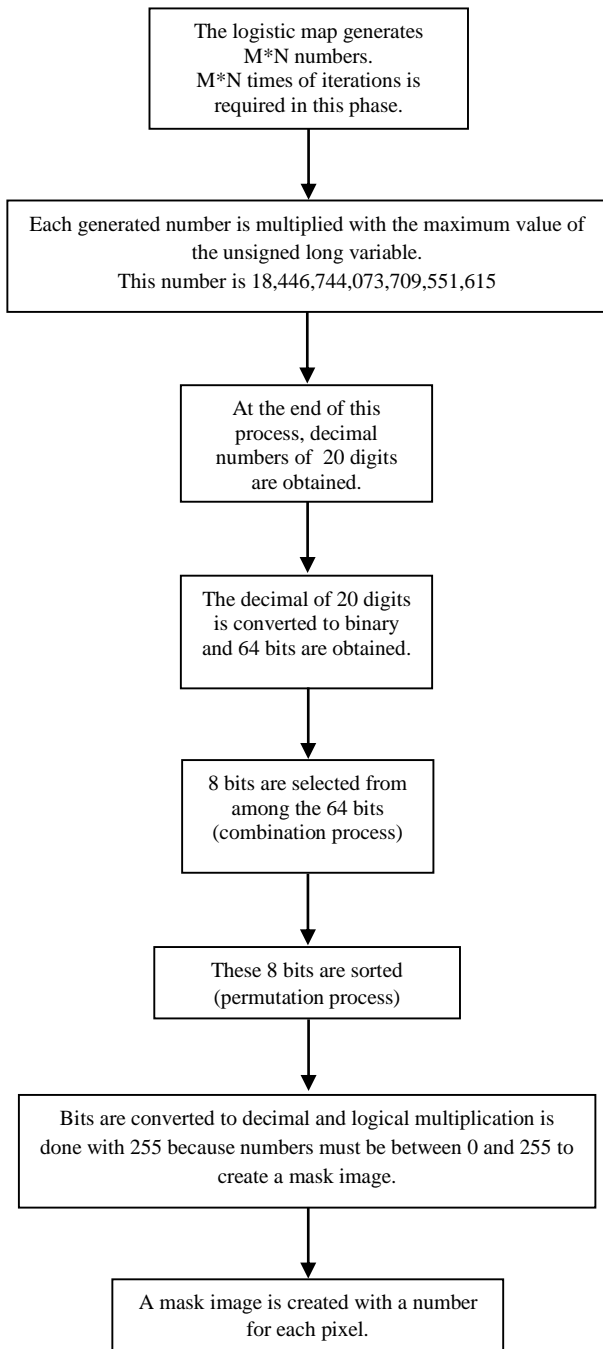


Figure 3. Details of step 2

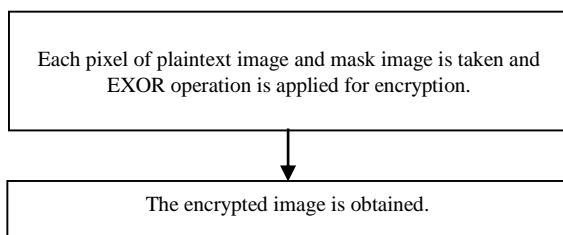
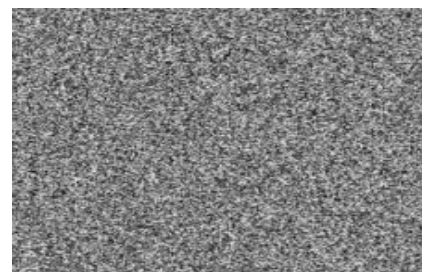


Figure 4. Details of step 3

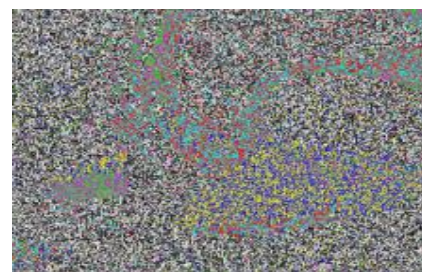
Initial parameters are required to generate numbers using a chaotic logistic map. The initial parameters are used as the encryption key. The key is randomly generated by the system and its length are 32 digit hexadecimal characters (128 bit). In the logistic map formula, the first 16 of these 32-digit hexadecimal characters are used as the X_0 parameter and the last 16 as the μ parameter. The logistic map is a pseudo-random number generator (PRNG). So a logistic map generates numbers using initial parameters. These numbers are combined to create a mask image. The mask image is made safer by increasing the randomness of the numbers with operations. The image is then encrypted with the mask image. EXOR operator is used for this operation. The EXOR operation is applied to each pixel in the same position of the image to be encrypted and the mask image. After this process, an encrypted image is obtained. Figure 5 shows the encryption process for the polar lights image.



(a)



(b)



(c)

Figure 5. a) Plain image of polar lights b) Mask image c) Encrypted image of polar lights

4. SECURITY ANALYSIS

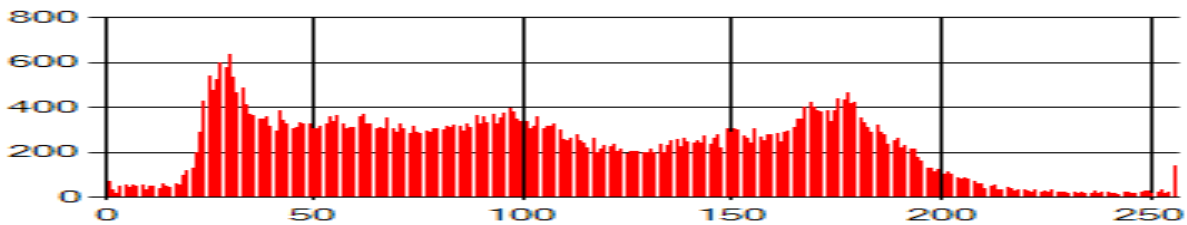
A good encryption system must be robust against to attacks. In this section, different tests are applied to proposed method in order to show the efficiency of proposed model. These tests are key space analysis, histogram analysis, information entropy analysis and correlation analysis. Key space analysis has been performed to ensure resistance against the brute force attacks. To have nearly equal distribution of pixels with the color intensity value, histogram analysis is carried out. The random distribution of pixels is analyzed with information entropy analysis. In order to resistant to statistical attacks, correlation analysis is implemented.

4.1. Key Space Analysis

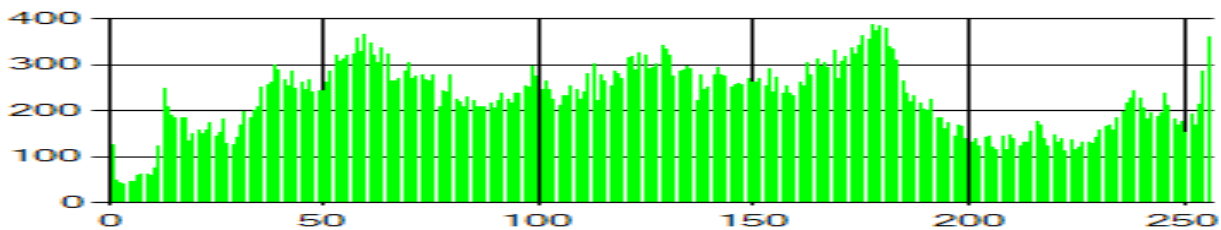
A robust encryption system relies on secret keys. The key space should be enormous enough to make the brute force attack impossible. There are 2128 different combinations of the secret key in our proposed image cipher. This indicates that it will be difficult to decrypt in the event of an attack.

4.2. Histogram Analysis

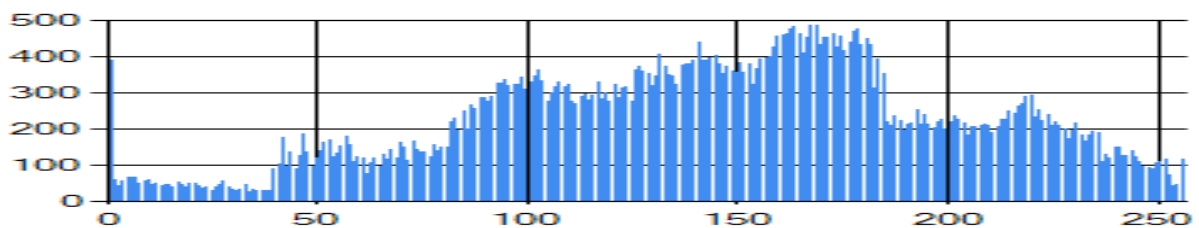
The experiments include encryption and decryption process, histogram analysis of plain-image and cipher-image. With the help of an image histogram, pixels are distributed by graphing the number of pixels at each color intensity. The histogram of the image must have nearly equal distribution of pixels with the color intensity value because ideal encryption requires this.



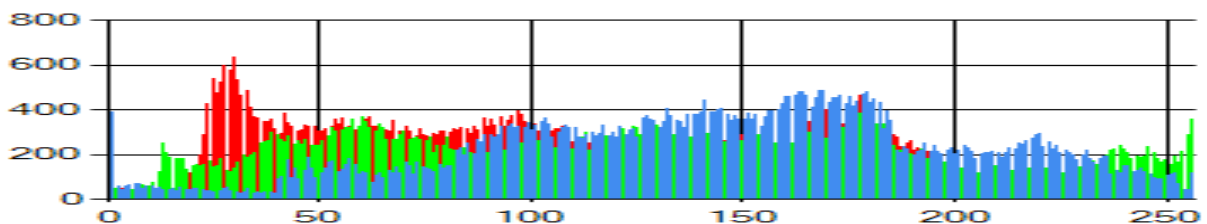
(a)



(b)

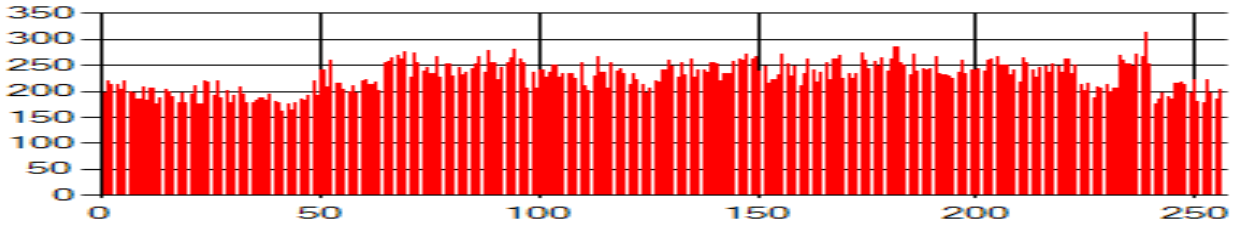


(c)

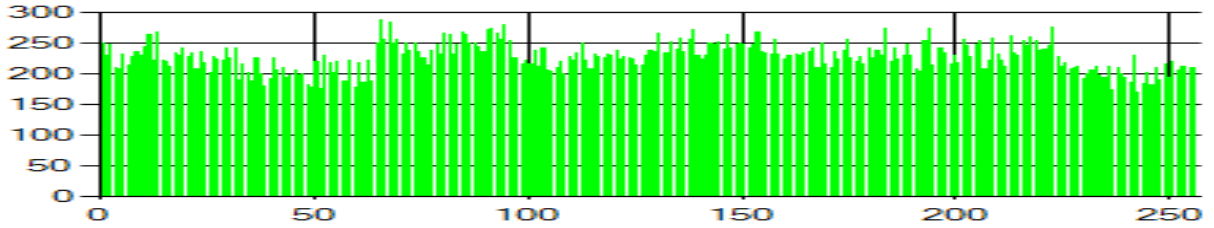


(d)

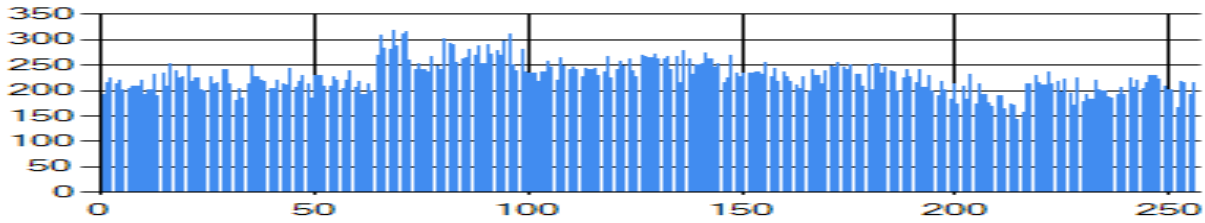
Figure 6. Histogram analysis of plain image of polar lights a) histogram of red channel b) histogram of green channel c) histogram of blue channel d) histogram of all channels



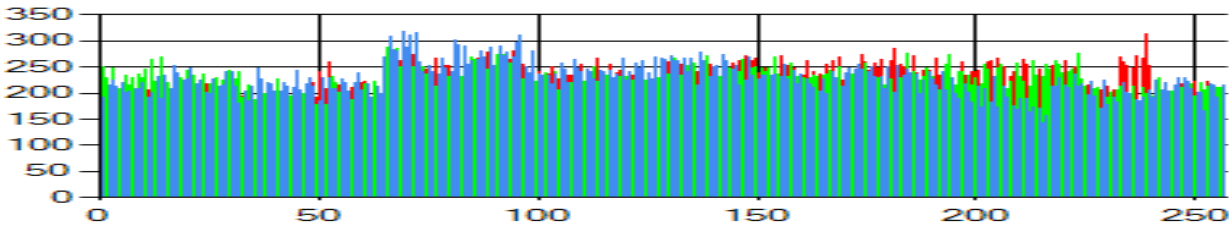
(a)



(b)

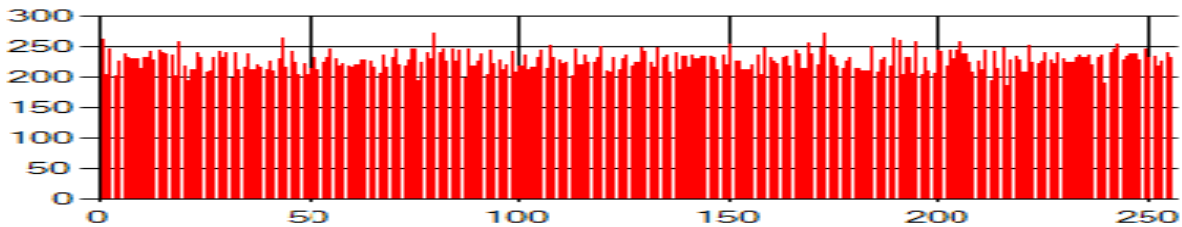


(c)



(d)

Figure 7. Histogram analysis of encrypted image of polar lights with the classic logistic map a) histogram of red channel b) histogram of green channel c) histogram of blue channel d) histogram of all channels



(a)

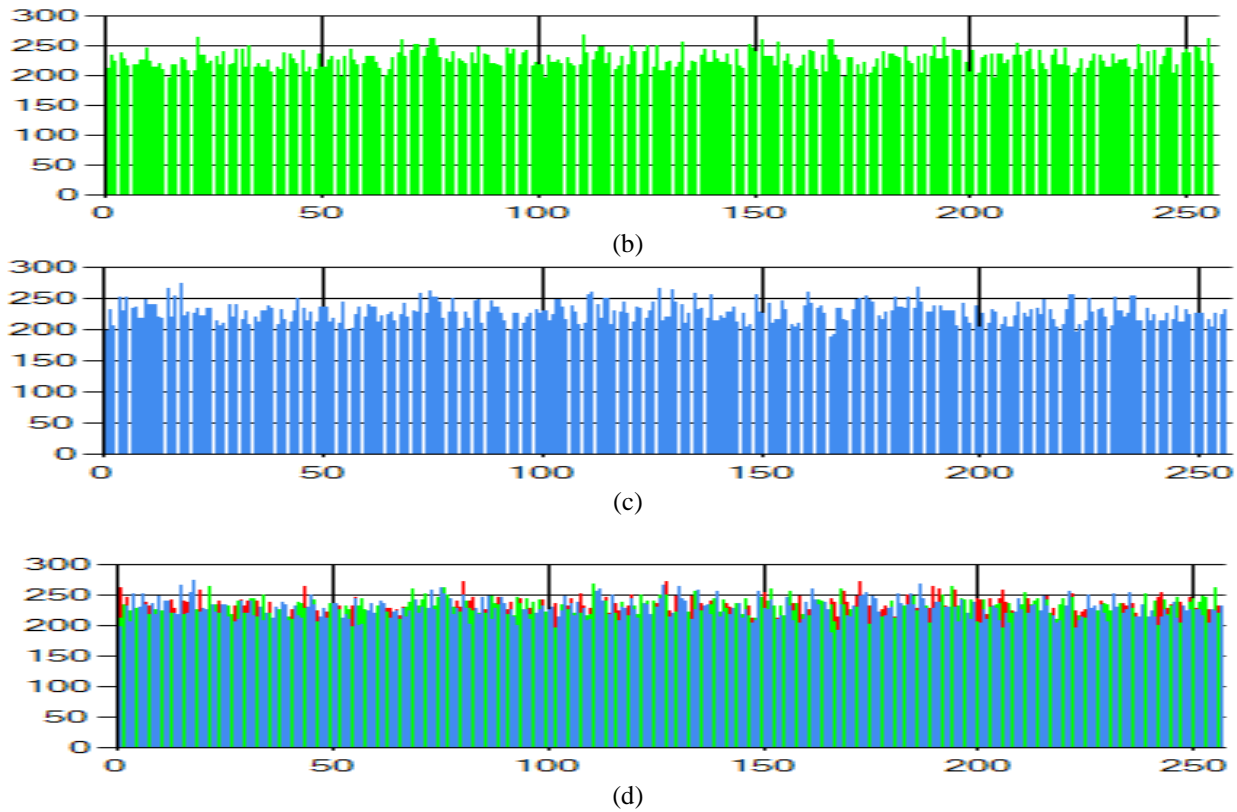


Figure 8. Histogram analysis of encrypted image of polar lights with the proposed logistic map a) histogram of red channel b) histogram of green channel c) histogram of blue channel d) histogram of all channels

Histogram analysis of encrypted image of polar lights with the proposed logistic map is shown in figure 8. As seen in the histogram results the proposed method provides secure encryption against possible attacks with the process of increasing the randomness of numbers during the mask image creation.

4.3. Information Entropies Analysis

Information entropy is a measure of the random distribution of pixels. Information entropy is calculated by the formula H in Formula 2.

$$H = \sum_{i=0}^{2^n-1} p(i) \log \frac{1}{p(i)} \quad (2)$$

Table 1. Information entropies of image of polar lights

	Classic Logistic Map	Proposed Logistic Map
Red Channel	7,98888623286609	7,9967689222797
Green Channel	7,99286580327609	7,99700676448519
Blue Channel	7,98711472072582	7,99670437348192

H being 8 means that the measured information is completely random. Therefore, after the encryption

process, the image in formation entropy value is expected to be about 8 in a robust encryption system. Table 1. shows information entropies of image of polar lights. The value of H being closer to 8 indicates the robust of the proposed logistic map when the encryption is done with the proposed logistic map

4.4. Correlation Analysis

The values between neighboring pixels of the picture are similar. In order to be resistant to statistical attacks to the encrypted image, the connection between neighbor pixels should be minimum. Correlation is calculated by the formula r in Formula 3.

$$r = \frac{\frac{1}{N} \sum_{i=1}^N \left(x_i - \frac{1}{N} \sum_{i=1}^N X_i \right) \left(y_i - \frac{1}{N} \sum_{i=1}^N y_i \right)}{\sqrt{\frac{1}{N} \left(x_i - \frac{1}{N} \sum_{i=1}^N X_i \right)^2 \left(y_i - \frac{1}{N} \sum_{i=1}^N y_i \right)^2}} \quad (3)$$

The r value can take values between -1 and +1. If the value is positive, it indicates that there is a positive relationship between the two images, and a negative relationship indicates that there is a negative relationship. In order for the algorithm to be resistant to attack, it must be just about zero. Table 2. shows correlation image of polar lights.

Table 2. Correlation between plain image of polar lights and encrypted image of polar lights

	Classic Logistic Map	Proposed Logistic Map
Red Channel	-0,140643516020314	-0,00360652334777824
Green Channel	-0,159809353249951	-0,00143491208873442
Blue Channel	-0,150111461352063	-0,00469082168427584

As seen table 2, comparing classic logistic map and proposed logistic map, proposed logistic map is a more robust encryption algorithm than the classic logistic map because correlation value is closer to zero.

5. CONCLUSION

In this paper, a new image encryption and decryption method based on the well-known Chaotic Logistic Map and the EXOR encryption similar to RC4 encryption methods is proposed. Using the complexity of the Chaotic Logistic Map and EXOR encryption, a secure and fast encryption method is proposed. It has been tried to find a way that is fast, robust and completely secure image encryption method with proposed algorithm. The reliability of the algorithm has been proven by statistically examining the numbers produced by the proposed algorithm. In order to prove the reliability of the algorithm statistically, statistical operations such as mean, variance and standard deviation have been applied and successful results have obtained in each of them. Moreover, the proposed method has been analyzed in terms of key space, histogram analysis, information analysis and correlation analysis. The proposed method has shown a nearly equal distribution of pixels with the color intensity value. When the picture was encrypted with the proposed algorithm, the results showed that the encrypted pictures have information entropy of close to 8 and the result of correlation analysis is around 0. When the results of the analysis are compared with the results of encryption with the classical logistic map, it is seen that the proposed method is more secure. The results showed that the algorithm can effectively resist different attacks.

REFERENCES

- [1] Li, C., Lin, D., Feng, B., Lü, J., & Hao, F., "Cryptanalysis of a chaotic image encryption algorithm based on information entropy", *Ieee Access*, 6, 75834-75842, 2018.
- [2] Flores-Vergara, A., Inzunza-González, E., García-Guerrero, E. E., López-Bonilla, O. R., Rodríguez-Orozco, E., Hernández-Ontiveros, J. M., ... & Tlelo-Cuautle, E., "Implementing a chaotic cryptosystem by performing parallel computing on embedded systems with multiprocessors", *Entropy*, 21(3), 268, 2019.
- [3] Rezk, A. A., Madian, A. H., Radwan, A. G., & Soliman, A. M., "Reconfigurable chaotic pseudo random number generator based on FPGA", *AEU-international Journal of Electronics and Communications*, 98, 174-180, 2019.
- [4] Safi, H. W., & Maghari, A. Y., "Image encryption using double chaotic logistic map", **2017 International Conference on Promising Electronic Technologies**, IEEE, 66-70, 2017.
- [5] Li, C., Xie, T., Liu, Q., & Cheng, G., "Cryptanalyzing image encryption using chaotic logistic map", *Nonlinear Dynamics*, 78(2), 1545-1551, 2014.
- [6] Mansouri, A., & Wang, X., "A novel one-dimensional chaotic map generator and its application in a new index representation-based image encryption scheme", *Information Sciences, Information Sciences*, 563, 91-110, 2021.
- [7] Gaata, M. T., & Hantoosh, F. F., "An efficient image encryption technique using chaotic logistic map and rc4 stream cipher", *International Journal of Modern Trends in Engineering and Research*, 3(9), 213-218, 2016
- [8] Ismail, S. M., Said, L. A., Rezk, A. A., Radwan, A. G., Madian, A. H., Abu-ElYazeed, M. F., & Soliman, A. M., "Image encryption based on double-humped and delayed logistic maps for biomedical applications", **2017 6th International Conference on Modern Circuits and Systems Technologies**, IEEE, 1-4, 2017.
- [9] Jerjees, S. A., Esttaifan, B. A., & Ismaeel, T. Z., "Hybrid ciphering method based on chaos logistic map and fingerprint information", *Journal of Engineering Science and Technology*, 15(5), 3013-3024, 2020.
- [10] Zhang, Y., "The unified image encryption algorithm based on chaos and cubic S-Box", *Information Sciences*, 450, 361-377, 2018.
- [11] Chai, X., Gan, Z., & Zhang, M., "A fast chaos-based image encryption scheme with a novel plain image-related swapping block permutation and block diffusion", *Multimedia Tools and Applications*, 76(14), 15561-15585, 2017
- [12] Wu, X., Wang, D., Kurths, J., & Kan, H., "A novel lossless color image encryption scheme using 2D DWT and 6D hyperchaotic system", *Information Sciences*, 349, 137-153, 2016.
- [13] Pareek, N. K., Patidar, V., & Sud, K. K., "Image encryption using chaotic logistic map", *Image and vision computing*, 24(9), 926-934, 2016.

Deep Learning-Based Traffic Accidents Severity Forecast*

Mustafa CAYMAZ¹, İ. Alper DOĞRU^{2*}, Sinan TOKLU³

¹ Electronic Computer Education, Graduate School of Informatics, Gazi University, 06100 Ankara, Turkey

² Department of Computer Engineering, Faculty of Technology, Gazi University, 06100 Ankara, Turkey

³ Department of Computer Engineering, Faculty of Technology, Gazi University, 06100 Ankara, Turkey

mustafa.caymaz@gazi.edu.tr, iadogru@gazi.edu.tr, stoklu@gazi.edu.tr

Abstract— Today, with the increase in the population, the number of vehicles in traffic also increases. In parallel with this increase, an increase is observed in traffic accidents. With the development of technology, many successes have been achieved on deep learning. With the development of GPUs, big data calculations have become easier in deep learning. Estimating the severity of traffic accidents is important for the prevention of traffic accidents. Also, revealing risk factors is crucial for implementing countermeasures. Most of the research on this subject examines traffic accidents statistically. Deep learning design provides a more comprehensive analysis of traffic accident severity. In this study, we propose a DNN framework to estimate the accident severity level. Our proposed model using traffic accident data from Turkey predicts the severity risks of traffic accidents with good accuracy.

Keywords— traffic accidents, severity forecast, deep learning, lstm

1. INTRODUCTION

The concept of the accident is defined as a bad event that occurs at an unpredictable time, resulting in loss of life and property, and an event that can be avoided and protected with previous precautions [1].

Many studies are carried out all over the world and in Turkey to reduce the number of traffic accidents. In these studies, the factors causing the accident are tried to be determined and precautions are taken accordingly. The correct determination of the factors causing the geese will also facilitate the solution of the problem [2].

Studies show that with the increasing population and number of vehicles in our country, the probability of a fatal or injured traffic accident is increasing day by day. According to the statistics published by the Turkish Statistical Institute in 2019; In 2018, 1,229 364 traffic accidents occurred, of which 1,042 832 were materially damaged and 186 532 fatal or injured, 6 675 people lost their lives and 307 071 people were injured in these accidents [3].

Table 1. According to the data obtained from the General Directorate of Security and the Gendarmerie General Command, "Numbers of accidents, dead and injured by years" published by the Turkish Statistical Institute [3]

Table 1. Numbers of accidents, dead and injured by years

Year	Total	M.D.	F.A.	D.T.	N.I.
2012	1296634	1143082	153552	3750	268079
2013	1207354	1046048	161306	3685	274829
2014	1199010	1030498	168512	3524	285059
2015	1313359	1130348	183011	7530	304421
2016	1182491	997363	185128	7300	303812
2017	1202716	1020047	182669	7427	300383
2018	1229364	1042832	186532	6675	307071

(T.A.: Total accidents, M.D.: Number of accidents with material damage, F.A.: Number of fatal and injury accidents, D.T.: Death toll, N.I.: Number of injuries)

In Table 1., it is observed that the number of accidents gradually increases and the number of dead and injured increases according to the years. It is not possible to detect the accidents in advance, but it seems possible to predict the accident risk with the information obtained by deep learning algorithms.

This study is planned to analyze and predict traffic accidents in Turkey with deep learning. In the study, data of approximately 2 million traffic accidents that occurred between 2013 and 2020 were discussed.

2. RELATED WORKS

In [4], was evaluated the application of Bayesian neural network (BNN) models to predict traffic accidents. To achieve this goal, it was implemented a model with

accident data collected from Texas. In the study, three types of models were compared, namely backpropagation neural network (BPNN), BNN, and negative binomial (NB) regression. In general, they were showed that both neural network models outperformed the NB regression model in terms of data prediction.

In [5], the temporal use of traffic accidents caused by traffic accidents in North America validated the choice of analysis and forecasting. Considering the modeling, they defended the phenomenon as a prediction.

In [6], the issue was said that real-time accident prediction usually takes place on highways. However, in this study, they were carried out on urban arterials rather than highways. In the study, they were proposed the long-short-term memory - convolutional neural network (LSTM-CNN) model for real-time collision risk estimation on arteries. They have been shown that this model can learn traffic flow characteristics from various features such as signal timing and weather conditions. They were claimed that they caught long-term dependence with LSTM, while CNN claimed that it extracted features that did not change over time. They stated that LSTM-CNN has a much higher AUC value than other methods.

In [7], They were used eXtreme Gradient Boosting (XGBoost) to detect accident occurrences using a set of real-time data from traffic, network, demographic, land use, and weather characteristics. Data were collected from Chicago metropolitan highways between December 2016 and December 2017 and was included 244 traffic accidents and 6073 non-accident cases. In particular, the speed difference between 5 minutes and 5 minutes before an accident has been found to have a relatively greater effect on the occurrence of accidents.

In [8], applied Backpropagation - Artificial Neural Network (BP-ANN) with Generalized Delta Rule (GDR) learning algorithm to model the factors affecting traffic accidents of elderly male and female drivers. A total of 95,092 accident records were reviewed in the West Midlands of the United Kingdom between 2006 and 2016. It was determined that the purpose of the trip, lighting conditions, pedestrian crossing, complex road geometry, extreme weather, and time severity are the most important factors in elderly driver accidents.

In [9], investigated the factors related to injury severity of pedestrian accidents according to the days of the week. Separate injury severity models have been developed for weekday and weekend accidents. They said that the model they developed (WOA-ANN) got better results than other ANN models.

In [10], the issue said traffic speed estimation is important to achieve a better transportation system in smart cities. They said that accurate traffic speed estimation is influenced by many contextual factors, such as abnormal traffic conditions, traffic incidents, lane closures caused by construction or incidents, and traffic congestion. To

overcome these problems, they proposed a new method called fuzzy optimized long short-term memory (FOLSTM) neural network for long-term traffic speed estimation. The FOLSTM technique is a hybrid method consisting of computational intelligence (CI), machine learning (ML), and metaheuristic techniques that can predict velocity for macroscopic traffic key parameters.

In [11], they used a country-wide traffic and weather dataset collected for the USA over two years. They found that long-term congestion, snow, rain, fog, severe cold, and flow events cause the most significant delay effects on traffic flow.

In [12], investigated the relationship between driver injury severity and driver, vehicle, road, and environmental characteristics. The use of multilayer perceptron (MLP) and fuzzy adaptive resonance theory (ART) neural networks have been investigated. The analysis focused on two-vehicle accidents at signalized intersections. As a result, they showed that rural intersections are more dangerous than urban intersections in terms of driver injury severity. They also showed that female drivers are more likely to experience a serious injury than male drivers. They said that the rate of speed increases the probability of injury severity.

In [13], they analyzed traffic accident data in England at the regional level over the period 2005-2013. They showed in which regions the accidents occurred in England occurred more intensely.

In [14], they provided a detailed discussion to show how different types of NN are used for different aspects of short-term traffic situation forecasting. They propose possible further research directions for additional applications of NN models, which use particularly deep architectures to address the dynamic nature of complex transport networks.

In [15], K-nearest neighbor was used to classifying traumatic brain injuries in traffic accidents.

In [16], They addressed the daily traffic flow forecasting problem based on historical data. They proposed a mapping estimation method integrating contextual mapping and time series estimation based on the group data processing (GMDH) algorithm.

In [17], many researchers say that they offer accident forecasting with a wide variety of forecasting techniques. Most of these methods argue that they are based on statistical studies, but often fail to explain their view of estimation results. He suggested that the prediction of accidents could be improved given poor data quality by using a classifier based on gray relational analysis, which is a similar data-based method. Evaluated the gray relational classifier with other state-of-the-art classifiers, including neural networks, classification and regression trees, k-nearest neighbor, linear discriminant analysis, pure Bayesian classifier, algorithm semi-optimal, and support

vector machines. He presented experimental results to demonstrate the efficiency and robustness of the gray associative classifier algorithm in terms of traffic accident prediction accuracy.

In [18], it was the first study in the literature to offer a modeling framework for analyzing traffic accident frequency based on a set of ensemble machine learning (EML) methods.

3. DEEP LEARNING-BASED TRAFFIC ACCIDENT SEVERITY FORECAST

Vehicle use is increasing day by day. Parallel to this increase, the number of accidents also increases. Many deaths and injuries occur as a result of traffic accidents. At the same time, much property damage occurs. Predicting traffic accidents will reduce these losses. Prediction of traffic accidents is a difficult problem. This problem is difficult to solve with statistical or classical mathematical methods. Future predictions are made by using deep learning methods in the proposed architecture for this problem. In this way, the authorities will be able to determine which factors have a greater effect on the occurrence of accidents at which points.

Traffic accidents that occurred between the years 2013-2020 in Turkey, obtained from the Traffic Department of the General Directorate of Security, were used as a data set [19]. The data set consists of 1,743,268 records. The Driver age feature in the data set was used by dividing it into age groups.

Table 2. Dataset parameters and descriptions

Parameters	Data Type	Notes
KazaYılı	Date-time	Accident Year
Arac_Cinsi	String	Vehicle Type
Sürücü_Cinsiyet	String	Driver Gender
Sürücü_Öğrenim Durumu	String	Driver Education Status
Sürücü_Alkol Sonucu	String	Driver Alcohol Result
Age_band	Integer	Driver Age Band
Sürücü_Emniyet Kemerı	String	Driver's Seat Belt Usage Status
Sürücü_Kusuru1	String	The Fault Of The First Driver In The Accident
Sürücü_Kusuru2	String	The Fault Of The Second Driver In The Accident
Arac_Kusuru1	String	The Fault Of The First Vehicle In The Accident
Arac_Kusuru2	String	The Fault Of The Second Vehicle In The Accident
Sürücü_KAZASO NUCU	String	Result of Traffic Accident: Dead/Injured, Property Damaged

Table 2 shows the parameters and explanations in the dataset.

Table 3. Model Parameters

Parameters	Value / Settings
Train Size	%67
Test Size	%33
Scaling	MinMaxScaler
Epochs	100
Activations	Tanh
Optimizer	Adam
Loss	mean_squared_error
Metrics	AUC, RMSE, Recall, BinaryAccuracy
Batch Size	20
Verbose	1

Table 3 shows parameters in the proposed deep learning model.

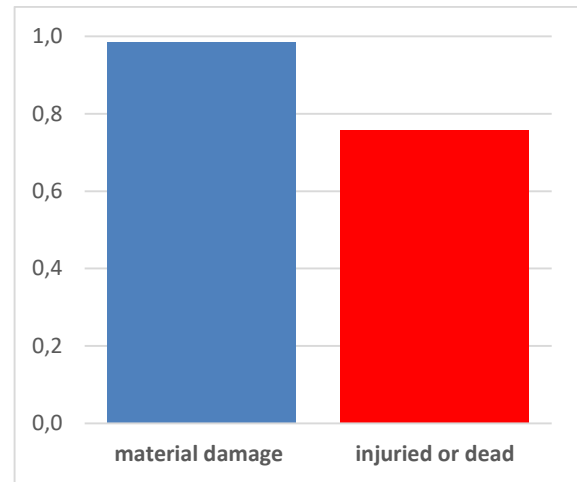


Figure 1. Number of fatal/injury accidents and property damage accidents

986,046 accidents occurred with material damage and 757,222 accidents with death/injury.

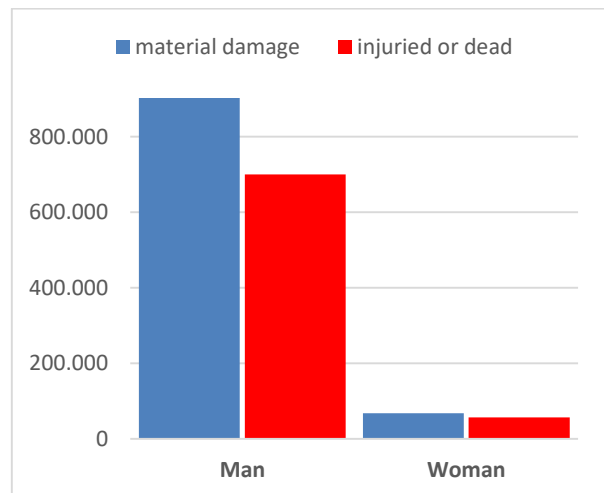


Figure 2. Number of accident types by Sex

Most of the accidents are male drivers. 43% of the accidents involving male drivers were fatal/injured. On the other hand, 45% of the accidents involving women were fatal/injured.

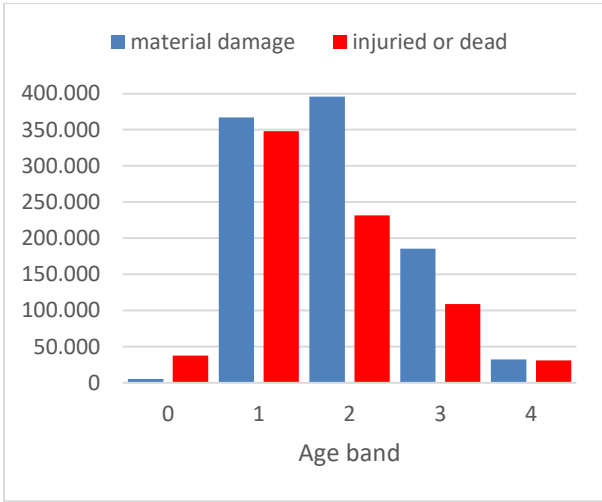


Figure 3. Number of accident types by Age-band

Drivers between the ages of 32-48 were responsible for most of the material damage. Fatal/injury accidents were caused by drivers between the ages of 18-32.

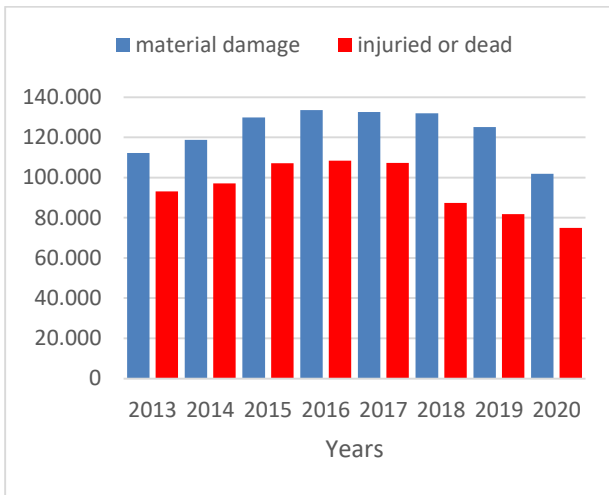


Figure 4. Number of accident types by year

The highest number of accidents occurred in 2016, gradually decreasing until 2021.

Table 4. Performance comparison for different features

Features	Accuracy (%)				
	LSTM	DNN	RF	KNN	SVM
Vehicle type, Sex, Impact section, Alcohol, Education status, Age-band	78.95	76.65	75.83	74.42	70.78
Sex, Education status, Age_band	61.25	59.87	60.17	60.11	60.02
Sex, Impact section, Alcohol, Age-band	73.15	70.63	70.15	67.30	65.28

Sex, Impact section, Age-band	60.18	58.80	57.77	56.13	55.62
-------------------------------	-------	-------	-------	-------	-------

Table 4 shows LSTM as the best method with 78.95% accuracy. The features used are Vehicle type, Sex, Impact section, Alcohol, Education status, Age-band.

The type of vehicle and alcohol use characteristics significantly affect performance.



Figure 5. Loss and Accuracy

4. CONCLUSION

Deep Learning models to prevent traffic accidents in the study of this article is an approach that makes risk assessment of points on a planned route using real accident data. Data of approximately 2 million accidents that occurred in Turkey were used.

Important features were found by making analyzes on the data set. Data that would make learning difficult or give erroneous results were tried to be cleaned. The meaningless accident number was removed from the data set.

The categorical data on the data set were transformed into dummy variables.

In this study, comparisons were made using various deep learning models. These are LSTM, DNN, RF, K-NN and SVM. Accuracy scores were 0.79, 0.77, 0.76, 0.74, and 0.71, respectively. LSTM and DNN gave the best results.

As observed from the literature studies, models in which many learning methods are applied, not just a single learning algorithm, have achieved high success.

For future studies, examining values such as driver's driving characteristics and attention level measurement together with the detection of accident blackspots will yield better results.

More such studies on traffic accidents will be carried out in our country, and success will be achieved in reducing traffic accidents.

REFERENCES

- [1] S. Kıran, S. Şemin, and A. Ergör, “Kazalar ve toplum sağlığı yönünden önemi”, *Sürekli Tıp Eğitimi Dergisi*, 10(2), 50–1, 2001.
- [2] I. Bilim, A., **Konya Şehir İçinde Meydana Gelen Kazaların Analizi ve Kritik Noktaların Belirlenmesi**, Master’s Thesis, Selçuk University, Institute of Science, 2006.
- [3] Internet: Türkiye İstatistik Kurumu, Karayolu Trafik Kaza İstatistikleri, <http://www.tuik.gov.tr/PreHaberBultenleri.do>, 11.11.2021.
- [4] Xie, Y., Lord, D., & Zhang, Y., “Predicting motor vehicle collisions using Bayesian neural network models: An empirical analysis.”, *Accident Analysis & Prevention*, 39(5), 922-933, 2007.
- [5] Rodríguez, J., Jattin, J., & Soracipa, Y., “Probabilistic temporal prediction of the deaths caused by traffic in Colombia. Mortality caused by traffic prediction.” *Accident Analysis & Prevention*, 135, 105332, 2020.
- [6] Li, P., Abdel-Aty, M., & Yuan, J., “Real-time crash risk prediction on arterials based on LSTM-CNN.”, *Accident Analysis & Prevention*, 135, 105371, 2020.
- [7] Parsa, A. B., Movahedi, A., Taghipour, H., Derrible, S., & Mohammadian, A. (Kouros)., “Toward safer highways, application of XGBoost and SHAP for real-time accident detection and feature analysis.”, *Accident Analysis & Prevention*, 136, 105405, 2020.
- [8] Amin, S., “Backpropagation – Artificial Neural Network (BP-ANN): Understanding gender characteristics of older driver accidents in West Midlands of United Kingdom.”, *Safety Science*, 122, 104539, 2020.
- [9] Mokhtarimousavi, S., Anderson, J. C., Azizinamini, A., & Hadi, M., “Factors affecting injury severity in vehicle-pedestrian crashes: A day-of-week analysis using random parameter ordered response models and Artificial Neural Networks.”, *International Journal of Transportation Science and Technology*, 9(2), 100-115, 2020.
- [10] George, S., & Santra, A. K., “An improved long short-term memory networks with Takagi-Sugeno fuzzy for traffic speed prediction considering abnormal traffic situation.”, *Computational Intelligence*, coin.12291, 2020.
- [11] Moosavi, S., Samavatian, M. H., Nandi, A., Parthasarathy, S., & Ramnath, R., “Short and Long-term Pattern Discovery Over Large-Scale Geo-Spatiotemporal Data.”, **Proceedings of the 25th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining**, 2905-2913, 2019
- [12] Abdelwahab, H. T., & Abdel-Aty, M. A., “Development of Artificial Neural Network Models to Predict Driver Injury Severity in Traffic Accidents at Signalized Intersections.”, *Transportation Research Record: Journal of the Transportation Research Board*, 1746(1), 6-13, 2001.
- [13] Boulieri, A., Liverani, S., de Hoogh, K., & Blangiardo, M., “A space-time multivariate Bayesian model to analyse road traffic accidents by severity.”, *Journal of the Royal Statistical Society: Series A (Statistics in Society)*, 180(1), 119-139, 2017.
- [14] Do, L. N. N., Taherifar, N., & Vu, H. L. (). “Survey of neural network-based models for short-term traffic state prediction.”, *WIREs Data Mining and Knowledge Discovery*, 9(1), 2019.
- [15] Liu, Q., Liu, J., Wu, X., Han, X., & Guan, F., “Evaluation and prediction of diffuse axonal injury based on optimization strategy in vehicle collision accidents.”, *Structural and Multidisciplinary Optimization*, 60(4), 1491-1508, 2019.
- [16] Song, X., Li, W., Ma, D., Wang, D., Qu, L., & Wang, Y., “A Match-Then-Predict Method for Daily Traffic Flow Forecasting Based on Group Method of Data Handling: A match-then-predict method for daily traffic flow forecasting based on group method of data handling.”, *Computer-Aided Civil and Infrastructure Engineering*, 33(11), 982-998, 2018.
- [17] Twala, B., “Extracting grey relational systems from incomplete road traffic accidents data: The case of Gauteng Province in South Africa.” *Expert Systems*, 31(3), 220-231, 2014.
- [18] Zhang, X., Waller, S. T., & Jiang, P., “An ensemble machine learning-based modeling framework for analysis of traffic crash frequency.”, *Computer-Aided Civil and Infrastructure Engineering*, 35(3), 258-276, 2020
- [19] Internet: Emniyet Müdürlüğü Trafik Başkanlığı, <http://www.trafik.gov.tr/veri-talebi>, 11.11.2021.

Systematic Literature Review of Detecting Topics and Communities in Social Networks*

Ömer Ayberk ŞENCAN^{1*}, İsmail ATACAK², İbrahim Alper DOĞRU³

¹Department of Computer Engineering, Gazi University, Ankara 06500, Turkey

²Department of Computer Engineering, Gazi University, Ankara 06500, Turkey

³Department of Computer Engineering, Gazi University, Ankara 06500, Turkey
omerayberk.sencan1@edu.tr, iatacak@gazi.edu.tr, iadogru@gazi.edu.tr

Abstract— The exponential increase in the number of active social media users has made social media one of the main sources for extracting information online. This has raised the importance of social media. As a result, it has been inevitable for researchers to work on concepts covering different topics and concepts on social media. This paper consists of a systematic and in-depth literature review on the concept of topic detection and community detection concepts on social media platforms to provide a comprehensive view of these areas. The results of our literature reviews, most of which focused on machine learning-based methods as the model to be used, have shown that a unique model cannot be proposed for the community and topic detection in terms of performance and that the appropriate model can only be selected or created by considering some specific characteristics of the given problem.

Keywords— social network, emerging topics, community detection, twitter

INTRODUCTION

The rapid increase in technology has led to increased interest in social networks and the frequency of use of these networks. Social networks that allow people to share content such as news, thoughts, and photos from anywhere are becoming more and more frequent. The number of monthly active social media users worldwide is expected to reach 3.43 billion by 2023, about a third of the world's entire population [1]. Twitter among these networks has become one of the most widely used social media apps due to its unique features, including a simple interface and a limit of 140 characters per post [2]. Today, it is known that almost every new piece of information is published first on social media and then on traditional media platforms. This makes social media platforms such as Twitter, a source of information worth examining. In line with this interest, multiple fields of study on Twitter have emerged. Some of these fields are, detection of topics and information, analysis of social media users, community identification, and detection. Apart from these areas, there are different workspaces as well. The main contribution of this study is the examination of studies in the fields of community detection and identification and topic/event detection and the presentation of an overview of these studies. The analysis that led to the findings is primarily based on community detection. This analysis is based on the criteria by which the groups in social networks are evaluated as a community and the characteristics of these groups. In addition to that, studies in the field of Topic Detection are also reviewed in this paper.

The rest of the paper is organized as follows: Section 2 consists of basic definitions of the research and gives information about the background of this study. The methodology we follow is given in Section 3. The results obtained in this study are presented in Section 4 and Section 5 summarizes and concludes the paper.

BACKGROUND

To understand the studies carried out in the field of community detection and topic detection, it is of great importance to first define the terms community and Topic as they are used in these studies. In this section, it is aimed to present these two terms and to give a broad idea about the usage of these terms.

Community Definition

The issue of community detection has gained an important place on social media platforms. This is mainly because a number of useful information about these users can be extracted and evaluated through the detection of users who has certain features in common. community detection aims to identify compatible clusters or groups in real-world graphs, such as social media networks [3]. It also aims to identify the modules of the graph as well [4]. Identifying user communities with similar interests, determining which advertisement to show to which users, and identifying the spam-posting communities are some of the results that are intended to be obtained [5]. These results are used for many different purposes such as blocking spam-posting accounts

[6], showing the related advertisement to the users with similar interests [7], and so on.

Emerging Topics Definition

The topic is considered simultaneously seen in word-groups [8]. The detection of topics is defined as the process of obtaining and summarizing trending topics in a form that will contain useful information. Therefore, the emerging topic detection concept can be defined as the detection of emerging issues on social networks. Also known as hot topic detection, aims to group thematically related documents from a temporal set of documents into an unknown number of topics, and then find a series of topics that are frequently seen over a period of time [9].

METHODOLOGY

In social media analysis, topic detection and community detection are known as some of the most common research areas. Thus, a large number of studies can be expected to exist which is related to the objective of this study. A systematic literature review (SLR) was chosen to be able to identify suitable articles from a large number of publications. Some useful features such as the method being accurate, robust, and transparent are taken into consideration when making this selection [10].

Planning the Review

During the study of a new field of information, researchers usually conduct a bibliographic examination to determine publications on a particular topic. However, such reviews do not use a systematic approach and do not offer any system to prevent factors such as deviation, bias, misapprehension, tendentiousness that may occur during the selection of publications to be analyzed [11]. While conducting the identification of publications, another way of analyzing primary studies is to use the systematic examination method [12], because it allows gaining clear and unbiased information about the research topic [13]. This systematic review study is conducted to summarize information about approaches involving topic and community detection through social networks and therefore to investigate the advantages and limitations of these studies. The process used in this systematic review is given below.

Table 2. Identification of the purpose of the study using the GQM Method [14]

Analysis	Studies on topic and community detection
Purpose	Characterizing and summarizing the studies
Scope	Originality and performance of the proposed model
Viewpoint	Researcher
Context	Primary studies involving topic and community detection

The following sections describe the basic information about the conduct of this systematic review. A more

detailed description can be found in the study of Barcelos & Travassos [15].

Table 3. Systematic review protocol master information based on the template of Biolchini et. al. [16]

Research Question Target	Review and evaluation of the studies in the field of community and topic detection.	
Research Questions (RQ)	RQ1	Is community and topic detection possible in social networks?
	RQ2	What are the methods/models used in this context?
	RQ3	Is there any specific model or models that are more successful than the others?
	RQ4	If so, which models are they?
Context	(1) The studies should be published in a journal with a Q1 Degree.	
	(2) The studies should be published after 2017.	
	(3) The studies should be primary.	
	(4) The studies should be tested with a dataset.	
	(5) The F-Score values should be specified in the article.	
	(6) The studies should have the SCI/SCIE/SSCI index.	
	(7) The studies should be in the English language.	
	(8) The data that used in the studies should be fetched from Twitter.	
Keywords	Community detection, emerging topic detection, topic detection, social networks, twitter, deep learning, machine learning	
Resource Selection Criteria	(1) It should have a search engine suitable for finding the articles.	
	(2) It should have official recognition in the academic field.	
	(3) It should consist of Q1, Q2 or at least Q3 indexed articles	
Resource List	IEEE Portal (IEEE Explore) [17], Science Direct Portal [18]	
Inclusion, Exclusion Criteria	The articles should be accessible and include the topics mentioned in the Question tab.	

3.2. Review Phase

There are two steps in the execution of this systematic review. The first of these steps consists of scanning scientific publications related to the above-mentioned topics. This scan is carried out by using search engines to find the related papers. This search was done with custom search strings specified using specific keywords defined in the protocol. The relevant information is provided in

Table 3.

After the identification process, the summary and introduction part of each publication were analyzed and it was determined whether they were selected for a more comprehensive analysis according to the criteria for inclusion or exclusion.

This systematic review is carried out in November 2021. In this review, 14 scientific studies were defined in the field of topic detection and only 3 were deemed to meet the relevant criteria after the analyses. The techniques used in these studies are as follows:

- I. Evolutionary context-aware sequential model (ECSM) [19]
- II. Enhanced heartbeat graph (EHG) [20]
- III. Dynamic topic model (DTM) [21]

The information about the main characteristics of these studies alongside the algorithms used has been given in Table 4.

Table 4. The Topic detection models and their main characteristics

Name of the study	Proposed method	Main characteristics
Evolutionary context-aware sequential model (ECSM)	Long short term memory (LSTM)	Both temporal dynamics and local semantic sequential dependencies are captured through two new context-aware layers integrated.
Enhanced heartbeat graph (EHG)	-	Introduced a new graphically based approach called EHG, which is effective in detecting events through Twitter data.
Dynamic topic model (DTM)	Latent Dirichlet allocation model	A geo-topic monitoring system has been proposed to identify, monitor,

		and relativize important local events in urban environments.
--	--	--

In this review, a total of 15 articles were found in the field of community detection. However, only 4 of these studies were accepted because they met the specified criteria, and the remaining 11 were excluded. The techniques used in these studies are as follows:

- I. Content-based approach to community detection [22]
- II. An incremental algorithm to detect communities in dynamic evolving social networks [23]
- III. A metaheuristic approach based on the Iterated greedy methodology for detecting communities in large social networks [24]
- IV. A community detection method that considers users' topical interests and their temporal evolution in tandem by learning neural users' representations [25]

The information about the main characteristics of these studies alongside the algorithms used has been given in Table 5.

Table 5. The community detection models and their main characteristics

Name of the study	Proposed method	Main characteristics
Content-based approach	-	Content analysis has been performed. Texts were examined using syntax and semantic properties.
Incremental algorithm	Dynamic Bayesian Non-Negative Matrix Factorization (DBNMF)	Used Bayesian matrixes to examine the Twitter data and detect communities.
Iterated Greedy methodology	Iterated Greedy search	Used Iterated Greedy Search method to detect communities.
Topical Interest and temporal evolution	-	Community detection was made by evaluating the subject-based interests and temporal evolutions of the users as a partner.

EVALUATION OF THE RESULTS

In this section, the results of the selected articles are evaluated in-depth and in-detail.

The Results of Topic Detection Studies

In the first of three articles selected on this area according to the selection criteria, the researchers developed an evolutionary context-sensitive sequential model to make topic detection. In the process of developing this model, LSTM, one of the machine learning models used to create context-awareness. This context-sensitive topic layer has helped capture semantic consistency sensitive to global context-awareness. The sequential layer, on the other hand, has been used to learn local dynamics and semantic dependencies throughout the dynamic evolutionary process. Experimental results from actual datasets show that the method achieves better results than state-of-the-art approaches. Accordingly, this study achieved better results based on PMI Score than DTM [26] and RCRP [27], previous studies in the field.

The second article, Enhanced heartbeat graph, proposes a graphical approach to detect topics. Researchers tested this method on three different sets of real-world data. According to the topic-recall of the proposed model in this study, it was seen that it achieved better results than previous studies in this field, such as SFPM, Doc-P, BNgram, and Exemplar. The comparison has been given in Table 5.

Table 5. Comparison of EHG with other models on a real-world dataset

Method	Top-K				
	20	40	60	80	100
LDA	.000	.130	.180	.280	.227
Doc-P	.310	.460	.500	.540	.680
Gfeat-P	.085	.227	.280	.280	.280
SFPM	.270	.325	.325	.325	.325
BNgram	.540	.540	.540	.540	.540
SVD+Kmean	.400	.547	.626	.666	.666
SNMF-Orig	.100	.277	.277	.320	.363
SNMF-KL	.183	.318	.366	.453	.410
Exemplar	.538	.586	.600	.638	.638
EHG	.466	.628	.674	.699	.711

The last paper examined in the topic detection area, dynamic topic model, focused on semantic similarity and geographical proximity to examine important local events. In the research, a data-driven system was proposed to detect and monitor the semantic, spatial, and temporal dynamics of these geographic topics, especially those designed for geotagged tweets. In order to perform these operations, a system called the dynamic topic model (DTM) was developed. According to the results of the research, the proposed system revealed that it is able to determine and monitor the semantic dynamics of emergencies during disasters accurately and effectively.

The Results of Community Detection Studies

The first study selected in the area of community detection, content-based approach, analyzes the community perception problem from the perspective of content

analysis. It is argued that the content produced in social interaction is a very distinctive feature of a community, so it can be used effectively for community perception. The problem was analyzed from a textual perspective using only syntax and semantic properties, including the top-level properties expressed as topics. In this context, it was tried to find three groups present in the Italian Parliament and followers of these groups. The values of the number of users and number of tweets found with the proposed model in the study are as follows:

Table 6. The results obtained from the study called the community perception problem from the perspective of content analysis.

	Number of Users	Number of tweets
Right parties	19	2174
Cinque Stelle	20	2295
Democratic Party	25	3452
Right parties' followers	126	4948
Cinque Stelle followers	289	16,145
Democratic Party followers	306	17,201

The other study evaluated, incremental algorithm, first proposed four types of incremental elements, and then worked on different update strategies. Finally, by combining these two concepts, a new algorithm has been proposed for community detection in dynamic developing networks. The difference of this study from similar studies in the field is that, unlike many studies, social networks assume that they are dynamic structures, not static ones. To achieve this, a system has been developed to update communities after each time period, taking into account the results of the last time period.

Another study examined, iterated greedy methodology, developed a super-intuitive algorithm based on the iterated greedy methodology to detect communities in large social networks. Accordingly, a new iterated greedy algorithm is proposed for the detection of communities in the ECO (ant colony optimization) networks. This algorithm starts from a first solution created by a heuristic procedure. It then develops through two well-parsed stages, called demolition and reconstruction, iteratively.

The last study reviewed in this paper, topical interest and temporal evolution, uses the multi-model attribute learning (embedding) method for community detection. To achieve this, a three-stage model has been developed. These stages are as follows:

- I. Temporal content-based and topological user representation learning
 - II. Adding user embeddings
 - III. User community detection
- The results obtained in this study are given in Figure 1.

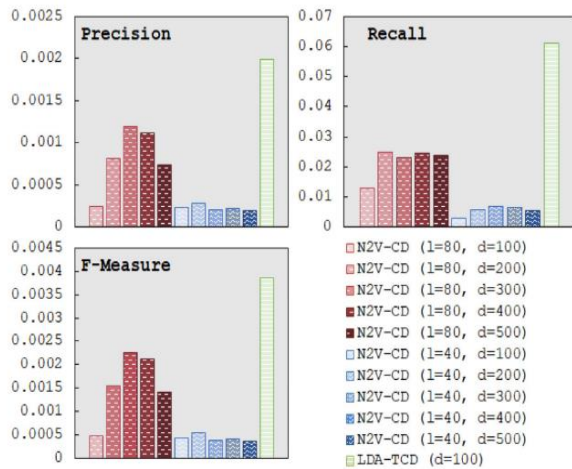


Figure 1. The performance metrics of topical interest and temporal evolution model

CONCLUSION

In this study, a systematic review was made on the studies of topic detection and community detection concepts on social media platforms, and the place of the relevant concepts on social media, the applied methods-methodologies and their performances on the concepts were examined. The review process was managed on the answers obtained in response to the questions about the topic detection and community detection concepts presented in Table 7. The results of this review shows that it is possible to make community detection and topic detection on social networks using different machine learning-based methods.

Table 7. Questions and the answers about the topic detection and community detection concepts

Research Question No.	Research Questions	Answers
Q1	Is community and topic detection possible in social networks?	Yes it is possible to detect communities and topics
Q2	What are the methods/models used in this context?	LSTM, EHG, DTM, DBNMF, content-based approach, iterated greedy search, topic interest, and temporal Evolution are some of the models used in this context.
Q3	Is there any specific model or	No there are no specific

	models that are more successful than the others?	model/models that are more successful than others.
Q4	If so, which models are they?	-

In this context, it was shown that the researchers succeeded with different methods on the specific topics they want to investigate. With reference to that, it is clear that there is no single method more successful than others in the area of topic detection or community detection. In addition to the results given above, it was also observed that the researchers can either build their own model or adjust another model for the specific problem but the model has to be built for the problem itself. Because there is no single model/method which can be applied successfully to every problem in the given area, the researchers can also choose the methods and the model based on the requirements and the specifications of the given problem.

REFERENCES

- [1]Internet: Social Media - Statistics & Facts, <https://www.statista.com/topics/1164/social-networks/#dossierKeyfigures>, 24.11.2021.
- [2] H.-J. Choi & C. H. Park, "Emerging topic detection in twitter stream based on high utility pattern mining", Expert Systems With Applications, 115, 27-36, 2018.
- [3] W. Wu, J. Zhao, C. Zhang, F. Meng, Z. Zhang, Y. Zhang & Q. Sun, "Improving performance of tensor-based context-aware recommenders using bias tensor factorization with context feature auto-encoding", Knowledge Based Systems, 128, 71-77, 2017.
- [4] S. Fortunato, "Community detection in graphs", Physics Reports, 3-5, 75-174, 2010.
- [5] X. Yao, Y. Zou, Z. Chen, M. Zhao & Q. Liu, "Topic-Based Rank Search with Varifiable Social Data Outsourcing", Journal of Parallel and Distributed Computing, 100, 1-12, 2019.
- [6] H. Byun, S. Jeong & C.-K. Kim, "SC-Com: Spotting Collusive Community in Opinion Spam Detection", Information Processing & Management, 58(4), 2021.
- [7] J. W. Kim, K. M. Lee, M. J. Shaw, H.-L. Change, M. Nelson & R. M. Easley, "A Preference Scoring Technique for Personalized Advertisements on Internet Storefronts", Mathematical and Computer Modelling, 44(1-2), 3-15, 2006.
- [8] H. Liu, Y. Ge, Q. Zheng, R. Lin & H. Li, "Detecting global and local topics via mining twitter data", Neurocomputing, 273, 120-132, 2017.
- [9] W. Ai, K. Li & K. Li, "An effective hot topic detection method for microblog on spark", Applied Soft Computing, 70, 1010-1023, 2017.
- [10] M. K. Linnenluecke, M. Marrone & A. K. Singh, "Conducting systematic literature reviews and bibliometric analyses", Australian Journal of Management, 45(2), 175-194, 2020.
- [11] B. Kitchenham, Procedures for Performing Systematic Reviews, Keele University, 2004.
- [12] B. Kitchenham & S. Charters, Guidelines for performing systematic

- literature reviews in software engineering, Durham University, 2007.
- [13] H. G. Gürbüz & B. Tekinerdoğan, "Model-based testing for software safety: a systematic mapping study", *Software Quality Journal*, 26(4), 1327-1372, 2018.
- [14] V. Basili, G. Caldiera & H. Rombach, "Goal Question Metrics Paradigm", *Encyclopedia of Software Engineering*, 528-532, 1994.
- [15] R. Barcelos & G. Travassos, "Evaluation Approaches for Software Architectural Documents: a Systematic Review", *Memorias de la IX Conferenci al Iberoamericana de Software Engineering (CIbSE)*, La Plata, Argentina, 433-446, 24-28 April, 2004
- [16] J. Biolchini, P. G. Mian, A. C. C. Natali & G. H. Travassos, *Systematic Review in Software Engineering*, Systems Engineering and Computer Science Department / UFRJ, Rio de Janeiro, 2005.
- [17] Internet: IEEE Explore, <http://ieeexplore.ieee.org/Xplore/guesthome.jsp> , 25 November 2021.
- [18] Internet: Science Direct, <https://www.sciencedirect.com/>, 25 November 2021.
- [19] Z. Lu, H. Tan & W. Li, "An Evolutionary Context-aware Sequential Model for topic evolution of text stream", *Information Sciences*, 473, 166-177, 2018.
- [20] Z. Saeed, R. A. Abbasi, I. Razzak, O. Maqbool & A. Sadaf, "Enhanced Heartbeat Graph for Emerging Event Detection on Twitter Using Time Series Networks", *Expert Systems with Applications*, 136, 115-132, 2019.
- [21] F. Yao & Y. Wang, "Tracking Urban Geo-Topics Based on Dynamic Topic Model", *Computers, Environment and Urban Systems*, 79, 2019.
- [22] G. Ramponi, M. Brambilla, S. Ceri, F. Daniel & M. D. Giovanni, "Content-based characterization of online social communities", *Information Processing and Management*, 57(6), 2019.
- [23] Z. Zhao, C. Li, X. Zhang, F. Chiclana & E. H. Viedma, "An incremental method to detect communities in dynamic evolving social networks", *Knowledge-Based Systems*, 163, 404-415, 2018.
- [24] J. Sánchez-Oro & A. Duarte, "Iterated Greedy algorithm for performing community detection in social networks", *Future Generation Computer Systems*, 88, 785-791, 2018.
- [25] H. Fani, E. Jiang, E. Bagheri, F. Al-Obeidat, W. Du & M. Kargar, "User community detection via embedding of social network structure and temporal content", *Information Processing and Management*, 57(2), 2019.
- [26] D. Blei & J. Lafferty, "Dynamic topic models", *Proceedings of the 23rd International Conference on Machine Learning*, New York, 2006.
- [27] A. Ahmed & E. Xing, "Dynamic non-parametric mixture models and the recurrent chinese restaurant process: with applications to evolutionary clustering", *Proceedings of the SIAM International Conference on Data Mining*, Atlanta, 2008.



95.yıl

