Proceedings of International Conference on

AI and Big Data in Engineering Applications

Edited by

Can BALKAYA

Proceedings of International Conference on

AI and Big Data in Engineering Applications

Edited by

Can BALKAYA

Department of Civil Engineering
Engineering Faculty
Istanbul Aydın University

Organized by
Istanbul Aydın University
14-15 June 2021 Istanbul, Turkey

İstanbul, 2021

Istanbul Aydın University Publications

Proceedings of International Conference on

AI and Big Data in Engineering Applications

Editorial Board Chairman: Assoc.Prof. Mustafa AYDIN

Editor: Dr. Can BALKAYA

Cover and Page Design: Istanbul Aydın University Visual Design Unit

Publication Year: 2021

Edition No: 1

Publication Place: C&B Basımevi
Litros Yolu 2.Matbaacılar Sitesi A Blok Zemin Kat No: ZA 1 34020, Istanbul, Turkey
Phone: +90 212 612 6522

E-ISBN: 978-625-7783-826

PREFACE

ICAIBDEA is an acronym for International Conference on Artificial Intelligence (AI) and Big Data in Engineering Applications. Today's artificial intelligence and big data is commonly used in every field in our lives as engineering applications including smart homes and smart cars. Artificial intelligence contributes to the system to be more planned, systematic and fast by adding different capabilities to solve many complex problems and decision support systems in engineering problems. Artificial intelligence decisions on collection and control of big data are widely used for practical engineering applications. Then the systems become smart and expert.

Today's companies have realized the situation and thought that artificial intelligence technologies can add value for their companies. With the increase in the use of artificial intelligence in daily life, many sectors tend to this direction as well as engineering applications. Big Data allows storage and increasing data space. This data is easy to access and safe because it is stored in cloud technology. Moreover, this data can be classified by machine learning. Advanced artificial intelligence algorithms are used to provide fast and effective solutions in line with the data received from real-time is used condition assessment, risk assessment and emergency and disaster managements.

Disaster management, industrial production, engineering design and construction, maintenance, sustainability, control and manage data in real time by using sensors, monitoring information, energy distribution, traffic congestion, environmental solutions are efficiently developed and implemented integrated solutions in AI and Big Data. Such as traffic and transportation data, infrastructure data, early warning systems to improve the management and efficiency of the urban environment and traffic safety.

Proceedings and presentations include case studies and practical applications, innovative techniques in engineering applications. On the other hand, the principal mission of ICAIBDEA this international conference is bridging the gap between universities and sectors for artificial intelligence applications in engineering fields, encouraging interdisciplinary research applications and bringing multidisciplinary researchers together.

Conference topics include engineering applications in industrial product design, industrial engineering, engineering management/disaster management, energy systems, smart systems, smart cities, transportation, cyber security, waste management, agriculture, mechatronics, robotics, control systems/monitoring/GIS, civil engineering/structures, infrastructure projects, information systems/expert systems, machine learning techniques, data collection systems, biosensors, computer aided design and applications, optimum designs, database/data mining, and other AI and Big Data analytics engineering applications. ICAIBDEA2021 is organized by Istanbul Aydın University.

We hope that the national and international academic community and industrial practitioners will continue for networking and future collaborations between the conference participants. We are also hoping to meet in healthy days and face-to-face meetings in the future.

Finally, thanks to invited key-note speaker, Distinguished Professor David Arditi from Illinois Institute of Technology for his valuable speech on AI and construction management applications.

Prof.Dr. Can Balkaya
Conference Chair and Editor

ORGANIZATION

Conference Chair and Editor

Prof.Dr. Can BALKAYA	Istanbul Aydın University

International Program and Technical Evaluation Committee

Prof.Dr Can BALKAYA	(Turkey) Istanbul Aydın University
Assoc.Prof.Dr. Onur Behzat TOKDEMİR	(TurkeyMiddle East Technical University
Prof.Dr. Ali OKATAN	(Turkey) Istanbul Aydın University
Prof.Dr. Mehmet GÜLGÜN	(Turkey) Sabancı University
Prof.Dr. Erol TUTUMLUER	(USA) University of Illinois at Urbana-Champaign, IL, USA
Prof.Dr. İlhami ÇOLAK	(Turkey) Nişantaşı University
Asst.Prof.Dr. Serkan AYVAZ	(Turkey) Bahçeşehir University
Prof.Dr. Oğuzhan HASANÇEBİ	(Turkey) Middle East Technical University
Assoc.Prof.Dr. Gökçe SARGUT	(USA) Governors State University, University Park, IL, USA
Dr. Ahmet ÇITIPITIOĞLU, PE	(Turkey) TAV Construction
Dr. Erol KALKAN, PE	(USA) QuakeLogic Inc., Roseville, CA, USA
Prof.Dr. Oktay ÖZCAN	(Turkey) Istanbul Aydın University
Prof.Dr. Ömer Faruk ERKENDİRCİ	(Turkey) Istanbul Aydın University
Prof.Dr. Zeynep Dilek HEPERKAN	(Turkey) Istanbul Aydın University

Supporting Team

Abdullah NİĞDELİOĞLU	(Turkey) Istanbul Aydın University
Didem ODABAŞI CİNGİ	(Turkey) Istanbul Aydın University
Birkan DAĞ	(Turkey) Istanbul Aydın University

LIST OF PARTICIPANTS

Prof.Dr. Can Balkaya	Structural and Earthquake Engineering, Istanbul Aydin University
Prof.Dr. David Arditi	Invited Keynote Speaker, Professor of Civil and Architectural Engineering, Professor of Industrial Technology and Management Director of Construction Engineering and Management Director of Engineering Management, Illinois Institute of Technology, Chicago, Illinois, USA
Prof.Dr. Erol Tutumluer	Transportation Engineering, University of Illinois at Urbana-Champaign, USA
Prof.Dr. Oğuzhan Hasançebi	Structural Engineering, Middle East Technical University
Prof.Dr. Ali Okatan	Computer and Software Engineering, Istanbul Aydin University
Prof.Dr. Dilek Heperkan	Food Engineering, Istanbul Aydin University
Assoc. Prof. Dr. Onur Behzat Tokdemir	Construction Management, Middle East Technical University
Assoc. Prof.Dr. Gökçe Sargut	Business Management, Governors State University, IL, USA
Dr. Ahmet Muhtar Çıtıpıtıoğlu	PE, Engineering and Design Director, TAV Construction
Dr. İ. Haluk Çeribaşı	Environmental Engineering, ENCON Environmental Consultancy Co.
Dr. Yunus Emre Ayözen	General Manager of Strategy Development Department in Republic of Turkey Ministry of Transport and Infrastructure
Bilge Büdeyri, M.Sc.	Head of Strategy Department, Strategy Development Department in Republic of Turkey Ministry of Transportand Infrastructure

A. Sinan Alemdar, M.Sc.	Program Manager, EMAY International Engineering and Consultancy
Ugur Kaya M.Sc.	Design Coordinator, EMAY International Engineering and Consultancy
Levent Sezgin	Electronic Engineer, SU YAPI Engineering & Consulting Inc.
Fetih Kaya	Computer Engineering, Bahcesehir University
Bedirhan Uygun	Construction Management, Middle East Technical University
Ali Bedii Candaş	Construction Management, Middle East Technical University
Murathan Saygılı	Construction Management, Middle East Technical University
Hasan Eser	Structural Engineering, Middle East Technical University
İrem Dağdemir	Big Data Analytics and Management, Bahcesehir University
Mehmet Emre Özengen	Big Data Analytics and Management, Bahcesehir University
Görkem Özdener	Big Data Analytics and Management, Bahcesehir University
Ali Mortazavi	Graduate School of Natural and Applied Science, Ege University
Mahsa Moloodpoor	Mechanical Engineering, Ege University
Necdet Özbalta	Mechanical Engineering, Ege University

Ahmet Doğan	Computer Engineering, Istanbul Gelisim University
Ali Çetinkaya	Electronic and Communication Engineering, Kocaeli University
Ayşe İpek Gözütok	Biomedical Engineering, Bahcesehir University
Gülnihal Çiçekdağı	Big Data Analytics and Management, Bahcesehir University
Erkan Caner Ozkat	Mechanical Engineering, Recep Tayyip Erdogan University
G. Moskvins	Institute of Energetic, Latvia University, Republic of LATVIA
E. Spakovica	Institute of Energetic, Latvia University, Republic of LATVIA
M Moskvins	Univ. of Life Sciences and Technologies, Latvia University, Republic of LATVIA
Prof.Dr. Hasan A. Heperkan	Mechanical Engineering, Dean of Engineering Faculty, Istanbul Aydın University
Prof.Dr. Ömer Faruk Erkendirci	Mechanical Engineering, Istanbul Aydın University
Asst. Prof.Dr. Kaveh Dehghanian	Geotechnical Engineering, Istanbul Aydın University
Asst. Prof.Dr. Ali Etemadi	Department of Civil Engineering, Esenyurt University
Hüseyin Sağlık	Department of Bridge Engineering, Tongji University, CHINA
Şükrü Baykan	Software Engineering, Project Manager, Psteknik Informatics and Engineering Inc.

Ahmet Yılmaz Aktan	Infrastructure Projects, Project Manager, Inpro Engineering and Consultant Inc.
Abdullah Niğdelioğlu	Research Assistant, Civil Engineering Department, Istanbul Aydın University
Birkan Dağ	Research Assistant, Civil Engineering Department, Istanbul Aydın University
Bassant Sayed	Civil Engineering Student, Istanbul Aydın University
Mitra Salehi	Civil Engineering Student, Istanbul Aydın University
Didem Odabaşı Cingi	Building Inspection Department, Istanbul Aydın University
Isaac Kwabena Adomako	Shardfo-Tech Company Limited, GHANA
Daniel Kwame Hiro	Shardfo-Tech Company Limited, GHANA
Richard Ofori	Shardfo-Tech Company Limited, GHANA
Alp Olcay Mis	Software Engineering, Bahçeşehir University
Prof.Dr. Mehmet Gülgün	Materials Science and Engineering, Sabancı University
Prof.Dr. İlhami Çolak	Electrical and Electronics Engineering, Nişantaşı University
Asst.Prof.Dr. Serkan Ayvaz	Software Engineering Department, Bahçeşehir University
Dr. Erol Kalkan	PE, Structural Engineer, QuakeLogic Inc., Roseville, CA, USA

TABLE OF CONTENTS

AI and Big Data Applications in Disaster Management

Prof. Dr. Can BALKAYA

Department of Civil Engineering, Istanbul Aydın University, Florya Campus, Kucukcekmece, Istanbul, Turkey, canbalkaya@aydin.edu.tr

Abstract

AI and Big Data has changed the disaster management strategies with the advantage of integration different source information technology monitoring systems, GIS, storing, accessibility, real-time, distribution, analysis, decision systems as well as social media communication. Artificial intelligence has many potential advantages and practical applications to prevent potentially loss of human lives and destruction of property damages in disaster management. AI and big data are also used in emergency to evacuate critical regions and to make rescue works in a fast way. Natural disasters will be due to earthquakes, landslides, flooding, storms and tornadoes, fires, tsunamis or multi-hazard disasters in large regions. AI and big data can be important tool for emergency and disaster management with the usage of integration of sensors, drones, satellite, geospatial and weather data to provide real-time information about the damaged structures and landscapes. Artificial intelligence for disaster response by using image process techniques are discussed for data from GIS and tweets/ photographs by using open source media platforms. Case studies in disaster management: 1. planning and applications of disaster management projects for emergency response planning, 2. rapid earthquake damage assessment of buildings, 3. risk assessment of landslide-bridge-tunnel are given. Simulation models with artificial intelligence and big data for infrastructure damage into planning and testing emergency response to natural or man-made disasters scenarios for a city were obtained by integrating infrastructure, transportation, communication, power supply, healthcare model and population behaviour with the probability of damage severity, national risk index and hazard details.

Keywords: Artificial Intelligence, disaster management, monitoring, emergency response, infrastructure damage.

1. Introduction

AI and big data has changed the disaster management strategies with the advantage and integration of different source information technology between monitoring control systems, information systems, data transfer, big data storing capacities, accessibility to information, real-time information, network systems, analysis and decision system as well as warning systems with sirens and mobile phones in social media communication warning messages used in disaster management. Artificial intelligence has many potential advantages and practical applications to prevent potentially massive loss of human lives and destruction of property damages in disaster management. AI and big data are used in emergency to evacuate the danger areas immediately and to make rescue efforts in a fast way. Disasters can occur any time without warning. Natural disasters across the cities even the along the country will be due to earthquakes, landslides, coastal flooding or riverine flooding, hurricanes, storms, tornadoes, fires, tsunamis, volcanic activities, avalanche or multi-hazard disasters. AI and big data can be a solution for emergency and disaster management with the usage and combination of sensors, robots, drones, satellite, geospatial and weather data to provide real-time information about the damaged infrastructure, structures and landscapes.

Artificial intelligence for disaster response tools utilizes machine learning to automatically identify text and tweets that relate to crises to collect and categorize. Rapid damage severity assessment is done in the early stages of disaster response by using image processing techniques based on deep neural networks [1] from photographs both GIS and social media about infrastructure damage.

Transportation infrastructure is important for emergency management. Deep learning for traffic management in smart cities in disasters can be used for the transportation infrastructure systems. Deep learning requires a large amount of data training the model and becomes experts about the severity of disaster in three categories: low, medium and high then act accordingly by using this information and stored in condition assessment and retrofitting databases.

Natural disasters are also increasing because of climatic changes. Artificial intelligence will also to improve prevention and mitigation of disasters, and risk reduction [2]. The coordination on the disaster areas including specific information databases such as structural systems, structural construction material, number of stories etc. as building information of the damaged building or transportation infrastructure/bridge structure is necessary. Building information or transportation infrastructure information from condition assessment before the disaster can be integrated in big databases and if necessary existing databases will be adopted to databases. Also, improving disaster resilience by adding intelligent devices especially for critical important structures as tall buildings, bridge structures, tunnels, dams can be done for earthquakes, flooding, landslides, tsunamis, wind storms. In disaster management especially for large areas in multi-hazard disasters to assessment of damage structures in a short time coordination with the disaster community planning, project management, strategic collaboration and international disaster management by using satellite pictures, sensors, unmanned aerial vehicles, Internet of Things (IoT) and other important information storage, accessibility, real-time information, people and damage structure distributions are valuable big data to be used during and after the disaster in the disaster management.

2. Emergency Response Planning in Disaster Management

Municipal, province and national authorities provide resources and manage emergency response operations following a major disaster/earthquake to save lives, prevent injury, and protect property. In the sample emergency response tasks the following items can be considered in the disaster management: secure incident site, emergency medical triage search for victims, access control medical care sites and hospitals, plan traffic control and transportation supplies, restore power, provide water and food, remove debris, provide emergency shelters and inform the locations, provide critical fuel and additional rescue machines/cranes, building/structural information for rescue works, people information, provide temporary communication station for mobile phones of victims, etc.

2.1. Planning Information in Disaster Management

The basic steps for planning information in disaster management are: incorporate infrastructure damage into emergency response planning: electrical power, communications, transportation, buildings and life support systems; incorporate population behaviors into emergency response planning: align response operations with survivors' behaviors and needs; locate medical care based on population injury patterns: i.e. burns, lacerations, crushing; plan logistics, evacuation, water, and food distribution based on population movements; prioritize and deploy response operations based on behavioral requirements. For rapid solutions, artificial intelligence and integration of all available information's with previous model studies and big databases can be used.

2.2. Artificial Intelligence for Disaster Response (AIDR)

AIDR [3] is the open source software system platform to filter and classify social media messages related to emergencies, disasters, and humanitarian crises. AIDR uses human and machine intelligence to automatically tag up to thousands of messages per minute. The steps are: 1. Collection: to collect tweets (context, location, user profile, etc.) by filtering tweets using keywords and/or hashtags such as "earthquake" or "hurricane" related to a disaster while the collector is a word-filter, 2. Classifier: the tagger such as "*infrastructure damage*" is a topic-filter. Crisis related tweets and photographs are based on categories: urgent needs, infrastructure damage and resource deployment needs. Rapid damage severity assessment is done by using image processing techniques based on deep neural networks [1]. By using AIDR, volunteers tagged crisis related tweets and photographs in 72 hours during Nepal earthquake 7.8 magnitude in April 2015.

In their study [1], AIDR image processing system used for Hurricane Dorian, Category 2 in the eastern Caribbean barreling toward the northern Bahaman Islands and central Florida on Aug. 30, 2019 and about two weeks approximately 6,890,106 tweets were collected among them 280,063 unique image URLs were found. The total number of downloaded images was 279,819 and around 244 images failed to download. The human experts agreed with the system 20,887 times, as shown in Table 1 the three diagonal colored cells of total 20,887 human experts agreed while for 7,163 images disagreed with the system accuracy as 74% for this case.

Table 1. Damage severity assessment confusion matrix: machine vs. human judgments [1]

N=28,050		Machine		
		Severe Damage	*Mild Damage*	*None*
Human	*Severe Damage*	710	384	357
	Mild Damage	113	881	355
	None	721	5,233	19,296

2.3. National Risk Index

National risk index is given in FEMA in the map [4] for different individual natural hazard risk details for counties in United States. Risk index levels for United States are shown in Fig.1. The National Risk Index (NRI) incorporates physical and social vulnerability data to identify communities more at-risk to the adverse impacts of natural hazards. Using the national risk index is to support resilience, mitigation planning and risk communications.

Fig.**1.** Risk index levels in United States.

a) Hurricane Katrina,
New Orleans, Aug. 2005

b) Hurricane Sandy, New York, Oct. 2012

Fig.2. Hurricane Katrina and Hurricane Sandy disaster damages.

Hurricane Katrina disaster in New Orleans on August 2005 is shown in Fig.2a and at least 1,833 people died in the hurricane and subsequent floods. Total property damage was estimated at $81 billion (2005 USD). Hurricane Sandy

(Fig.2b) in New York, October 2012 killed more than 100 people. In the national risk index for New Orleans, hurricane plays an important disaster risk.

3. AI Applications of Disaster Management

3.1 Emergency Response Planning Application in Disaster Management

Virginia Tech research program models and studies infrastructure and populations under stress resulting from natural disasters. The National Risk Index app is intended to provide quick display of data. Hazard details [4] for Washington D.C. (in parenthesis number of events) is shown in Fig.3: Avalanche (0), Coastal Flooding (0), Cold Wave (2), Drought (98), Earthquake (0), Hail (435), Heat Wave (15), Hurricane(9), Ice Storm(26), Landslide(1), Lightning(1394), Riverine Flooding (61), Strong Wind (521), Tornado (5), Tsunami (0), Volcanic Activity (0), Wildfire (0), Winter Weather (47).

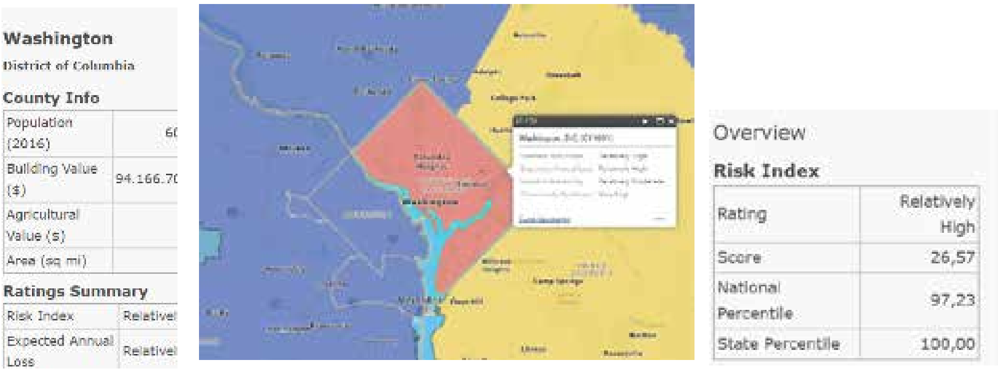


Fig. 3. Hazard risk details map for Washington D.C. [4].

At the Virginia Tech and Virginia Tech Arlington Research Lab in Washington D.C. simulation of different models are shown in Fig. 4 infrastructure model, transportation model, communication model, power supply model, health care models for emergency response planning in disaster management. These models were prepared with geographical information systems (GIS), High-Performance Computing (HPC) models and big data applications. Infrastructure damage and transportation network damage are given in Fig. 5 with damage levels showing different colors red: completely damage; orange: highly damage; green: medium damage; blue: light damage; white: no damage for disaster scenarios.

Incorporate study at Virginia Tech regarding infrastructure damage into planning and testing emergency response options to natural or man-made disasters were coupled with the behavior of individuals within an environment employing agent-based modelling at Virginia Tech Arlington Research Lab. Population behaviour on disaster scenarios are shown in Fig.6. for spontaneous evacuation, people evacuate away from disaster and individual selected more hazardous evacuation routes.

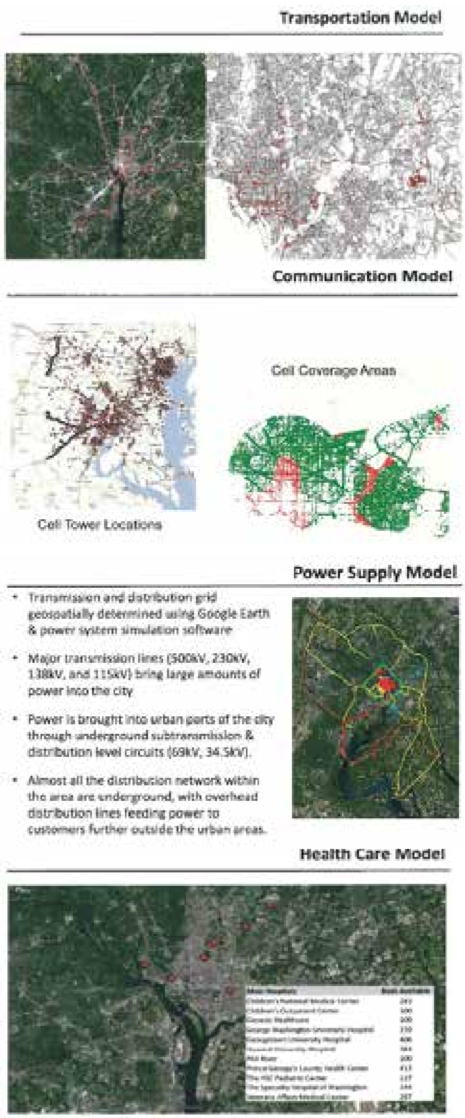
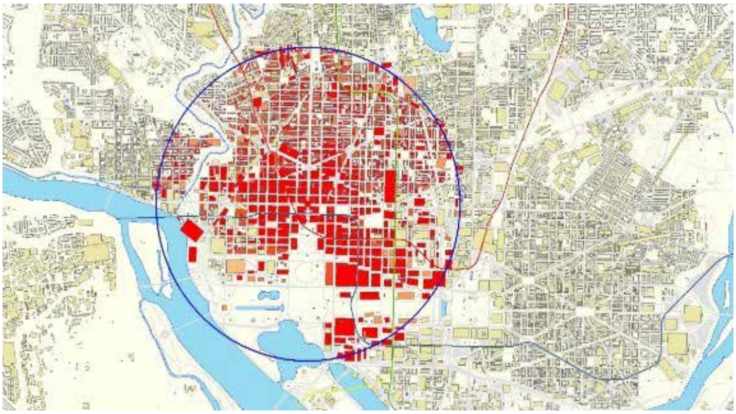
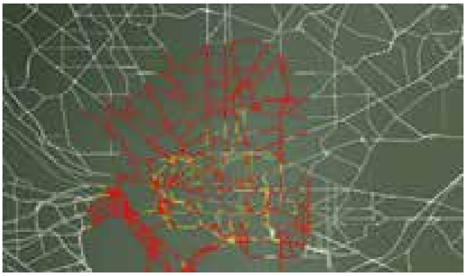


Fig. 4. Models for emergency response planning in disaster management.



Infrastructure Damage

- Red: completely damaged
- Orange: highly damage; reduced travel
- Green: medium damage
- Blue: light damage
- White: No damage



Transportation Network Damage

Fig. 5. Infrastructure damage scenarios with probability of damage.

- Spontaneous evacuation
- Blue tracking shows heavy *east-west* movement as people evacuate away from disaster.
- Individual selected *more hazardous* evacuation routes

Fig. 6. Population behaviour on disaster scenarios

3.2 Rapid Earthquake Damage Assessment of Buildings in Disaster Management

3.2.1. Conza Earthquake, Irpinia, Southern Italy

It is a well-known fact that post earthquake damage assessment is of great importance to develop an effective risk management after a strong earthquake. In Fig.7 shows the pre and post event satellite images of Conza Earthquake (Irpinia, Southern Italy). Artificial intelligence and big data can play as a vital tool to determine rapid damage assessment by integration of GIS and image process. On the other hand, due to limitations on the vertical information and weather conditions in satellite images, by using open access multi-views image/fusion systems are integrated in the image process real-time damage assessment of disaster areas [5].

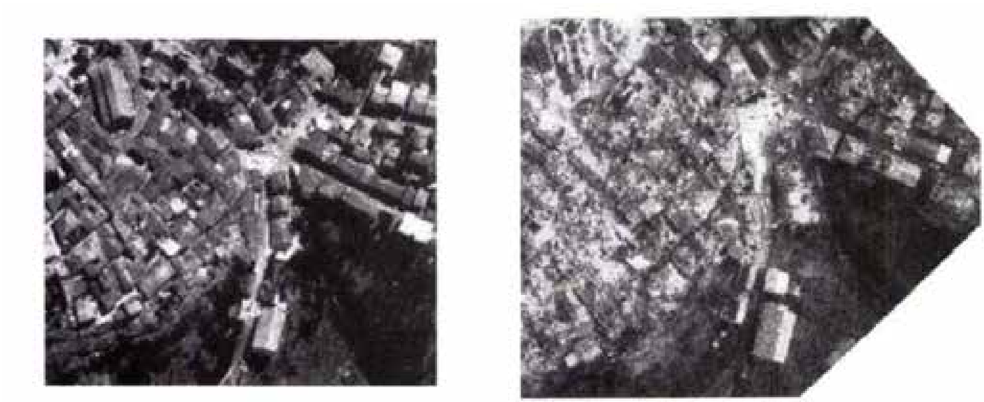


Fig. 7. Pre- and post-event images of Conza (Irpinia, Southern Italy) struck by tremendous earthquake [5].

3.2.2. Adana-Ceyhan Earthquake (post- earthquake buildings) and Marmara Region (pre-earthquake buildings)

Building information's are: spatial position, connection with other buildings, dilatation, building structural system, material quality, number of floors, floor types, basement, foundation system, soil information, wall indices, people living inside and contacted people information can also be integrated with GIS and image process. Damage index or structural wall index in x and y directions for the previous applications/case studies are stored for artificial intelligence expert systems for rapid evaluations [6]. 92 buildings damaged in the 1998

Mw = 6.3 Adana-Ceyhan Earthquake (post-earthquake structures) and 160 existing buildings in the western part of Turkey and Marmara region (pre-earthquake structures) were investigated. Seismic analyses were conducted at the Earthquake Research Center, METU. Using the results of seismic performance analysis of investigated buildings as low seismic performance level/strengthening required, low seismic performance level but strengthening not economic or normal seismic performance level strengthening not required and corresponding critical wall index in x and y directions as shown in Fig.8 were calculated for the RC building stock in Turkey. Knowledge-based information systems were developed.

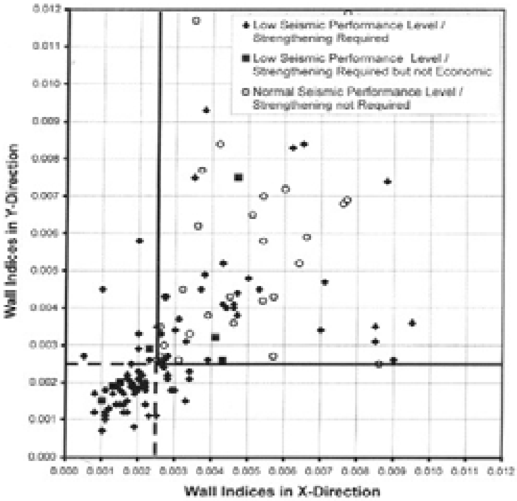


Fig. 8. Wall indices for seismic performance evaluation database.

3.2.3 RC Buildings in Zeytinburnu, Istanbul

A new model [7] is developed for rapid damage assessment of buildings by using integration of image processing of satellite images before and after the earthquake disaster and the structural information databases obtained from the field. The seismic displacement demand is calculated based on the spectral acceleration obtained from deterministic and probabilistic seismic hazard analysis. The capacity of the structure is defined by the relationship between base shear vs top displacement of each building type. The "seismic displacement demand" determined in terms of top displacement for both deterministic and probabilistic seismic hazard are compared with the "displacement capacity" of each building type for "life safety" and "collapse prevention" performance criteria. The performance of the buildings is determined through comparison of "seismic displacement demand" and "displacement capacity" for each building type. Demand/capacity ratio is determined for is seismic performance criteria define as "life safety" and "collapse prevention". Performance damage level relationship is defined as slight, moderate/extensive and collapse. The methodology has been used in a pilot project in Zeytinburnu district in Istanbul to assess the vulnerability of the building stock (Fig.9). Control systems can also be used for seismic behaviour of high rise/tall building structures [8].

Fig.9. Damage distribution of RC buildings in Zeytinburnu (probabilistic seismic input) [7].

3.3 Risk Assessment of Landslide-Bridge-Tunnel to Structural Control of Transport Network in Disaster Management

In this case study, an interactive database was used to control the consequences of catastrophic events in the risk assessment of land sliding GIS based real-time monitoring the environmental situation of an area of the Val Susa Valley (Province of Torino-Italy) [9]. For this purpose, the transport network infrastructure in the tunnel, bridge and land sliding site located in Pont-Ventoux can be controlled by using satellite images and monitoring sensors. Structural information of railway bridge and tunnel is shown in Fig.10. Management of information ArcMAP and implementation of data-base GIS dynamic, structural analysis of the bridge and simulation of disaster events for estimation of vulnerability curves are the phases of the project. Satellite images and area subjected to landslide are shown in Fig.11.

Fig. 10. Structural Information of Railway Bridge and Tunnel in Val-Susa Valley, Italy

Val Susa Valley Area Satellite Image Area Subject to Landslide

Fig. 11. Satellite Images and Landslide Area of Val-Susa Valley

System architecture for interactive database is given in Fig.12 with product characteristics- text communication, voice communication, GIS application, seismometer data. It is proposed that this system will be integrated with other control assessment sensors/devices, drones, robots and photographs from social communication and an artificial intelligence system with deep learning and experts' systems. Proposed system will be a rapid solution for disaster managements.

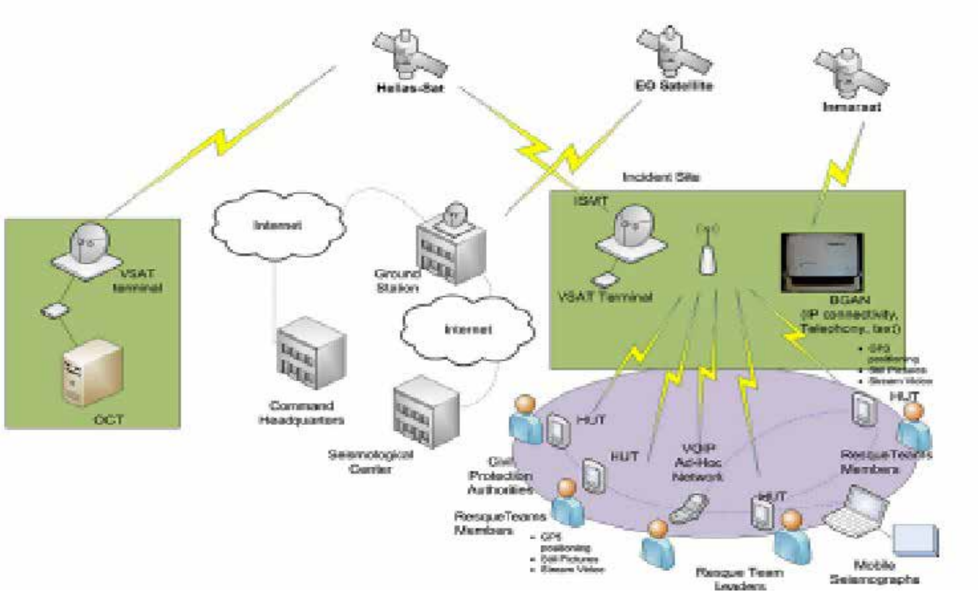


Fig. 12 Interactive database in disaster management [9].

4. Conclusion

Artificial intelligence plays an important role for fast decision of integrated big databases by using real time and stored information's with machine learning for damage assessment of infrastructure and structures. Monitoring system for condition assessment, GIS and image process techniques, warning systems, health monitoring systems, drones, robots and social media communication all information will be stored as big data for a city or large areas for a disaster even multi-hazards. Disaster response program filter and classify the social media messages by using image processing techniques based on deep neural network (deep learning). On the other hand, national risk index/maps with hazard details natural or man-made disasters can be used in the disaster simulation models, mitigation planning and risk communications.

Infrastructure model, transportation model, communication model, power supply model, health care models were considered for a city or critical regions for emergency response planning in disaster management. Building/bridge/ tunnel structure information, wall indices (for buildings), damage index and damage information can be stored for machine learning from previous studies. Real-time GIS information, sensors, cameras, drone video pictures, social media tweets and photographs will be integrated and used for fast artificial intelligence decisions in disaster management. Population behavior on disaster scenarios for spontaneous evacuation and evacuation routes can be studied for different cultural behaviors.

To prevent disaster that is mostly occurred without early warning, important building structures, bridges, dams and tunnels will be continuously real-time monitoring for condition assessment and retrofitting studies to reduce the risk. Risk assessment and disaster management planning by using AI and big data for critical regions and cities can be done for different natural disasters, man-made disasters or multi-hazard disasters.

Acknowledgements
Author thanks to Virginia Tech. Arlington Director Profesor Bruce Lawlor in 2012 for disaster management works and University of Pavia, Department of Civil Engineering, Professor A. Nieto for Val-Susa Valley case study applications. Author acknowledged 'Rapid Damage Assessment of Buildings to Avoid Demolishing' European Union LdV Transfer of Innovation (Project No: 2011-1-TR1-LEO05-27938).

References

[1] Muhammad Imran, Firoj Alam, Umair Qazi, Steve Peterson, and Ferda Ofli, "Rapid damage assessment using social media images by combining human and machine intelligence",17th International Conference on Information Systems for Crisis Response and Management (ISCRAM), Virginia, USA, 2020.

[2] Amina Khan, Sumeet Gupta, Sachin Kumar Gupta, "Multi-hazard disaster studies: monitoring, detection, recovery, and management, based on emerging technologies and optimal techniques", International Journal of Disaster Risk Reduction, vol. 47, August 2020.

[3] http://aidr.qcri.org/, artificial intelligence for digital response, 2015.

[4] https://hazards.geoplatform.gov/portal/apps/MapSeries/index.html?appid=ddf915a24fb24dc8863eed96bc3345f8

[5] C. Balkaya, F. Casciati, S. Casciati, L. Faravelli, M. Vece, "Real-time identification of disaster areas by an open access vision-based tool", Advances in Engineering Software, 2015.

[6] C. Balkaya, "Information systems for repair alternatives and initial cost estimation of damaged building structures", International Journal of Engineering and Technologies, IJET, vol. 4(2), pp. 80-88, June 2018.

[7] C. Balkaya, M. Erdik, C. Tuzun, "Rapid seismic damage assessment of existing buildings by integrating structural information in a GIS", ASCE, Journal of Performance of Constructed Facilities, accepted for publication, 2015.

[8] A. Okatan, C. Balkaya and M. Arslan, "Investigation of seismic behavior of high rise buildings by using control systems engineering", ISITES2017 5th International Symposium on Innovative Technologies in Engineering and Science, Baku-Azerbaijan, 799-808, 2017.

[9] C. Balkaya, F. Casciati, D. Bortoluzzi, G. Kostopoulos, N. Bogonikolos, "System Architecture for an interactive database to structural control of transport network infrastructure: A case study of railroad bridge", EACS2012, 5th European Conference on Structural Control, Genoa, Italy, 2012.

Comparison of Machine Learning Techniques for Text Classification

[1]Bedirhan Uygun, [2] Ali Bedii Candaş, [3]Onur Behzat Tokdemir

[1]Undergraduate Student e-mail: bedirhan.uygun@metu.edu.tr

[2]Graduate Student e-mail: ali.candas@metu.edu.tr

[3]Associate Professor e-mail: tokdemir@metu.edu.tr

Middle East Technical University, Faculty of Engineering,
Department of Civil Engineering, Ankara, Turkey

Abstract

Contract documents are one of the most crucial elements of a construction project that describes the parties' roles, obligations, and responsibilities involved within the project. A detailed description of the work, specifications, conditions, or even exceptions are also included in these documents. Most of the time, a construction contract consists of a set of documents because of its detailed structure to describe every project stage with no clues. Therefore, it is needed to be checked completely, and the crucial points should be classified. Because of the detailed structure of documents, manually reading the documents is arduous and lengthy to classify all the vital points with no mistake. Since manual checking is a time-consuming and open-to-error burden, automated systems have been developed and used several times. NLP (Natural Language Processing), a research area of artificial intelligence, enables machines to process natural language texts using several machine learning techniques. Machine learning-based text classification can be categorized into two major types: Supervised ML and Unsupervised ML. In this paper, the comparison of supervised and unsupervised ML techniques for text classification was discussed in terms of several aspects.

Keywords: Classification, machine learning, NLP, supervised learning, unsupervised learning.

1. Introduction

A construction contract is a legal agreement between the parties involved in a construction project. It clarifies all roles, wishes, and specifications about all construction stages, including design, procurement, construction, and execution. Despite any other contract types, construction contracts are considerably complex because of their high level of uncertainty [1]. Not only mega projects but also regular-sized construction projects bring a high complexity level because of their organizational complexity [2]. Unlike routine projects, many mega construction projects have failed to meet client's requirements in construction history. Although some of them get completed on schedule and budget, they fail to meet clients' needs because of incomplete identification of the project requirements [3]. Moreover, Squires and Murphy [4] state that if the general contractor misses the client's requirements' critical points, the subcontract would be incompetent. When a subcontract is incompetent, it can create a dispute between subcontractors and be resulted in undesirable conditions. Therefore, manual checking may lead to human errors resulting in any unwilling situations, such as disputes.

On the contrary, creating an automated system using NLP (Natural Language Processing) introduces an effective, quick, and safe way for the classification process. Thus, AI-based automated systems can help contract management by mitigating risks and saving time.

2. Background

2.1. Natural Language Processing

Natural language processing (NLP) is a subfield of artificial intelligence that enables machines to process natural language texts using several approaches. According to Chowdhury [5], NLP aims to develop specific methods and techniques to make the machines understand texts written in natural languages. This process's source is based on the information gathered by observing how human beings understand and use language. Information extraction (IE), information retrieval (IR), and speech recognition are some major usage areas of NLP, and text classification (TC) is a subdomain of NLP. Text classification aims to classify documents into specific predefined categories [6]. There are two types of approaches used for text classification that are rule-based and machine learning-based. Hand-coded rules are included in the rule-based method, while algorithms are used in the machine learning-based method to classify texts. However, the biggest drawback of the rule-based method is that it depends on exhaustive labor [7].

NLP can be used for text classification in several application areas in contract management, such as contract clause and meta-data analysis, identification and checking compliance and regulatory terms, etc.

2.2. Text Classification using Machine Learning Techniques

Until the ’90s, knowledge engineering (KE) was the most popular approach to text classification, where the hand-coded rules were manually defined by the direct knowledge of experts. In the ’90s, this approach has significantly lost its popularity due to an increasing trend of machine learning [8]. Machine learning (ML) is one of the numerous subfields that have been developed in computational intelligence. Researchers from different scientific disciplines have widely used it to deal with pattern recognition, prediction, optimization, and classification [9]. Text classification is categorizing the given data into a predefined set of categories. The classifying process is performed between the accumulation of texts and a given set of categories. Although the machine learning systems used for text classification (TC) are automated systems, they are not fully automated because these categories are usually created by the direct knowledge of the expert [10]. Machine learning-based text classification can be categorized into two major types: supervised ML and unsupervised ML. Human guidance is the distinctive point of supervised ML. The classifier algorithm is trained with training data set before the actual testing, where the human guidance is involved in the form of labeled data set and preparation of training data set. As a result, higher manual effort is needed for supervised ML algorithms. Unsupervised ML algorithms do not require human guidance. Instead, they cluster documents into potentially logical categories without training, and they discover and identify unlabeled datasets by themselves. Thus, the computational complexity of unsupervised ML is expected to be more than supervised ML.

3. Supervised Learning Algorithms

Supervised ML works based on the idea that algorithms are trained before testing in order to learn patterns and hypotheses. It is relatively expensive and demanding regarding the unsupervised ML because human labor is necessary while labeling the data [11]. The following mentioned are the most common supervised algorithms.

Naïve Bayes (NB) is a probabilistic classifier that uses Bayes’ Theorem. It is one of the simplest algorithms used in ML. It is commonly used in the training data because of its simplicity and success [12]. The drawback of NB is that it can show bias when training data is non-homogenous [13].

Logistic Regression (LR) in ML is used for classification problems, it is a predictive analysis algorithm, and it tries to select the best subjects for labeling to achieve higher classification accuracy. Active learning is utilized in ML models, so it requires a large training data size to achieve consistent results [14].

Support Vector Machine (SVM) is a supervised algorithm that has been widely used in classification tasks. Its application areas are text classification, credit risk analysis, medical diagnosis, and information extraction [15]. A parameter that is learned from training data is the critical property of the model when making predictions. SVM mostly achieves the highest precision among the other algorithms because of having more parameters. However, it is also time-consuming relative to the other algorithms because of its higher number of parameters and thus higher computational time [10].

Decision trees (DT) are described as a tree consisting of separate branches labeled individually. Internal nodes are labeled by term, branches are labeled by term weight, and leaves are labeled by categories [16]. Because of the structure of the trees, the computational complexity increases when the relationship increases among the variables, which may result in a decrease in the accuracy.

K-Nearest Neighbor (kNN) works based on the idea of finding the closest training data to classify unlabeled data [17]. Its efficiency depends on the selecting k value, which indicates the number of the nearest neighbors, and sometimes deciding the value of k can be complicated. Tree-based k-NN can be an option to reduce this complexity [18].

Artificial Neural Networks (ANN) are inspired by biological neural networks. They work in the same logic but in a highly simplified way. In an ANN model, processing units come together with interconnections and form a structure [19]. ANN usage is common where hierarchical multi-label classification is needed and where the relationships are not linear because the nature of ANN allows us to capture many relationships between data [20].

Random Forest (RF) is the combination of several tree classifiers on different subsets. The classification is determined by the majority of the votes. Its idea is basically deciding the correct output by looking at all outputs from different classifiers and selecting the majority as it is correct [21].

Table 1. Summary of the supervised algorithms [10,12,13,14,15,16,22]

Method	Advantages	Disadvantages
LR	Easy to implement, efficient for training.	Requires large training data size to work efficiently, not suitable for non-linear problems.
NB	Works fast and be able to solve multi-class prediction problems, works efficiently even with less training data.	Assumes all the features as independent features so that no interactions can be achieved between features.
SVM	Effectively work in high dimensional spaces, mostly achieves high precision.	For training and testing more time-consuming relative to the other algorithms.
DT	Easy to understand and offers clear visualization.	Not suitable for large datasets, affected by noise.
kNN	Easy to implement, requires no training period.	Searching for the best k value, lazy learning model

Table 2. Statistical comparison of supervised algorithms [23,24,25,26]

Results reported by	Technique	Classifier Used	Performance (%)		
			Accuracy	Recall	Precision
Caldas et al. [23]	Vector representation	SVM	91.12	-	-
		NB	58.82	-	-
		kNN	49.11	-	-
Ul Hassan et al. [24]	N-gram (unigram)	SVM	98.51	-	-
		NB	95.53	-	-
Zhou et al. [25]	Semantic Hierarchy	SVM	-	85.2	93.5
		kNN	-	74.8	93.7
		RF	-	86	91.5
Ul Hassan et al. [26]	N-gram (unigram)	NB	87.48	75	96
		SVM	93.17	99	93
		LR	94.18	92	97
		kNN	92.03	86	93
		FNN (Feedforward Neural Networks)	92.8	84	89
		DT	89.38	98	93

Table 2. is constructed to show the variation of the performance of the supervised algorithms in different studies. Accuracy, precision, and recall are the most common measures of performance. Accuracy is the ratio of the correct observations to the total, while precision and recall are the ratios of correct positive observations to the positive total and total, respectively. According to the table, it seems that one algorithm can show different behaviors in different studies because of its characteristics. However, we can come up with the idea that some algorithms can show consistently good performance on classification problems such as SVM. According to the overall evaluation, SVM yields the best performance, whereas NB achieves the least. However, except for NB, there is no significant performance difference considering other algorithms. Performance is a crucial measure for selecting the algorithm; however, it is not the only one. Because of their characteristics, the algorithms might have some problems related to implementation, execution time, or fitting. Depending on the problems, the algorithms are needed to be considered and compared in several ways. For example, if decision trees and neural networks are compared, advantages and disadvantages are almost the opposite. Decision trees are easy to be understood by humans, but they are troubling when there is noise in data. However, for neural networks, it is just the opposite. They are hard to be understood by humans, but they have almost no deal with noisy data [10]. Thus, we need to select an algorithm that depends on not only accuracy but also its characteristics.

4. Unsupervised Learning

Unsupervised ML, also called cluster analysis, is a type of ML technique that clusters data. Unlike supervised learning, no training data is introduced to the system to execute the clustering process [16]. Unsupervised algorithms have a principle that works as clustering documents in a group that are more similar to each other. Although it seems complex initially, it will increase its efficiency when much data is introduced [27].

Hierarchical Clustering Algorithms are among the most common unsupervised algorithms that try to build hierarchy between the clusters. The classification process is operated by merging and splitting the clusters formed in either a bottom-up or top-down fashion. The algorithm is not capable of performing adjustments after a merge or split has been executed. Therefore, a merge, or split decision may lead to low-quality output if it is not chosen at an appropriate step [28].

K-means Clustering is another standard and relatively simple unsupervised algorithm. The purpose of the algorithm is to collect data points into k unknown clusters where k stands for cluster number. It is relatively fast and straightforward and usually gives good results [29].

5. Comparison of Supervised and Unsupervised Methods

In this study, the comparison of machine learning techniques for text classification is discussed. According to Sebastiani [8], there is no specific type of best ML algorithm for all domains, depending on the conditions. Besides, supervised ML techniques usually yield better performance but need extra human guidance rather than unsupervised ML. However, these comparisons are reliable only if the same author performs the studies under the same controlled conditions.

In their research related to automated compliance checking, Salama and El-Gohary [30] selected a supervised ML approach instead of an unsupervised one. The authors argued that supervised TC is more time-consuming than unsupervised TC because preparing a training data set requires time and effort. However, supervised TC is advantageous in the case of performance. Since it works with human guidance, it usually gives higher performance (in terms of precision and recall) than unsupervised TC. Also, they claimed that domain-specific labels are needed since the problem they have is domain-specific, so the supervised approach fits better in these situations. In addition, Zhou and El-Gohary [25] also stated that unsupervised ML's accuracy is often less accurate concerning supervised ML because of the lack of training and human guidance.

Ozgur [31] has compared the quality of clusters produced by supervised and unsupervised approaches. Several supervised and unsupervised algorithms have been evaluated on five different tests consist of different datasets, and F-measure has been used for performance measurement. F-measure (or F-score) is a cluster evaluation metric that combines the precision and recall values as shown in Eq. 1. The highest value of the F-measure is 1.0, which represents the perfect precision and recall.

$$\textit{F-Measure} = \frac{(2*\text{Precision}*\text{Recall})}{(\text{Precision}+\text{Recall})} \quad (1)$$

Table 3. Results of the comparison study [31]

	F-measure				
Test / Algorithm	1	2	3	4	5
k-means	0.984	0.497	0.597	0.196	0.473
kNN (traditional)	0.983	0.593	0.763	0.247	0.58
NB	0.983	0.485	0.525	0.111	0.463
SVM	0.991	0.666	0.816	0.297	0.652

According to Table 3., one can say that supervised algorithms tend to yield higher performances in comparison to the unsupervised ones. Unlike unsupervised techniques, supervised techniques use the similarity between texts and the information of class labels they have been learned in the training session. However, we can observe that k-means as an unsupervised algorithm usually yields better performance than a supervised algorithm NB. As another observation, it can be said that SVM is the most successful algorithm in terms of performance according to the overall average. However, it is seen from Table 3. that algorithms' performances changed significantly on different datasets. For instance, NB has an F-score of 0.983 on Test 1, whereas this number was decreased to 0.111 on Test 4. As a result, the search for a best-fitting algorithm according to the problem is more logical than the search for the algorithm with the best overall scores.

Ko and Seo [32] have compared supervised and unsupervised learning approaches for text categorization. The authors aimed to see the difference between hand-labeled training data and automated labeled training data by unsupervised learning. The proposed method in the study was preparing the training data by unsupervised learning because of the high cost and time required for hand-labeling. They resulted in that supervised and unsupervised (proposed method) systems reached 75.6% and 71.8% F-scores, relatively. Because of the insignificant difference between methods, the authors argued that the proposed system could be used in the areas where the cost of hand-labeling is high.

In another research [33], the authors developed an unsupervised approach and compared its performance with the previous supervised approaches. They stated that the highest achieved performance is 98.69% for recall and 92.70% for precision, which are slightly higher than those achieved in the previous supervised approaches.

SVM's and neural networks usually achieve higher grades of performance when they are dealing with multi-dimensional problems. Differently, logic-based algorithms such as DTs are usually better when dealing with categorical/discrete featured problems. K-NN is hypersensitive to irrelevant features because of the algorithms' nature so that they are very inefficient when the problem has irrelevant features [20].

6. Application Areas in Construction Industry

Automated TC has almost reached the effectiveness of the human experts working in that area. Human experts' effectiveness is not even 100%, and machines can complete the classification tasks in just seconds. Thus, the classification process can be improved with the combined work of humans and machines [8]. As a result, automated TC techniques have also been tested and used several times in the construction industry. Some of the studies, other than construction contracts, related to application areas of ML approaches in the construction industry are listed in Table 4.

Table 4. Examples of application areas of ML Approaches [23,24,25,34,35,36,37,38,39]

Study	Technique	Application Area
Ur-Rahman et al. [34]	Supervised ML	Classification of Post Project Reviews (PPRs)
Chi et al. [35]	ML-based TC	Classifying safety standards for field inspection
Chi et al. [36]	Supervised ML	Extracting safety text resources to support Job Hazard Analysis
Sajadfar et al. [37]	Deep Learning	Text detection and classification of construction documents
Fang et al. [38]	Deep Learning	Classification of near-misses from safety reports
Caldas et al. [23]	Supervised ML	Organization and access in construction management information systems based on automatic hierarchical classification of construction project documents
Zhong et al. [39]	Deep learning	Automatic classification of Building Quality Complaints (BQCs)
Ul Hassan et al. [24]	Supervised ML	Automated identification and extraction of client requirements from construction contracts
Zhou et al. [25]	Supervised ML	Automated Compliance Checking of Environmental regulatory documents
Zhou et al. [33]	Unsupervised ML	Automated compliance Checking of Construction regulatory documents

Construction contracts bring certainty by transforming uncertainties into defined risk events. When a contract is incomplete or miss-checked, it is called suboptimal. Suboptimal contracts can cause erosion of profit margin if contractual incompleteness arises [40]. However, according to Table 4., ML-TC's application area is not limited to contracts. The reports of past events and projects have a big potential for efficiency increase in construction projects, and some of the researchers have already studied those areas [34, 35, 38].

Although contract documents have a significant usage area for machine learning classification, as stated in previous parts of this study, there are many other machine learning TC applications in the construction industry. According to the study [34], the author claimed that past projects' unstructured data could be helpful by transforming into usable formats. Decision-makers can use this identified knowledge, and it can be adapted for future projects. Moreover, in their research [33], the authors stated that automated systems could save time and resources and reduce the error probability in compliance checking.

The effective usage of ML-based systems depends on the collected data. The process consists of two main steps. The first step is to collect as much data as possible. It can be previous work analysis and reports or contracts etc. The second step is turning these unstructured data into a structured readable format for automated systems. As a result, a systematic usage of the ML-based classification system can be achieved and utilized in many ways, as stated in Table 4.

7. Conclusion

This paper describes text classification methodologies by different machine learning algorithms and their application areas in the construction industry, especially for the contract documents, based on the related literature review. Since manual checking is time-consuming, resource-consuming, and an open-to-error burden because of humans' nature, an automated TC approach has been considered. TC process may be more efficient and productive by combining humans and automated systems [20]. The practical usage the machine learning for contract document checking can reduce the costs for either short-term by reducing the effort and time required for manual reading, or long term by preventing disputes or any other unwilling situations arising from contractual misunderstandings. We have discussed and compared the TC algorithms' characteristics, usage areas, and performance. Our aim was not to select a superior algorithm but to understand algorithms' natures to select the most suitable one under different conditions. We have come up with the main idea that there is no algorithm working with the best performance in all conditions, but some have higher overall performance, such as SVM.

Since all algorithms have their strong and weak sides, creating a hybrid system by integrating two or more algorithms may be a solution to increase the existing algorithms' performance. Based on a related literature search, hybrid approaches were developed in several studies; however, they are not included since they are out of the scope of this study.

In this study, a systematic approach was explained for selecting suitable algorithms under different conditions instead of performing countless numbers of trial errors to find the optimum. However, further study is needed to apply TC algorithms to different cases and document the results.

References

[1] Klee, L. (2015). International construction contract law. (1st ed.). Hoboken, NJ: John Wiley & Sons.

[2] Murdoch J. & Hughes W. (2008). Construction contracts: law and management (4th ed.). Madison Ave, New York/USA

[3] Zhang, Y., and H. Liu. "Discussion on issues related to the explanation of contract for construction project." Int. Conf. on Information Management, Innovation Management and Industrial Engineering, ICIII 2012, vol. 2, 454–457. New York: IEEE, 2012.

[4] Squires III, W. R., and Murphy, M. J. (1983). "The impact of fast track construction and construction management on subcontractors." Law & Contemp. Probs., 1(46), 55.

[5] Chowdhury, G. (2003) Natural language processing. Annual Review of Information Science and Technology, 37. pp. 51-89. ISSN 0066-4200

[6] Manning, C. D., and Schutze, H. (1999). Foundations of statistical natural language processing, MIT Press, Cambridge, MA

[7] Manning, C., Raghavan, P., and Shutze, H. (2009). An introduction to information retrieval, Cambridge University Press, Cambridge, U.K.

[8] Sebastiani, F. (2002). "Machine learning in automated text categorization." J. ACM Comput. Surv., 34(1), 1–47.

[9] Jain, A. K., Mao, J., and Mohiuddin, K. M. (1996). "Artificial Neural Networks: A Tutorial," In IEEE *Computer*, Vol. 29, No. 3, pp. 31-44.

[10] Pawar, P. Y., & Gawande, S. H. (2012). A Comparative Study on Different Types of Approaches to Text Categorization. International Journal of Machine Learning and Computing, 423–426. https://doi.org/10.7763/ijmlc.2012.v2.158

[11] Liu, X., Wang, J., Yin, M., Edwards, B., & Xu, P. (2017). Supervised learning of sparse context reconstruction coefficients for data representation and classification. Neural computing and applications, 28(1), 135-143.

[12] Sucar, L.E. (2015). Bayesian classifiers. In L. E. Sucar, Probabilistic graphical models (pp. 41-62). Springer. https://doi.org/10.1007/978-1-4471-6699-3_4

[13] Rennie, J., Shih, L., Teevan, J., and Karger, D. (2003). "Tackling the poor assumptions of naïve Bayes text classifiers." Proc., 20th Int. Conf. on Machine Learning (ICML), Vol. 20, ACM, New York, 616–623.

[14] Maalouf, M., & Siddiqi, M. (2014). Weighted logistic regression for large-scale imbalanced and rare events data. Knowledge-Based Systems, 59, 142–148. https://doi.org/10.1016/j.knosys.2014.01.012

[15] Altınel, B., Ganiz, M. C., &Diri, B. (2015). A corpus-based semantic kernel for text classification by using meaning values of terms. Engineering Applications of Artificial Intelligence, 43, 54–66. https://doi.org/10.1016/j.engappai.2015.03.015

[16] Greiner, R., & Schaeffer, J. (2001). "AIxploratorium – Decision Trees", Department of Computing Science, University of Alberta, Edmonton, ABT6G2H1, Canada.2001. http://www.cs.ualberta.ca/ ~aixplore/ learning/ DecisionTrees

[17] Lorena, A. C., Jacintho, L. F., Siqueira, M. F., Giovanni, R. D., Lohmann, L. G., de Carvalho, A. C., & Yamamoto, M. (2011). Comparing machine learning classifiers in potential distribution modelling. Expert Systems with Applications, 38(5), 5268–5275. https://doi.org/10.1016/j.eswa.2010.10.031

[18] Maillo, J., Ramfrez, S., Triguero, I., & Herrera, F. (2016). kNN-IS: An iterative spark-based design of the knearest neighbors classifier for big data. Knowledge-Based Systems, 117, 3-15. https://doi.org/10.1016/j.knosys.2016.06.012

[19] Yegnanarayana, B. (2009). Artificial neural networks. PHI Learning Pvt. Ltd.

[20] Kotsiantis, S. B., Zaharakis, I., & Pintelas, P. (2007). Supervised machine learning: A review of classification techniques. Emerging artificial intelligence applications in computer engineering, 160(1), 3-24.

[21] Oshiro, T. M., Perez, P. S., & Baranauskas, J. A. (2012, July). How many trees in a random forest? In International workshop on machine learning and data mining in pattern recognition (pp. 154-168). Springer, Berlin, Heidelberg.

[22] Yesilbudak, M., Sagiroglu, S., & Colak, I. (2017). A novel implementation of kNN classifier based on multi-tupled meteorological input data for wind power prediction. Energy conversion and management, 135, 434-444.

[23] Caldas, C. H., Soibelman, L., & Han, J. (2002). Automated classification of construction project documents. Journal of Computing in Civil Engineering, 16(4), 234-243.

[24] Ul Hassan, F. U., & Le, T. (2020). Automated requirements identification from construction contract documents using natural language processing. Journal of Legal Affairs and Dispute Resolution in Engineering and Construction, 12(2), 04520009.

[25] Zhou, P., & El-Gohary, N. (2016b). Domain-Specific Hierarchical Text Classification for Supporting Automated Environmental Compliance Checking. Journal of Computing in Civil Engineering, 30(4), 04015057. https://doi.org/10.1061/(asce)cp.1943-5487.0000513

[26] Ul Hassan, F., Le, T., & Tran, D. H. (2020). Multi-Class Categorization of Design-Build Contract Requirements Using Text Mining and Natural Language Processing Techniques. In Construction Research Congress 2020: Project Management and Controls, Materials, and Contracts (pp. 1266-1274). Reston, VA: ASCE.

[27] Shafiabady, N., Lee, L. H., Rajkumar, R., Kallimani, V. P., Akram, N. A., & Isa, D. (2016). Using unsupervised clustering approach to train the support vector machine for text classification. Neurocomputing, 211, 4–10. https://doi.org/10.1016/j.neucom.2015.10.137

[28] Rani, Y., Rohil, H. (2013). “A Study of Hierarchical Clustering Algorithm”. International Journal of Information and Computation Technology, 3(10).

[29] Fahim A M,Salem A M,Torkey F A, “An efficient enhanced k-means clustering algorithm” Journal of Zhejiang University Science A, Vol.10, pp:1626-1633,July 2006

[30] Salama, D. M., & El-Gohary, N. M. (2016b). Semantic Text Classification for Supporting Automated Compliance Checking in Construction. Journal of Computing in Civil Engineering, 30(1), https://doi.org/10.1061/(asce)cp.1943-5487.0000301

[31] Ozgur, A. 2002. Supervised and Unsupervised Machine Learning Techniques for Text Document Categorization. M.Sc. thesis. Bogazici University.

[32] Ko, Y., & Seo, J. (2000). Automatic text categorization by unsupervised learning. In COLING 2000 Volume 1: The 18th Int. Conference on Computational Linguistics.

[33] Zhou, P., & El-Gohary, N. (2016). Ontology-Based Multi-label Text Classification of Construction Regulatory Documents. Journal of Computing in Civil Engineering, 30(4), 04015058. https://doi.org/10.1061/(asce)cp.1943-5487.0000530

[34] Ur-Rahman, N., & Harding, J. A. (2012). Textual data mining for industrial knowledge management and text classification: A business-oriented approach. Expert Systems with Applications, 39(5), 4729-4739.

[35] Chi, N. W., Lin, K. Y., El-Gohary, N., & Hsieh, S. H. (2016). Evaluating the strength of text classification categories for supporting construction field inspection. Automation in Construction, 64, 78-88.

[36] Chi, N. W., Lin, K. Y., & Hsieh, S. H. (2014). Using ontology-based text classification to assist Job Hazard Analysis. Advanced Engineering Informatics, 28(4), 381-394.

[37] Sajadfar, N., Abdollahnejad, S., Hermann, U., & Mohamed, Y. (2019). Text detection and classification of construction documents. In ISARC. Proceedings of the International Symposium on Automation and Robotics in Construction (Vol. 36, pp. 446-452). IAARC Publications.

[38] Fang, W., Luo, H., Xu, S., Love, P. E., Lu, Z., & Ye, C. (2020). Automated text classification of near-misses from safety reports: An improved deep learning approach. Advanced Engineering Informatics, 44, 101060.

[39] Zhong, B., Xing, X., Love, P., Wang, X., & Luo, H. (2019). Convolutional neural network: Deep learning-based classification of building quality problems. Advanced Engineering Informatics, 40, 46-57.

[40] Agrawal, A. K., Jagannathan, M., & Delhi, V. S. K. (2021). Control Focus in Standard Forms: An Assessment through Text Mining and NLP. Journal of Legal Affairs and Dispute Resolution in Engineering and Construction, 13(1), 04520040.

AI-Based Electricity Demand Prediction: A Case Study for An Industrial Area

Murathan Saygili[1], İbrahim Levent Sezgin [2], Onur Behzat Tokdemir [3]

[1]Department of Civil Engineering, Faculty of Engineering,
Graduate Student, Middle East Technical University, Ankara, Turkey,
e-mail: murathan.saygili@metu.edu.tr

[2]Electronic Engineer, SU YAPI Engineering & Consulting Inc., Ankara, Turkey,
e-mail: levent.sezgin@suyapi.com.tr

[3]Department of Civil Engineering, Faculty of Engineering,
Assoc. Prof. Dr., Middle East Technical University, Ankara, Turkey,
e-mail: tokdemir@metu.edu.tr

Abstract

Electricity consumption fluctuates in a day due to changes in various parameters. The changes in consumption rates are important for production planning since the current state of technology does not allow storing the electricity. Hence what is produced should be consumed simultaneously. In the event of predicting the demand inaccurately, redundant energy creates losses in money and energy resources; similarly, insufficient energy causes power shortages in the grid. Thus, the demand should be predicted accurately so that electricity production can be planned correctly. Current prediction methods are mostly based on expert opinions and manual spreadsheets. However, these methods are not accurate due to unpredictable variables affecting electricity consumption. Other methods using innovative techniques such as artificial intelligence (AI) usually predict up to daily frequency. Therefore, there is a gap in hourly electricity consumption prediction in the literature. This study provides a literature review including several prediction models and gives the early results of an AI-based prototype. The research shows significant variables that affect electricity consumption, and it reveals the accuracy score of different AI models, including decision trees, support vector machines, and deep learning tools. Finally, the changes in the consumed amount of electricity due to COVID-19 are shown. The data used in the study was taken from an industrial region, including hourly electricity consumption for three years. This study contributes to the literature by gathering different studies, presenting the results of an AI prototype, emphasizing the importance of data, and revealing insights into electricity consumption changes in pandemics.

Keywords: Electricity Demand Prediction, Industrial Electricity Consumption, Artificial Intelligence, Machine Learning,

1. Introduction

Electricity is one of the mainstream life sources of civilized societies. It is needed for both producers and consumers to keep industrial plants running and provide sufficient energy to residents. The increasing population and industrial activity demand more and more energy from the grid, making electricity production planning more important [1,2]. In addition, preventing energy waste became a must to have sustainable cities and grids, so energy management is the key [3]. To have a properly functioning production plan, the demand for electricity should be known by the electricity producers so that a balanced supply-demand relationship can be maintained [4]. The reason is that current methods do not allow storing the electricity, although there are ongoing studies pursuing that goal [5]. As a result, the generated electricity is highly dependent on the consumer side. If the electricity generation is higher than demand, an excessive amount is wasted, as well as money and energy resources. On the contrary, if the electricity supplied to the grid is lower than the demand, consumers suffer from power shortages, and additional power plants need to be activated due to emergencies. Similarly, the result is inevitable financial loss. Obviously, both scenarios lead to inefficiencies and losses [6]. Therefore, accurate demand prediction is crucial.

There is a wide range of prediction methods composed of traditional techniques such as regression-based approaches and innovative techniques such as artificial intelligence. A regular electricity consumption data set possesses a nonlinear structure; therefore, innovative techniques are more adaptive to such nonlinearity in the data [1]. Also, the fact that electricity consumption is affected by a plethora of variables makes accurate prediction more challenging [7–9]. The literature includes a variety of research trying to reach the least prediction error in electricity consumption so that electricity generation can be planned wisely. However, they focus on a standard set of input variables, and they do not address the changes in the consumption due to COVID-19 [10,11]. Furthermore, their prediction frequencies are mostly daily or monthly, and the studies trying to forecast hourly consumption are not too many. It is known that electricity demand prediction can be in different frequencies. These prediction frequency intervals are called short-term, medium-term, and long-term forecasting [12]. Hence, the studies are also divided into three, and this research pursues short-term prediction in hourly intervals so that hourly load given to an industrial facility can be observed. Also, this paper aims to show the relevance of parameters affecting the electricity demand rather than trying to reach minimum prediction error. Although the tool's primary goal is to predict electricity consumption with minimum error, this study provides the initial findings of the research.

2. Literature Review

The literature consists of different studies using a wide range of intelligent prediction tools aimed to predict electricity demand. Jawad et al. (2020) [6] aimed to achieve the least cost electricity load forecasting through AI and ML models. Their technique includes multiple linear regression, k-nearest neighbors, support vector machines, random forest, and AdaBoost. Similarly, Ozer et al. (2021) [12], Xu et al. (2020) [13], and Ullah et al. (2020) [14] used deep learning and smart meter sensors for short-term load prediction. Deep learning has been the focus of researchers, and there are studies aiming at the short-term forecast for electricity consumption [15]. Also, Bedi and Toshniwal (2019) [1] and Liao et al. (2020) [16] created a deep learning-based model to predict long-term electricity demand, and they compared the findings with the result of other models based on artificial neural network, recurrent neural network, and support vector machines. Similarly, Marino et al. (2016) [3] and Jeyaraj and Nadar (2020) [17] used deep learning to reveal building-level electricity consumption and compared the method's accuracy with the outputs of artificial neural networks and support vector machine tools. However, what they [3] did is to predict consumption with a one-minute frequency. Their [3] following study used convolutional neural networks with the same purpose but focusing on one-hour frequency rather than one minute [18]. Kheirkhah et al. (2013) [2] combined AI techniques, including artificial neural networks to explain monthly and seasonal changes in the consumption trends. Atef and Eltawil (2019) tried to decrease peak load demand through an effective consumer demand response strategy. Their method was based on deep learning and hourly demand prediction. According to the findings, deep learning yielded significantly better predictions than other tools such as support vector machines. As for generalized linear models, Vaghefi et al. (2015) [19] used the method to model industrial electricity loads. Also, Ahmad et al. (2018) [20] used different machine learning tools, including generalized linear models, to predict short, medium, and long-term energy usage. Considering decision trees, gradient boosted trees, and random forest, Galicia et al. (2018) [21] used these three AI methods to observe and forecast the electrical demand. Likewise, Bellahsen and Dagdougui

(2021) [22] used random forests, k-nearest neighbor and deep learning tools for short term load forecasting. Other than comparing different AI tools, studies conducted a comparison using the same technique but changing learning parameters such as the activation function [23]. To clarify the consumption trends during the pandemic, Lu et al. (2021) [24] proposed a prediction framework that accounts for the impact of COVID-19 on electricity consumption. The previous studies are gathered in the study of Kuster et al. (2017) [25]. There are several other publications aiming at similar objectives, and these studies are provided in Table-1.

3. Methodology

This study shows the early results of an AI-based prediction tool for electricity demand. The tool is fed with data including past hourly electricity consumption values of an industrial area (dependent variable) from 2018 to 2020, and the independent variables are collected from several sources. To understand what type of variables are needed, two main sources of information are used. One of them is past studies aiming to achieve accurate electricity demand prediction. The studies are obtained from search engines such as Scopus, Web of Science, and universities' thesis databases. The second source is expert knowledge. Experts on electricity and electricity prediction participated in this study to create a sophisticated tool. As a result, the decision variables are selected. The data includes hourly electricity consumption, which is the dependent variable of the prototype. The independent variables are also collected using external sources. While creating the model, RapidMiner Studio is used to preprocess the data and set up AI models. The software is chosen considering the wide range of models it provides, the possibility of changing model parameters, and the easy-to-use interface with visuals. The outputs taken from the software are provided in this study. Finally, to observe the changes during the COVID-19 period (after March 2020), the monthly sum of consumption data is revealed from the perspective of the industrial plant.

4. Types of AI Techniques and Variables

Artificial intelligence is an umbrella term covering multiple techniques. Each of these methods addresses different needs and datasets, and their accuracy varies accordingly [6]. Therefore, different AI tools and different dataset combinations are utilized in this study.

4.1. Prediction Techniques

To measure the accuracy of different prediction tools under different data combinations, these tools are tested in RapidMiner Studio. These techniques are generalized linear model, deep learning, decision tree, random forest, gradient boosted trees, and support vector machines. Table-1 shows the studies using the corresponding tools. It should be stated that the table includes publications that are not mentioned in the literature review due to their scope but given considering their prediction technique.

Table 1: Prediction Tools Used in This Study

Techniques	Literature
Generalized Linear Model	[16,19]
Deep Learning	[1,3,23,26,12–18,22]
Decision Tree	[20,21]
Random Forest	[8,16,21,22,27]
Gradient Boosted Trees	[21,28]
Support Vector Machines	[6,8,16,24,29–40]

4.2 Input Variables

As mentioned earlier, electricity consumption is affected by several different parameters. The studied AI tool is planned to include a wide range of datasets incorporating different variables. The reason for that is to observe the relevancy of these variables with electricity consumption through different data combinations. Although this is one of the primary goals of the tool, this study only includes a limited number of data types. The types of data and their frequency are given in Table-2.

5. The Early Results of the Study

Three models are created in this study. The combinations of data variables are given in Table 2. The AI tools are applied to each combination, and the results of the models are given in Fig. 1, Fig. 2, and Fig. 3. In the first combination, the variables whose impact on electricity is widely acknowledged are given as input. In the second combination, wastewater treatment data that belongs to the region covering the location of the industrial area is added to the data used in Combination-1. In the last combination, drinking water consumption data is added to the data used in Combination-2. In this way, it is aimed to observe new variables that are not commonly studied but have a correlation with electricity consumption. Finally, the impact of COVID-19 on the electricity consumption trends is shown in Fig. 4.

Table 2: Data Combinations

Variables	**Role**	**Frequency**	**Combination -1**	**Combination -2**	**Combination -3**
Electricity	Output	Hourly	X	X	X
Calendar Data	Input	Daily	X	X	X
Exchange Rate (USD/ TL)	Input	Daily	X	X	X
Industrial Production Index	Input	Monthly	X	X	X
COVID-19	Input	Daily	X	X	X
Wastewater Treatment	Input	Monthly		X	X
Drinking-Water	Input	Monthly			X

5.1 Combination -1

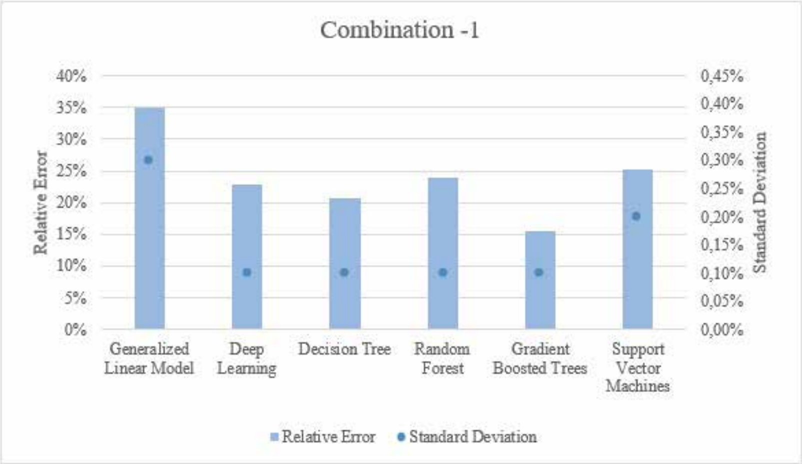


Fig. 1. RapidMiner Studio Output - Model 1

The Most Accurate Tool: Gradient Boosted Trees
Relative Error: 15.4%
Standard Deviation: ±0.1%

5.2 Combination -2

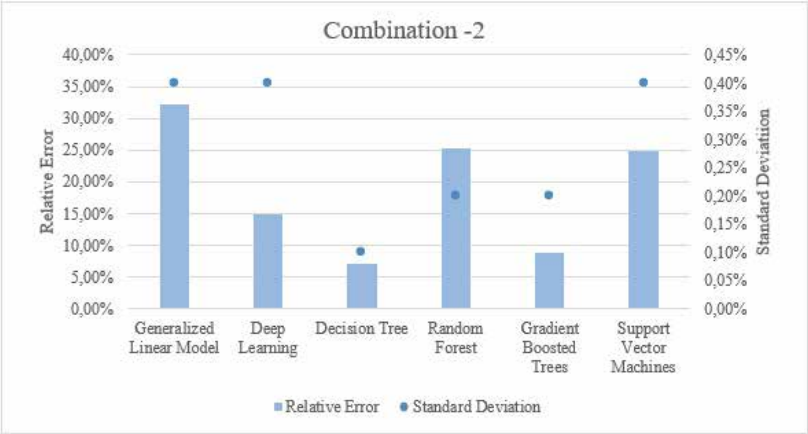


Fig. 2: RapidMiner Studio Output - Model 2

The Most Accurate Tool: Decision Tree
Relative Error: 7.2%
Standard Deviation: ±0.1%

5.3 Combination -3

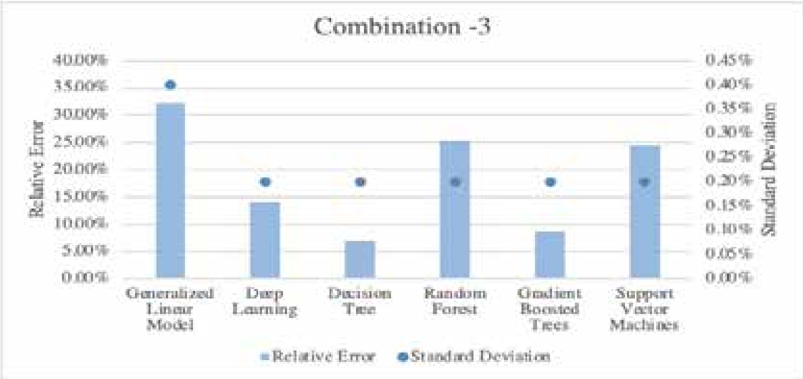


Fig. 3: RapidMiner Studio Output - Model 3

The Most Accurate Tool: Decision Tree
Relative Error: 6.9%
Standard Deviation: ±0.2%

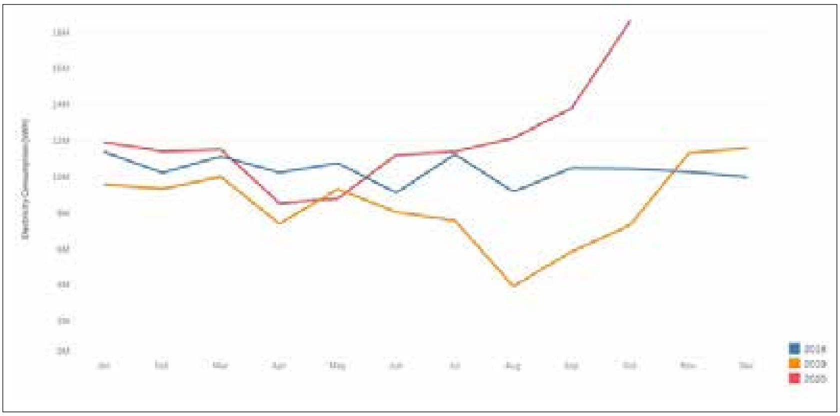


Fig. 4: Monthly Electricity Consumption Values for Three Years

6. Discussion of the Results

Having different data combinations and different prediction tools enables observing the accuracy of prediction under different conditions. The results can be interpreted from two different perspectives: in terms of input types and prediction tools.

The first combination shows that the gradient boosted trees model gives the most accurate prediction with a relative error of 15.4%. This error dramatically decreases to 7.2% with the decision tree model of the second combination. This result shows that wastewater treatment has an important impact on electricity consumption. When the drinking water amount is added as input, the relative error slightly decreases to 6.9% with decision trees, but the standard deviation of the prediction rise by 0.1%

The literature includes studies showing that the deep learning technique outperforms many AI techniques [26], and this study also shows an increase in accuracy when the model was changed to deep learning. However, as mentioned earlier, the accuracy of different AI tools may vary depending on the dataset, and the decision tree (in Combination -2 and -3) and the gradient boosted trees (in Combination -1) outperformed other tools. This output implies that the ultimate tool can be based on one of the two prediction models.

Finally, Fig. 4 shows the electricity consumption trends in the past three years (2018 to 2020). It is visible that the trends in 2018 and 2019 show a similar behavior although the values change. However, after May 2020, when COVID-19-related restrictions are applied in Turkey, the trendline showing electricity consumption behaves differently than the previous years. During this period, industry-based electricity consumption constantly increased. Although the data for November and December 2020 is missing, based on the past trend, a similar rise can be expected.

2. Conclusion

Consumers are the most critical node of an electricity grid since their demand directly defines the amount of electricity that needs to be generated. Therefore, accurate prediction models are crucial for the market where companies generate and distribute electricity. Studies trying to predict such demand should adopt artificial intelligence and machine learning-based techniques since they are adaptive to nonlinear datasets. However, a wide range of prediction techniques and the fact that electricity consumption is affected by several parameters create inaccuracies due to poor adoption of these tools. Predictions can yield worse predictions than traditional techniques due to insufficient knowledge of their types and use cases. Therefore, this study shows the differences between AI-based forecasting tools and the impact of different types of inputs on accuracy. As part of more extensive research that aims to create a tool with the minimum hourly prediction error, this paper exhibits the early findings for selecting the appropriate machine learning technique and datasets. The variables of wastewater treatment and drinking water consumption are considered unique to this research since they lack in similar studies. Although not required, this study incorporated them to understand the relevance between them and electricity consumption. It is believed that this study will shed light on not only its future but also similar studies having a similar purpose.

Acknowledgments

This study is supported by Su Yapi Engineering and Consulancy Inc. The authors would like to thank for the given data, expert opinion, and the fund given to researchers.

References

[1] J. Bedi, D. Toshniwal, Deep learning framework to forecast electricity demand, Appl. Energy. 238 (2019) 1312–1326. https://doi.org/10.1016/j.apenergy.2019.01.113.

[2] A. Kheirkhah, A. Azadeh, M. Saberi, A. Azaron, H. Shakouri, Improved estimation of electricity demand function by using of artificial neural network, principal component analysis and data envelopment analysis, Comput. Ind. Eng. 64 (2013) 425–441. https://doi.org/10.1016/j.cie.2012.09.017.

[3] D.L. Marino, K. Amarasinghe, M. Manic, Building energy load forecasting using Deep Neural Networks, IECON Proc. (Industrial Electron. Conf. (2016) 7046–7051. https://doi.org/10.1109/IECON.2016.7793413.

[4] C. Hamzaçebi, F. Kutay, Yapay Sinir Ağları ile Türkiye Elektrik Enerjisi Tüketiminin 2010 Yılına Kadar Tahmini, J. Fac. Eng. Archit. Gazi Univ. 19 (2004) 227–233.

[5] Electricity Storage, United States Environ. Prot. Agency. (n.d.), Report.

[6] M. Jawad, M.S.A. Nadeem, S.O. Shim, I.R. Khan, A. Shaheen, N. Habib, L. Hussain, W. Aziz, Machine Learning Based Cost Effective Electricity Load Forecasting Model Using Correlated Meteorological Parameters, IEEE Access. 8 (2020) 146847–146864. https://doi.org/10.1109/ACCESS.2020.3014086.

[7] B. Başoğlu, M. Bulut, Kisa dönem elektrik talep tahminleri için yapay sinir ağlari ve uzman sistemler tabanli hibrit sistem geliştirilmesi, J. Fac. Eng. Archit. Gazi Univ. 32 (2017) 575–583. https://doi.org/10.17341/gazimmfd.322184.

[8] T. Vantuch, A.G. Vidal, A.P. Ramallo-Gonzalez, A.F. Skarmeta, S. Misak, Machine learning based electric load forecasting for short and long-term period, IEEE World Forum Internet Things, WF-IoT 2018 - Proc. 2018-Janua (2018) 511–516. https://doi.org/10.1109/WF-IoT.2018.8355123.

[9] E. Boltürk, Elektrik Talebi Tahmininde Kullanılan Yöntemlerin Karşılaştırılması, İstanbul Teknik Üniversitesi, 2013, Master's Thesis.

[10] J. López Prol, S. O, Impact of COVID-19 Measures on Short-Term Electricity Consumption in the Most Affected EU Countries and USA States, IScience. 23 (2020). https://doi.org/10.1016/j.isci.2020.101639.

[11] TEPAV, COVID-19'un Ekonomiye Etkisinin Elektrik Tüketimi İle Takibi, 2020, Report.

[12] I. Ozer, S.B. Efe, H. Ozbay, A combined deep learning application for short term load forecasting, Alexandria Eng. J. 60 (2021) 3807–3818. https://doi.org/10.1016/j.aej.2021.02.050.

[13] C. Xu, W. Li, M. Yu, J. Xu, J. Liu, Y. Wang, L. Zhu, A Correlation Sorting-LSTM Method for High Accuracy Short-term Load Forecasting based on Smart Meter Data, 2020 7th Int. Conf. Information, Cybern. Comput. Soc. Syst. ICCSS 2020. (2020) 755–760. https://doi.org/10.1109/ICCSS52145.2020.9336912.

[14] A. Ullah, K. Haydarov, I.U. Haq, K. Muhammad, S. Rho, M. Lee, S.W. Baik, Deep learning assisted buildings energy consumption profiling using smart meter data, Sensors (Switzerland). 20 (2020) 1–15. https://doi.org/10.3390/s20030873.

[15] S. Ryu, J. Noh, H. Kim, Deep neural network based demand side short term load forecasting, Energies. 10 (2017) 1–20. https://doi.org/10.3390/en10010003.

[16] J.M. Liao, M.J. Chang, L.M. Chang, Prediction of air-conditioning energy consumption in R&D building using multiple machine learning techniques, Energies. 13 (2020). https://doi.org/10.3390/en13071847.

[17] P.R. Jeyaraj, E.R.S. Nadar, Computer-assisted demand-side energy management in residential smart grid employing novel pooling deep learning algorithm, Int. J. Energy Res. 45 (2021) 7961–7973. https://doi.org/10.1002/er.6449.

[18] K. Amarasinghe, D.L. Marino, M. Manic, Deep neural networks for energy load forecasting, IEEE Int. Symp. Ind. Electron. (2017) 1483–1488. https://doi.org/10.1109/ISIE.2017.8001465.

[19] A. Vaghefi, F. Farzan, M.A. Jafari, Modeling industrial loads in non-residential buildings, Appl. Energy. 158 (2015) 378–389. https://doi.org/10.1016/j.apenergy.2015.08.077.

[20] T. Ahmad, H. Chen, R. Huang, G. Yabin, J. Wang, J. Shair, H.M. Azeem Akram, S.A. Hassnain Mohsan, M. Kazim, Supervised based machine learning models for short, medium and long-term energy prediction in distinct building environment, Energy. 158 (2018) 17–32. https://doi.org/10.1016/j.energy.2018.05.169.

[21] A. Galicia, J.F. Torres, F. Martínez-Álvarez, A. Troncoso, A novel spark-based multi-step forecasting algorithm for big data time series, Inf. Sci. (Ny). 467 (2018) 800–818. https://doi.org/10.1016/j.ins.2018.06.010.

[22] A. Bellahsen, H. Dagdougui, Aggregated short-term load forecasting for heterogeneous buildings using machine learning with peak estimation, Energy Build. 237 (2021) 110742. https://doi.org/10.1016/j.enbuild.2021.110742.

[23] H. Duong-Ngoc, H. Nguyen-Thanh, T. Nguyen-Minh, Short term load forcast using deep learning, 2019 Innov. Power Adv. Comput. Technol. i-PACT 2019. (2019) 1–5. https://doi.org/10.1109/i-PACT44901.2019.8960036.

[24] H. Lu, X. Ma, M. Ma, A hybrid multi-objective optimizer-based model for daily electricity demand prediction considering COVID-19, Energy. 219 (2021) 119568. https://doi.org/10.1016/j.energy.2020.119568.

[25] C. Kuster, Y. Rezgui, M. Mourshed, Electrical load forecasting models: A critical systematic review, Sustain. Cities Soc. 35 (2017) 257–270. https://doi.org/10.1016/j.scs.2017.08.009.

[26] S. Atef, A.B. Eltawil, Real-time load consumption prediction and demand response scheme using deep learning in smart grids, 2019 6th Int. Conf. Control. Decis. Inf. Technol. CoDIT 2019. (2019) 1043–1048. https://doi.org/10.1109/CoDIT.2019.8820363.

[27] T. Ahmad, H. Chen, Nonlinear autoregressive and random forest approaches to forecasting electricity load for utility energy management systems, Sustain. Cities Soc. 45 (2019) 460–473. https://doi.org/10.1016/j.scs.2018.12.013.

[28] A.M. Pirbazari, E. Sharma, A. Chakravorty, W. Elmenreich, C. Rong, An Ensemble Approach for Multi-Step Ahead Energy Forecasting of Household Communities, IEEE Access. 9 (2021) 36218–36240. https://doi.org/10.1109/ACCESS.2021.3063066.

[29] R.K. Jain, K.M. Smith, P.J. Culligan, J.E. Taylor, Forecasting energy consumption of multi-family residential buildings using support vector regression: Investigating the impact of temporal and spatial monitoring granularity on performance accuracy, Appl. Energy. 123 (2014) 168–178. https://doi.org/10.1016/j.apenergy.2014.02.057.

[30] D. Demren, Electrical Load Demand Forecasting Application Using Support Vector Machines, Istanbul Technical University, 2011. https://doi.org/10.16194/j.cnki.31-1059/g4.2011.07.016. Master's Thesis.

[31] F. Kaytez, M.C. Taplamacioglu, E. Cam, F. Hardalac, Forecasting electricity consumption: A comparison of regression analysis, neural networks and least squares support vector machines, Int. J. Electr. Power Energy Syst. 67 (2015) 431–438. https://doi.org/10.1016/j.ijepes.2014.12.036.

[32] O. Kaynar, H. Özekicioğlu, F. Demirkoparan, Forecasting of Turkey's Electricity Consumption with Support Vector Regression and Chaotic Particle Swarm Algorithm*, J. Adm. Sci. Cilt. 15 (2017) 211–224.

[33] T. Küçükdeniz, Long Term Electricity Demand Forcesting : An Alternative Approach With Support Vector Machines, 1 (2010) 45–53.

[34] A. Baziar, A. Kavousi-fard, Short Term Load Forecasting Using A Hybrid Model Based On Support Vector Regression, International Journal of Scientific & Technology Research 4 (2015) 189–195.

[35] Y. Chen, P. Xu, Y. Chu, W. Li, Y. Wu, L. Ni, Y. Bao, K. Wang, Short-term electrical load forecasting using the Support Vector Regression (SVR) model to calculate the demand response baseline for office buildings, Appl. Energy. 195 (2017) 659–670. https://doi.org/10.1016/j.apenergy.2017.03.034.

[36] P. Jiang, R. Li, N. Liu, Y. Gao, A novel composite electricity demand forecasting framework by data processing and optimized support vector machine, Appl. Energy. 260 (2020) 114243. https://doi.org/10.1016/j.apenergy.2019.114243.

[37] J. Nagi, K.S. Yap, F. Nagi, S.K. Tiong, S.K. Ahmed, A computational intelligence scheme for the prediction of the daily peak load, Appl. Soft Comput. J. 11 (2011) 4773–4788. https://doi.org/10.1016/j.asoc.2011.07.005.

[38] Y. Fu, Z. Li, H. Zhang, P. Xu, Using Support Vector Machine to Predict Next Day Electricity Load of Public Buildings with Sub-metering Devices, Procedia Eng. 121 (2015) 1016–1022. https://doi.org/10.1016/j.proeng.2015.09.097.

[39] M.K. Azad, S. Uddin, M. Takruri, Support vector regression based electricity peak load forecasting, 11th Int. Symp. Mechatronics Its Appl. ISMA 2018. (2018) 1–5. https://doi.org/10.1109/ISMA.2018.8330143.

[40] K. Kavaklioglu, Modeling and prediction of Turkey's electricity consumption using Support Vector Regression, Appl. Energy. 88 (2011) 368–375. https://doi.org/10.1016/j.apenergy.2010.07.021.

Applying Artificial Neural Networks to Capture, Rationalize and Generalize Nonlinear Response of Structural Elements from Advanced Analytical Simulations and Tests for Performance Based Design

Ahmet Muhtar Çıtıpıtıoğlu, Ph.D. P.E. [1]

[1]Engineering & Design, TAV Construction, Ayazağa Mah., Azerbaycan Cad., Vadistanbul 1B, Sarıyer, İstanbul, citipit@gmail.com

Abstract

Modern seismic structural design codes utilize performance-based procedures to assess new and existing structures. Nonlinear response of each element is defined and assembled to simulate the overall response of the structure under given seismic loads to evaluate against the limit states assigned per performance criteria. For this assessment the nonlinear response of individual components modeled in the structural analysis under varying loading conditions is needed. Detailed nonlinear finite element analysis has been widely available for this purpose, yet the level of detail to simulate the response for each element individually can be cumbersome and impractical for several cases in regular design assessment. ANN are presented as an alternative means to derive accurate nonlinear element response of structural elements from a dataset based on both detailed nonlinear simulations and experimental tests by way of a bolted angle component representative of the range used in practice. It is demonstrated that it is possible to consolidate a wide array of test and analytical data into a practical tool to accurately predict nonlinear response. Although there are many complex configurations of neural networks possible for even more complex problems, it is shown that simple neural network architecture can create very effective predictions of the nonlinear response of bolted steel connections which can be generalized for a wider range of structural components.

Keywords: artificial neural networks (ANN), nonlinear response, seismic design, performance-based design, element response, bolted connections.

1. Introduction

Modern seismic structural design codes utilize performance-based procedures to assess new and existing structures [1,2,3]. The nonlinear response of each element under static or dynamic loads are defined and assembled to simulate the overall response of a structure under given seismic loads. Defined structural elements are then evaluated against their limit states assigned according to the designated performance criteria. There is a need for representative nonlinear response of individual components modeled in structural simulations for performance assessments under varying loading conditions. Design codes provide prescriptive means to model element response or allow alternative testing and analysis to derive realistic, idealized element response. Detailed nonlinear finite element analysis has been widely available for this purpose, yet the level of detail to simulate the response for each element individually can be cumbersome and impractical for regular design assessment.

This study presents an approach to capture, rationalize, and generalize the response of structural elements from a set of detailed nonlinear simulations and tests by way of a bolted angle component using a trained artificial neural network (ANN). The methodology of an ANN is inspired by actual biological nervous systems that consists of several neurons connected to each other. The neuron, also referred to as a node, is the building block of the network which is a simple mathematical function which takes a scalar input and converts it to a desired output. By linking these neurons in layers, the cumulative behavior of a network of neurons can be tailored to take inputs and process them to lead to a desired behavior.

In the scope of Artificial Intelligence (AI), ANN is a machine learning (ML) approach in which algorithms process signals via interconnected nodes. Neural networks are the backbone of deep learning which is a subfield of ML. ML itself is a subfield of AI. ANN is considered “weak” AI. Weak AI is defined by its ability to complete a very specific task, whereas stronger forms of AI will be able to incorporate more human behaviors like intuition. Research on “strong” AI continues [4]. Work on AI is rapidly evolving. Current applications are only scratching the surface of the potential of AI in general for engineering applications.

The application of ANN for several different types of civil and structural engineering problems have been studied since the early 1990s [5,6]. Recent studies continue to demonstrate the effectiveness of ANN’s to predict the overall nonlinear hysteresis behavior structural connections [7] as well as global structural responses under earthquake excitations using less parameters, data and computation utilizing more advanced algorithms [8,9] compared to traditional simulations.

This study is based on the approach presented by Ghaboussi et. al. [10] utilizing the Matlab Neural Network Toolbox [11]. There are several alternative open-source programming libraries available to similarly create and train ANN models for a wide range of engineering problems.

A dataset of nonlinear force-displacement responses of 270 bolted steel angle connection simulations are used to train an ANN which is used to predict the response data excluded from the training dataset. For reference the same data is used to fit a curve function to compare prediction capabilities of the different approaches. The fitted function and ANN are also used to predict the response of experimentally tested bolted angle connection members to truly assess the ANN prediction capability and practical usefulness beyond the simulations used for training. Although there are many complex configurations of neural networks possible for even more complex problems, it is shown that a simple neural network architecture can create very effective predictions of the nonlinear response of bolted steel connections which can be generalized for any structural component.

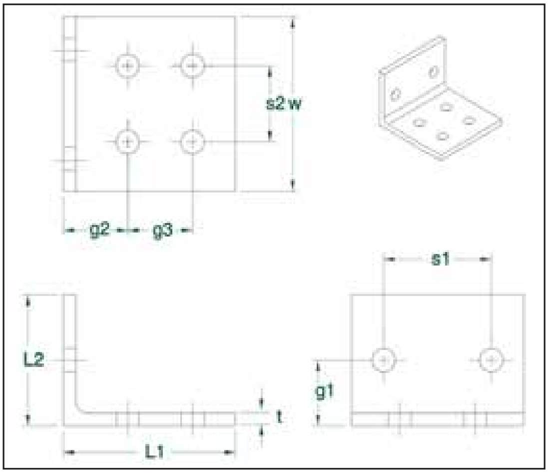
2. Modeling Nonlinear Response of Bolted Angle Connection Components

2.1. Connection Topography and Parametric Variables

The geometrical and topological parameters used to create the data set are explained in this section. Table 1-7 in the AISC LRFD Specification [12] lists all the geometrical properties of angles available for use in structural steel connection design. A rational data set of bolted angle geometry and topography that covers the whole range of parameters used in design practice is selected. Bolted angles are commonly used as connection components to connect the top and bottom flange and web of a beam to the web or flange of a column.

A common bolted angle connection component configuration shown in Fig. 1, is used in this study where two tension bolts connect the angle to a column and four bolts connect the angle to a beam. The angle sizes and bolt gage govern the size of bolts that can be used in a connection base on whether it can be accommodated for the given configuration. Based on the angle profile and bolt size a topography is determined per AISC LRFD Specifications.

The parameters that describe the topography of the bolts on the angle are a function of the size of the bolt and the size of the angle chosen for the connection. Bolt sizes from 1/2 inch (*1.27 cm*) to 1 inch (*2.54 cm*) are used in this study. The full range of each parameter per angle and bolt size is determined and included in the data set. All common angles sizes reasonable for a connection are used in this study. The table in Fig. 1 lists the geometric and topographic parameters of angles in the dataset which are analyzed to creating a large dataset of connection responses.



Angle	Thickness	Bolt Diameter	Tension gage (g1)	Bolt gage (g2)	Bolt gage (g3)	Width (w)	Bolt column spacing (s1)	Bolt beam spacing (s2)	Number of Runs
L8x8	1/2 5/8 7/8	3/4 7/8 1	2.50 4.50 6.75	3	3	6	3.5	3.5	81
						8	5.5 5.5	3.5 3.5	
	1	3/4 7/8 1	2.635 4.500 6.750	3	3	6	3.5	3.5	27
						8	5.5 5.5	3.5 3.5	
L8x6	1/2 3/4 1	3/4 7/8 1	2.500 3.625 4.750	3	3	6	3.5	3.5	81
						8	5.5 5.5	5.5 3.5	
L8x4	1/2 3/4 1	3/4 7/8 1	2.50 2.75	3	3	6	3.5	3.5	54
						8	5.5 5.5	3.5 3.5	
L6x6	5/16 1/2 3/4	1/2 3/4	2.2500 3.5625 4.8750	2.5	2.25	6	3.5	3.5	54
						8	5.5 5.5	5.5 3.5	
	5/16 3/8 1/2	7/8	2.2500 3.5625 4.8750	2.5	2.25	6	3.5	3.5	27
						8	5.5 5.5	5.5 3.5	
L6x4	5/16 1/2 3/4	1/2 3/4	2.2500 2.5625 2.8750	2.5	2.25	6	3.5	3.5	54
						8	5.5 5.5	5.5 3.5	
	5/16 3/8 1/2	7/8	2.2500 2.5625 2.8750	2.5	2.25	6	3.5	3.5	27
						8	5.5 5.5	5.5 3.5	

Fig. 1. Bolted angle connection topography and design set parameters. (*inch units*)

2.2. Construction of Parametric Data Set Variables

The refined 3D finite element modeling for the bolted angle connections shown in Fig. 2 follows the principals outlined in the [13] where the force response of the connections is obtained by imposing a displacement to the edge of the pull plate connected to the angle.

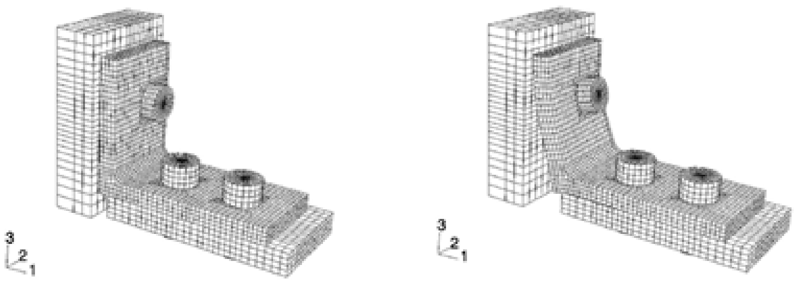


Fig. 2. Typical half symmetry bolted angle mesh model in initial and loaded condition.

2.3. Curve Fitting Based Connection Response Prediction

Previous studies have used various parametric formulas to fit the moment-rotation response of various types of connections. Such fitted equations are usually only reliable in the range of parameters they were fitted with. The bolted angle connection response dataset created using the refined 3D finite element simulations give a considerably wider range of data to work with to fit a parametric function to model and predict connection response.

Following a similar methodology followed by Bahaari and Sherbourne [14], the Richard-Abbott function [15] is used to characterize the force-displacement response of the bolted angle connections as a reference to gage the prediction of the trained ANN. Error in the function predictions compared to the analysis models reveal the inherent limitation of using functions to fit the response of angles. In general, simple functions do not have the flexibility to be representative for the response all the connection cases.

2.4. Neural Networks Based Connection Response Prediction

Although there are many complex configurations of neural networks possible for even more complex problems, it will be shown that simple neural network architecture can create very effective predictions of bolted connection response.

A single neuron, shown in Fig. 3, individually takes on one or more input, **p** which are multiplied by a scalar weight, **w**. All the weighted inputs are summed and a scalar bias, **b** is added to forms the argument, **n** for the transfer function, **f**. The transfer function, also known as an activation function, is usually a step function or a sigmoid which gives the final output,

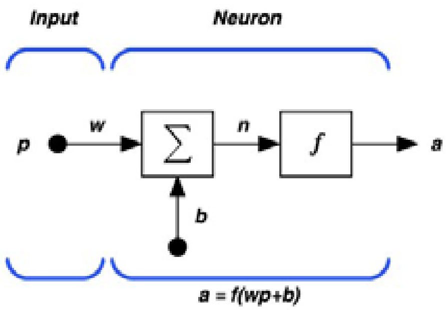


Fig 3. Schematic representation of the mathematical functioning of a neuron.

The weight adjusts the effect of an input to a particular node while the bias adjusts the cumulative input into the transfer function of a particular node. The act of adjusting the weights and bias so the output of a neuron or a network matches the targeted response is referred as training. Usually, the objective of the training is to minimize the mean square error (MSE) between the output and the target data. Once the neural network is trained for a given set of data it is expected to give reasonable output for cases (input) it has not encountered during the training process. The network of neurons can be used to predict the behavior of systems with limited available data. Once the architecture along with the numerical values of the weights and biases are known, the application of a neural network is mathematically straight forward.

There are many training algorithms developed in literature which is beyond the scope of this study, though all these training algorithms use the gradient of the performance function (the mean square error) iteratively to adjust the weights and biases to minimize the performance function. In each iteration the weights and biases are adjusted in the direction of the negative gradient of the performance function. A technique named backpropagation is used where the computations are performed backwards through the network using the derivatives of the network error to determine this gradient. Various optimization methods are implemented to speed up the training process.

The architecture and method of training affects the success of a network to “fit” the data. While having a larger neural network means more computation time for training, more nodes and layers does not particularly mean a better fit. The balance of the number of layers and nodes of the network are important as these affect the overall fitting and prediction performance.

A sigmoid transfer function is commonly used in multilayer networks. The function takes input, **n** from an infinite range and compresses it into a finite output range for **a** of [-1,1]. The expression for the hyperbolic tangent sigmoid function used in this study is shown in Fig. 4.

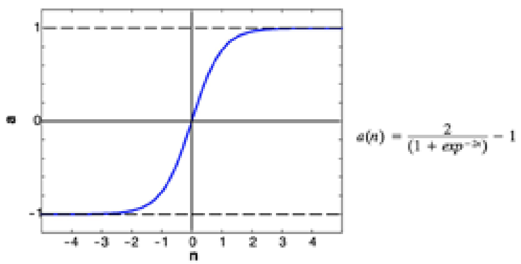


Fig 4. Hyperbolic tangent sigmoid transfer function.

To have the transfer functions work efficiently, all data used in training the network are normalized to the range of [-1,1]. Thus, the weights and biases are tuned to handle normalized input data. The minimum and maximum data values of each parameter are required to make use of the neural network. The simple relationship shown in Eq. (1) is used to preprocess the input parameter, **p** by condensing it between the minimum and maximum value scaled to a range between -1 and 1 while Eq. (2) post process the output data by mapping the output value, **p** to its actual numerical value.

$$p_n = 2\frac{(p - p_{min})}{(p_{max} - p_{min})} - 1$$

$$p_{un} = 0.5(p + 1)(p_{max} - p_{min}) + p_{min}$$

Input parameters used to train the ANN are as follows (shown in Fig. 1): **L1**, the beam side leg length of the angle; **L2**, the column side leg length of the angle; **t**, the thickness of the angle; **db**, the bolt diameter; **w**, the angle width; **g1**, the tension bolt gage; and **d**, the displacement of the connected pull plate bolted to the angle which are mapped to the force response by the neural network.

2.5. Trained Neural Network for Connection Response

The full architecture of the 3 layer 7-7-1 node ANN capable to model the force-displacement of bolted steel angle connection components is shown schematically in Fig. 5. The input is an array of variables which describes the various properties of the angle shown in Eq. (3). The input vector, **p** includes the displacement, **d** in inches as a component for which the NN computes a response force **F** in kips. To create a response curve the displacement, **d** in the input vector is looped from 0 up to 2.5 inches (*6.35 cm*) is inputted to the ANN at a desired increment resulting in each loop giving a single point on the curve. The vectors **pmin** and **pmax** describe the boundaries for the input variables with which the ANN was trained. These vectors are used to scale the input before entering them into the NN using Eq. (1). Similarly, the output **Fn** is scaled with Eq. (2) using the minimum and maximum values given in Eq. (4).

$$p = \begin{Bmatrix} L1 \\ L2 \\ t \\ db \\ w \\ g1 \\ d \end{Bmatrix} \qquad p_{min} = \begin{Bmatrix} 6.0 \\ 4.0 \\ 0.3125 \\ 0.5 \\ 6.0 \\ 2.25 \\ 0 \end{Bmatrix} \qquad p_{max} = \begin{Bmatrix} 8.0 \\ 8.0 \\ 1.0 \\ 1.0 \\ 8.0 \\ 6.75 \\ 2.5 \end{Bmatrix}$$

$$f_{min} = \{ 0 \} \qquad f_{max} = \{ 221.66 \}$$

All the layer weights and bias for the trained ANN are given in the Eqs. (5) to (8). The normalized input vector is multiplied with the input weight matrix, $\mathbf{IW}^{1,1}$ which relates the input to the first hidden layer. The resulting nodal values are summed with their respective bias and processed by the sigmoid transfer function shown in Fig. 4. This set of operation results in the hidden layer vector, $\mathbf{a}^1$. The same operations are repeated by multiplying the $\mathbf{a}^1$ vector with the layer weight matrix $\mathbf{LW}^{2,1}$ which relates the second hidden layer to the first layer. With the same set of operations, the summed nodal values are summed up with their respective bias values and processed by the transfer function resulting in the second hidden layer, $\mathbf{a}^2$. The ANN output is a single node representing the force related to the given displacement, so the layer weights form a vector $\mathbf{LW}^{3,2}$. Summing the result with the bias gives the scaled output. The last node is set so the transfer function is a simple linear relation which requires no function transformation.

The algorithm to compute a single point on the force-displacement using the ANN as described above is presented compactly in Eq. (9). This algorithm can simply be implemented in a code with the angle parameter variables entered and placed in a loop with the displacement variable plotting point by point the nonlinear force response of a bolted angle connection.

2.6. Verification of Model Predictions Using Analysis Data

The response of connection configurations not used for fitting and training are predicted by both the fitted Richard-Abbott function and neural network model and compared to assess their prediction performance. The variables used in the verification cases are given in bold in Table 2 in comparison with the variables used in training the models. In total, 18 connection configurations are simulated and compared to the function and ANN predictions. The mean average of the square root of the summed squares of the residual between the dataset and both the fitted equations and ANN is presented together in Fig. 6. Nonlinear response curves simulated and predicted are compared in Fig. 7. Clearly the ANN response predictions for all 18 cases are more accurate when compared to the parametric fitted function predictions.

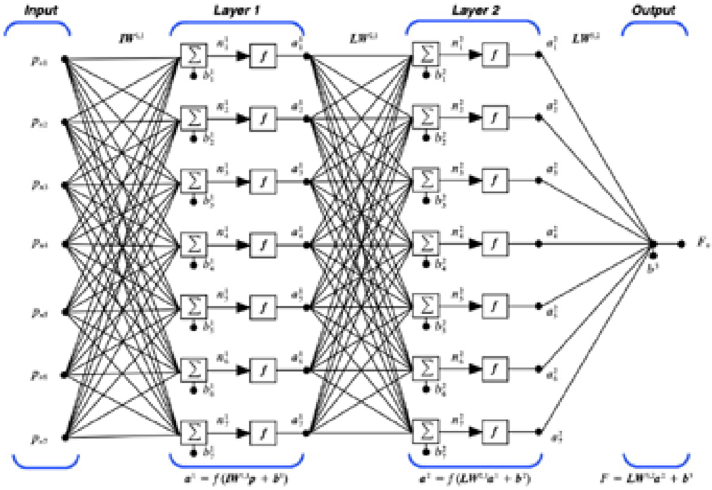


Fig. 5. Schematic representation of trained neural network for simulating the force-displacement response of bolt steel angle components.

$$IW^{1,1} = \begin{bmatrix} -9.2049e-03 & -2.3363e-03 & 3.2209e-01 & 5.7958e-02 & 7.4493e-03 & -8.2541e-01 & -9.4140e-03 \\ 5.7716e-03 & 7.7493e-01 & -1.0662e+00 & 9.0034e-01 & -1.3951e-01 & 1.1172e-01 & -2.7095e-01 \\ -8.9598e-03 & 2.8439e-02 & 4.3011e-01 & 1.7497e-02 & 3.1336e-02 & -6.1448e-01 & -6.2392e-01 \\ 5.5178e-03 & 4.7227e-02 & 3.2548e-01 & -1.3060e-03 & 6.0043e-02 & -3.0730e-01 & 7.4289e-02 \\ 3.5722e-03 & 2.3253e-02 & -2.8827e-02 & -1.1207e-01 & -1.4010e-03 & 1.6256e-03 & -8.9809e-03 \\ -1.1503e-02 & -2.5148e-02 & 1.6036e-01 & -5.7560e-02 & -3.6942e-02 & 3.4671e-01 & -7.2420e-02 \\ -5.1555e-03 & -4.0536e-03 & 1.8636e-01 & 9.7689e-02 & 9.0623e-03 & -2.0649e-01 & -9.8108e+00 \end{bmatrix}$$

$$LW^{2,1} = \begin{bmatrix} 3.4489E+00 & -8.7356E-02 & 3.3919E+01 & 4.9325E+00 & 9.6414E+00 & 4.5368E+00 & -9.1602E+00 \\ -6.1752E-01 & 7.6718E-02 & 2.7598E+00 & -4.6614E+00 & -1.5284E+01 & 1.2800E+00 & -6.0508E+00 \\ 5.7098E+01 & -7.3301E-02 & -6.4352E+01 & 7.3562E+01 & 3.7002E+01 & 2.3694E+01 & 8.6821E+00 \\ 3.9269E-01 & 7.6054E-02 & 1.9058E+00 & -2.8954E+00 & -1.5366E+01 & 1.5236E+00 & -6.5538E+00 \\ -5.5981E+00 & -5.8615E-01 & -2.7616E+00 & -4.9857E+00 & -1.5872E+01 & 2.4126E+00 & 1.7379E-01 \\ -1.3225E+01 & 3.5202E-01 & 1.4254E+01 & -2.3250E+01 & -2.2385E+01 & -3.8846E+00 & -2.3306E+00 \\ -1.6078E+01 & 3.7971E-01 & 1.6711E+01 & -2.7406E+01 & -2.6175E+01 & -4.5362E+00 & -2.5224E+00 \end{bmatrix}$$

$$LW^{3,2} = [\,1.5332E+00 \quad -5.2877E+01 \quad 3.8710E-01 \quad 4.7227E+01 \quad 1.7550E+00 \quad 2.0776E+01 \quad -1.5451E+01\,]$$

$$b^1 = \begin{Bmatrix} -3.0699E+00 \\ 5.8391E-01 \\ -3.5791E+00 \\ -2.8396E+00 \\ 1.8368E+00 \\ 2.2911E+00 \\ -1.1178E+01 \end{Bmatrix} \qquad b^2 = \begin{Bmatrix} 2.0754E+01 \\ 4.4315E+00 \\ 1.6454E+01 \\ 5.6858E+00 \\ -1.2452E-02 \\ 7.5637E-01 \\ 3.1786E-01 \end{Bmatrix} \qquad b^3 = [\,-4.5405E+00\,]$$

$$F_n = LW^{3,2}(f(LW^{2,1}(f(IW^{1,1}p + b^1)) + b^2)) + b^3$$

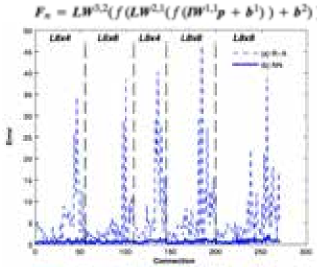


Fig. 6. Comparison of the mean average of the square root of the summed squares of the residual between the simulated bolted angle connections response and the (a) the fitted parametric Richard-Abbot function and (b) the trained ANN predictions.

Table 2. Connection verification data set parameters in bold

Angle	Thickness	Bolt Diameter	Tension gage (g1)	Bolt gage (g2)	Bolt gage (g3)	Width	Bolt column spacing	Bolt beam spacing	Number of Runs
L8x8	1/2 **9/16** 5/8 **3/4** 7/8	3/4 **7/8*** 1	2.50 **3.00** 4.50 **4.00** 6.75	3	3	6	3.5	3.5	4
L8x6	1/2 **5/8** 3/4 **7/8** 1	**3/4*** 7/8 1	2.50 **3.00** 3.625 **4.00** 4.75	3	3	8	5.5	3.5	4
L8x4	1/2 **5/8** 3/4 **7/8** 1	3/4 7/8 **1***	2.50 **2.60** 2.70	3	3	6	3.5	3.5	2
L6x6	5/16 **3/8** 1/2 **5/8** 3/4	1/2 **5/8** 3/4	2.50 **3.00** 3.5625 **4.00** 4.875	2.5	2.25	8	5.5	3.5	4
L6x4	5/16 **3/8** 1/2 **5/8** 3/4	1/2 **5/8** 3/4	2.25 **2.50** 2.5625 **2.75** 2.8750	2.5	2.25	6	3.5	3.5	4

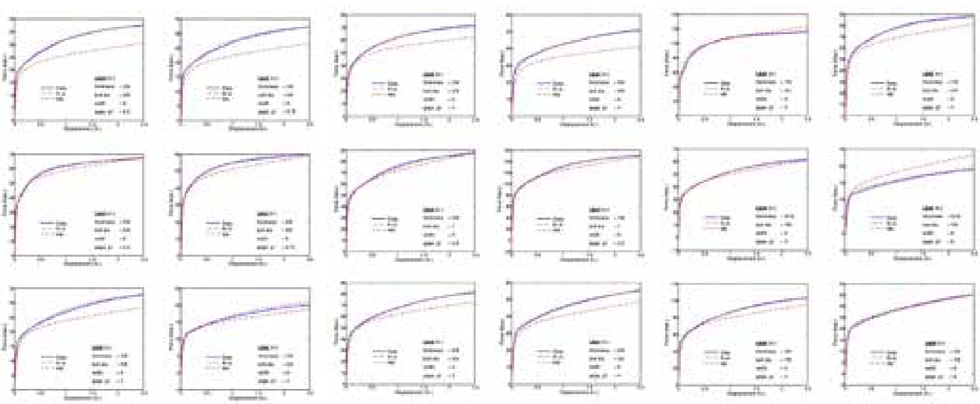


Fig. 7. Nonlinear behavior of bolted steel angle connections generated by FE simulation compared to prediction by the fitted Richard-Abbot (R-A) equation and the trained ANN for cases not used in the fitting and training.

2. 3. 1 Model Prediction of Experimental Data

The parametric fitted Richard-Abbott function and trained ANN are used to predict the response of experimentally tested bolted angle connection members to truly assess their prediction capability and practical usefulness beyond the dataset they are based on. As part of the extensive experimental work done by Swanson [16] on T -stub connections which involved testing individual steel connection components under cyclic loading, thick angle components were also tested as possible alternatives to using T -stubs. The geometry and topography of the tested angles are within the scope of the fitted equation and trained ANN and offers an ideal test bed for their performance in predicting their response. The angles were designed to be connected within a beam-column assembly with a pair of tension bolts on the column side leg and four bolts on the beam side leg as shown in Fig. 8.

All the angles tested were of A572 grade 50 steel and had a thickness of one inch. The beam leg was 8 inches (*20.32 cm*) while the column leg had a varying length of 4 and 6 inches (*10.16 and 15.24 cm, respectively*). Bolt sizes of 7/8 and 1 inch (*2.22 and 2.54 cm, respectively*) were used and based on the diameter bolt spacing on the beam side leg were varied slightly. The geometrical parameters of each test specimen are listed in the table in Fig. 8.

Specimen	Angle	Bolt Diameter	Bolt Grade	Setback(in)	Tension gage (g1)	Gage (g2)	Gage (g3)
CA01	L8x4x1	7/8	A490	1.875	2.5	3.1875	2.625
CA02	L8x6x1	7/8	A490	1.875	2.5	3.1875	2.625
CA04	L8x6x1	7/8	A490	1.875	4.0	3.1875	2.625
CA09	L8x4x1	1	A490	1.5	2.5	3.0	3.0
CA10	L8x6x1	1	A490	1.5	2.5	3.0	3.0
CA12	L8x6x1	1	A490	1.5	4.0	3.0	3.0
CA14	L8x6x1	1	A325	1.5	2.5	3.0	3.0
CA16	L8x6x1	1	A325	1.5	4.0	3.0	3.0

Fig. 8. Bolted angle connection topography and experiment test specimens.

The specimens tested by Swanson were subjected to cyclic axial loads following a protocol consisting several constant cycles of increasing displacement amplitudes. The predicted monotonic response from using the fitted Richard-Abbot (RA) equation and trained ANN is compared to the cyclic response of the tested steel angle specimens in Fig. 9. The neural network successfully predicted the envelope of the cyclic response of the tested bolted angle connection, including bolt slippage seen as the plateau in the response curves.

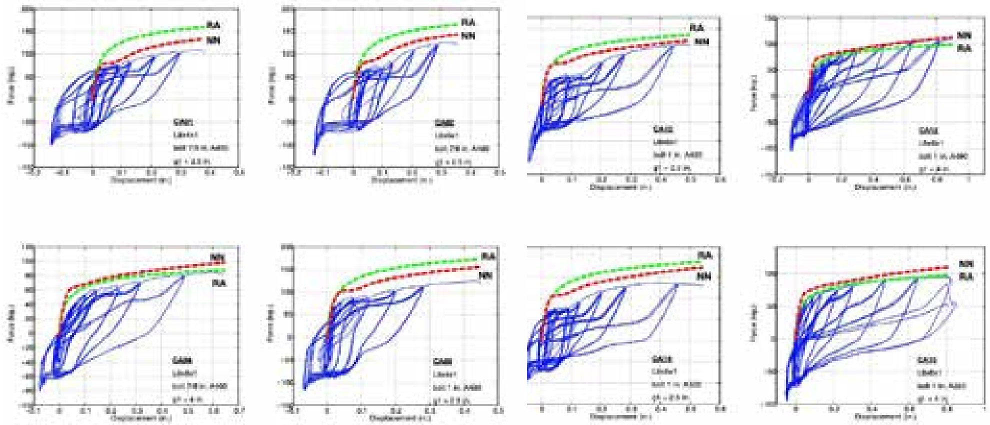


Fig. 9. Fitted Richard-Abbott (RA) equation and trained ANN monotonic response compared to the cyclic response angle component specimens tested by Swanson [16].

3. Conclusion

A large dataset of nonlinear bolted angle connection force-displacement response is generated using parametric refined 3D finite element simulations. The parameters of the bolted angles are representative of the range commonly used in practice. The response dataset is used train an ANN to demonstrate their use to practically predict the response of bolted angle connections. It is demonstrated that it is possible to consolidate a wide array of test and analytical data into a practical tool to predict connection response. Comparing predictions of a fitted function and trained ANN for cases not included in the dataset demonstrate that the trained ANN is particularly successful in predicting the full nonlinear connection response and even capable in predicting the impact of bolt slippage to the overall connection response. ANN predictions envelope the cyclic test response well. ANN is clearly a viable approach for a wider range of structural components.

References

[1] Afet ve Acil Durum Yönetimi Başkanlığı, Türkiye bina deprem yönetmeliği, Resmi Gazete, Sayı 30364, 18 Mart 2018.

[2] ASCE 41-17 2017, Seismic evaluation and retrofit of existing buildings. American Society of Civil Engineers (ASCE), DOI: 10.1061/9780784414859, Reston, Virginia.

[3] Comite Européen de Normalisation, Eurocode 8: Design of structures for earthquake resistance. Part 1: General rules, seismic actions and rules for buildings. CEN, Brussels, EN 1998-1.

[4] E. Kavlakoglu, AI vs. Machine Learning vs. Deep Learning vs. Neural Networks: What's the Difference?, accessed 15 Feb 2021, https://www.ibm.com/cloud/blog/ai-vs-machine-learning-vs-deep-learning-vs-neural-networks.

[5] H. Adeli, "Neural Networks in Civil Engineering: 1989–2000", Computer-Aided Civil and Infrastructure Engineering, DOI:10.1111/0885-9507.00219, Vol. 16, pp. 126–142, 2001.

[6] R. D. Vanluchene, R. Sun, "Neural Networks in Structural Engineering", Neural Netw. Struct. Eng. Microcomp. Civil Eng., DOI: 10.1111/j.1467-8667.1990.tb00377.x, Vol. 5 , p. 207, 1990.

[7] G. Abdollahzadeh, SM. Shabanian, "Experimental and numerical analysis of beam to column joints in steel structures", Front. Struct. Civ. Eng., DOI: 10.1007/s11709-017-0457-z, Vol. 12(4), pp. 642–661, 2018.

[8] T. Kim, OS. Kwon, J. Song, "Response prediction of nonlinear hysteretic systems by deep neural networks", Neural Networks, DOI: 10.1016/j.neunet.2018.12.005, Vol. 111, pp. 1–10, 2019.

[9] SS. Eshkevari, M. Takac, SN. Pakzad, M. Jahani, "DynNet: Physics-based neural architecture design for nonlinear structural response modeling and prediction", Eng. Struc., DOI: 10.1016/j.engstruct.2020.111582, Vol. 229, 2021.

[10] J. Ghaboussi, DA. Pecknold, M. Zhang, RM. Haj-Ali, "Autoprogressive training of neural network constitutive models", Int. J. Number. Meth. Eng., Vol. 42(1), pp. 105-126, 1998.

[11] H. Demuth, M. Beale, Neural Network Toolbox User's Guide, Vol. 4, The MathWorks, Inc., 2005.

[12] AISC 360-2005, Specification for structural steel for buildings, 13th Ed., American Institute of Steel Construction, Chicago.

[13] A. Çıtıpıtıoğlu, "Development and assessment of response and strength models for bolted steel connections using refined nonlinear 3d finite element analysis" Ph.D. Thesis, Georgia Institute of Technology, Atlanta GA., 2009.

[14] MR. Bahaari, AN. Sherbourne, "Finite element prediction of end plate bolted connection behavior. II: Analytic formulation." J. Struc. Engrg., ASCE, Vol. 123(2), pp. 165-175, 1997.

[15] A. Richard, BJ. Abbott, "Versatile elastic-plastic stress-strain formula" J. Engrg. Mech., ASCE, Vol. 101(4), pp. 511-515, 1975.

[16] JA. Swanson, "Characterization of the strength, stiffness, and ductility behavior of t-stub connections" Ph.D. Thesis, Georgia Institute of Technology, Atlanta GA., 1999.

A Comparison of Optimal Force-Based and Performance-Based Designs for A Planar Ordinary Moment-Resisting Steel Frame

Hasan Eser[1], Oğuzhan Hasançebi[2]

[1]Department of Civil Engineering, Middle East Technical University,
Ankara, Turkey,
hasan.eser@metu.edu.tr

[2] Professor, Department of Civil Engineering, Middle East Technical University,
Ankara, Turkey,
oguzhan@metu.edu.tr

Abstract

In past earthquakes, seismic performance of the structures designed in accordance with conventional force-based design (FBD) methodology has been observed to vary significantly from one structure to another. The pursuit of designing structures with foreseeable damage protection under a particular hazard level or a set of hazard levels resulted in the development of performance-based design (PBD) methodology. Since the first introduction of the PBD concepts into structural engineering with the SEAOC Vision 2000, other guidelines such as FEMA 273/356, ASCE 41-06, ASCE 41-13 and ASCE 41-17 are followed; however, the use of these guidelines is limited to the evaluation of the existing structures rather than being complete design codes for new buildings. Nevertheless, the seismic performance criteria in these guidelines have been adopted by many researchers and have found a broad application especially under structural optimization framework. In this study, a planar ordinary moment-resisting steel frame is optimally designed under conventional FBD and PBD methodologies using exponential big bang – big crunch (EBB-BC) algorithm. Then, these optimum designs are compared in terms of design costs and seismic performance. The results suggest that it is possible by PBD approach to produce more economical and safer designs than by the conventional FBD approach.

Keywords: Structural optimization, metaheuristics, performance-based design, steel frames.

1. Introduction

In conventional force-based design (FBD), a structure is expected to endure service loads without showing excessive deformations and vibrations. The ductility of the structural system is ensured via prescriptive capacity design approach and in order to provide economical designs, the structure is allowed to undergo inelastic deformations under seismic action by using strength (force) reduction factors. Although this approach generally produces satisfactory designs, past earthquakes have shown that seismic performance may vary significantly from one structure to another even they are designed under the same code regulations. Over the past few decades, the pursuit of predicting the seismic performance of structures in a reliable manner resulted in the development of several guidelines for the seismic performance evaluation of existing buildings such as FEMA 273, FEMA 356, ASCE 41-06, ASCE 41-13 [1], and ASCE 41-17.

In these guidelines, there are mainly three structural performance levels to classify the post-earthquake state of a structure. Immediate occupancy (IO) performance level is the state in which a structure retains nearly all its pre-earthquake strength and stiffness therefore it is safe to occupy. Life safety (LS) performance level states that some structural members may suffer extensive damage, yet some protection margin is maintained against the risk of partial or total collapse therefore life-threatening injury risk is very low. Collapse prevention (CP) level corresponds to a state where the structure is on the verge of collapse and poses severe risk of injury and life loss. In addition to structural performance levels, there are also performance levels introduced for non-structural components of a structure such as operational, position retention, and life safety.

Performance evaluation of force-controlled and displacement-controlled members are performed separately. For example, force-controlled members are required to satisfy some strength checks and deformation-controlled members are subject to certain plastic deformation limits. Non-structural components are also subject to some strength and drift ratio limitations. Besides, the reliable and precise estimation of these engineering demand parameters necessitates using computationally expensive nonlinear analysis tools such as pushover analysis or nonlinear time history analysis.

Performance criteria in the forementioned guidelines have been adopted and used as design constraints in performance-based design (PBD) optimization of steel moment frames [2,3,4,5,6,7]. In these studies, various inter-story drift ratio (IDR) limits are considered as the only performance criterion for performance-based design. From the structural optimization point of view, it is understandable as it reduces the computational burden associated with numerous analyses and response calculations performed during optimization process such that instead of trying to satisfy performance constraints separately for each structural component in a story, their performance is handled roughly as whole through limiting IDR. On the other hand, although IDR is a good measure of both structural and non-structural damage [8], it does not guarantee that local performance criteria for force-controlled and deformation-controlled members are also satisfied. In fact, it is shown that even under the same inter-story drift ratio, plastic deformation demands can vary considerably from one member to another [9].

Therefore, for a fair comparison between optimum FBD and PBD of a structure, it is crucial to consider both local (e.g., strength, deflection, deformation, etc.) and global (e.g., IDR, etc.) design criteria while evaluating these designs. In this study, a planar ordinary moment-resisting steel frame (OMRSF) is optimally designed under conventional FBD and PBD approaches separately, by using exponential big bang-big crunch algorithm (EBB-BC) [10]. Then, a comparison of these designs is presented in terms of design cost and seismic performance.

2. Problem Formulation

In practice, frame members are selected from a set of commercially available steel sections. For a steel frame consisting of N members collected in N_g member groups for practicality and constructability, the optimization problem for minimum weight design can be formulated as follows:

The objective is to find a vector of design variables I,

$$\boldsymbol{I^T} = \left[I_1, I_2, \ldots, I_{N_g}\right] \boldsymbol{I^T} = \left[I_1, I_2, \ldots, I_{N_g}\right] \boldsymbol{I^T} = \left[I_1, I_2, \ldots, I_{N_g}\right] \quad (1)$$

which minimizes the weight (W) of a structure,

$$W = \sum_{i=1}^{N_g} \gamma_i A_i \sum_{j=1}^{N_m} L_j \quad (2)$$

In Equations (1) and (2), the design variable vector *I* holds the sequence numbers of N_g member groups from the section pool, which is usually constructed from selected AISC standard wide-flange steel sections, g_i and A_i are the unit weight and area of a standard section adopted for a member group *i* respectively, N_m is the total number of members in group *i*, and L_j is the length of the member *j* of the group *i*.

The constraints of optimization problem consist of various requirements imposed by design codes and construction practices. The objective function given by Equation (2) is modified to handle problem constraints via penalty functions where the designs violating these constraints are penalized and their objective functions are calculated using the following equation:

$$\emptyset = W[1 + Penalty(\vec{a})] = W\left[1 + r\left(\sum_{j=1}^{N_c} \max\,(0, g_j)\right)\right] \qquad (3)$$

In Equation (3), ϕ and *W* represents the penalized (constrained) objective function and the unconstrained objective function, respectively; the subscript $j = 1, 2, \ldots, N_c$ denotes the j^{th} normalized design constraint (g_j), where N_c is the total number of constraints on the design. The penalty coefficient *r* is used to adjust the intensity of penalization as a whole. It is usually taken as 1.0.

The design constraints of steel moment frames according to FBD and PBD methodologies are formulated separately in the following subsections.

2.1 Constraints for FBD

In compliance with LRFD methodology in ANSI/AISC 360-10 design code [11], the following strength requirements are considered for the design of the structural steel members:

$$g_1 = \left(\frac{P_r}{P_c}\right)_j + \frac{8}{9}\left(\frac{M_{rx}}{M_{cx}} + \frac{M_{ry}}{M_{cy}}\right)_j - 1.0 \le 0 \quad for \left(\frac{P_r}{P_c}\right)_j \ge 0.2 \qquad (4)$$

$$g_1 = \left(\frac{P_r}{2P_c}\right)_j + \left(\frac{M_{rx}}{M_{cx}} + \frac{M_{ry}}{M_{cy}}\right)_j - 1.0 \le 0 \quad for \left(\frac{P_r}{P_c}\right)_j < 0.2 \qquad (5)$$

$$g_2 = \frac{V_{r,j}}{V_{c,j}} - 1 \le 0 \qquad (6)$$

where Pr, Mr, and Vr are the required axial, flexural, and shear strength demands calculated using LRFD load combinations, respectively; Pc, Mc, and Vc are the axial, flexural, and shear strength capacities determined in accordance with AISC 360-10 LRFD provisions, respectively; x and y are the subscripts which represent the strong and the weak axis of bending for a member, respectively; subscript j = 1, 2, …, Nt denotes the jth member of the structure, where Nt is the total number of structural members. The traditional "1/360" limit in practice is adopted and formulated as a design constraint for the deflection of beams as follows:

$$g_3 = \delta_j - \frac{L_j}{360} = \frac{360\delta_j}{L_j} - 1 \leq 0 \tag{7}$$

where dj and Lj are the calculated deflection in a beam and its corresponding length, respectively; the subscript j = 1, 2, …, Nb denotes the jth member of the structure, where Nb is the total number of beam elements.

In compliance with ASCE 7-10 [12], the inter-story drift is limited as follows:

$$g_4 = \frac{\Delta_i}{\Delta_a} - 1 \leq 0,\ where\ \Delta_i = \frac{C_d \delta_{xe}}{I_e} \tag{8}$$

In Equation (8), D_i and D_a represent the drift of the i^{th} story under seismic action and allowable design story drift, respectively; d_{xe} is the maximum difference between the horizontal displacements of vertically aligned points at the top and bottom of a story along any of the edges of the structure and is determined by the elastic analysis; C_d and I_e are the deflection amplification factor and the importance factor, respectively. The subscript $i = 1, 2, ..., N_s$ denotes the i^{th} story, where N_s is the total number of stories.

In order to ensure the practicality of the resulting design from the constructability point of view, following geometric constraints are considered for the connections between different members (Fig. 1):

$$g_5 = \frac{h^c_{upper}}{h^c_{lower}} - 1 \leq 0 \quad \text{for column} - \text{to} - \text{colum connections} \tag{9}$$

$$g_6 = \frac{b^b_f}{b^c_f} - 1 \leq 0 \quad \text{for beams connecting to column flange} \tag{10}$$

$$g_7 = \frac{b^b_f}{\left(h^c - 2t^c_f\right)} - 1 \leq 0 \text{ for beams connecting to column web} \tag{11}$$

where h, b_f, and t_f are the section depth, flange width, and flange thickness of the doubly symmetric steel section assigned to a structural element, respectively; and the superscripts b and c denote the corresponding beam and column element, respectively.

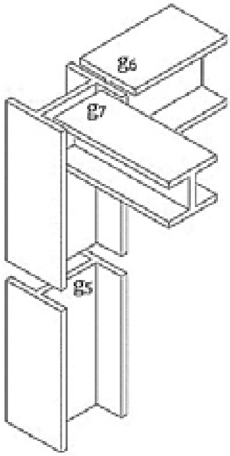


Fig. 1. Geometric (constructability) constraints

The following load combinations are considered for the numerical examples in this study, as specified in ASCE 7-10, Section 2.3:

(1) 1.4D

(2) 1.2D + 1.6L

(3) 1.2D + 1.0L ± 1.0E

(4) 0.9D ± 1.0E

where D, L, and E represent the dead, live, and earthquake loads, respectively. The earthquake loads are calculated and applied in accordance with the equivalent lateral load procedure in ASCE 7-10.

1.2. Constraints for PBD

PBD methodology requires two stages of design checks. In the first stage, it is checked and ensured that all constraints given in the previous subsection for the FBD methodology are satisfied under non-seismic load combinations only. On the other hand, the second stage design constraints are introduced to ensure a satisfactory seismic performance of a structure under a specified hazard level or a set of hazards. Only the second stage design constraints are presented in this subsection to avoid repetition.

In this study, the pushover analysis method with a lateral load pattern based on fundamental mode shape in the direction under consideration is conducted to assess the seismic performance of the structures. The plastic hinges are defined at the start and the end node of each structural element. Modeling parameters and acceptance criteria of the hinges are determined in accordance with ASCE 41-13.

The deformation-controlled members of steel moment-resisting frames are subject to the following constraint:

$$g_1 = \frac{(\theta_p)_j}{(\theta_a)_j} - 1 \leq 0 \tag{12}$$

where q_p is the plastic rotation of a hinge at the target displacement, and q_a is the corresponding allowable plastic rotation for the targeted performance level (IO, LS, or CP); the subscript $j = 1, 2, \ldots, N_h$ denotes the hinge j, where N_h is the total number of hinges defined on the deformation-controlled members.

The members subjected to compressive forces in the excess of 50% of its axial load carrying capacity are designated as force-controlled members, and they must satisfy the following constraint:

$$g_2 = \left(\frac{P_t}{P_{CL}} + \frac{M_{tx}}{M_{CLx}} + \frac{M_{ty}}{M_{CLy}}\right)_j - 1 \leq 0 \tag{13}$$

where P_t and M_t are the axial load and bending moment capacities at the target displacement, respectively; P_{CL} and M_{CL} are the lower bound compressive and flexural strengths, respectively; x and y are the subscripts which represent the strong and the weak axis of bending for a member, respectively; the subscript $j = 1, 2, \ldots, N_f$ denotes the j^{th} force-controlled member of a structure where N_f is the total number of the force-controlled members.

The following constraint on the inter-story drift is also considered for limiting damage to nonstructural components at the target displacement:

$$g_3 = \frac{\Delta_i}{\Delta_a} - 1 \leq 0 \tag{14}$$

In Equation (14), D_i and D_a represent the drift of the i^{th} story and its allowable design story drift, respectively; d_i is the maximum difference between the horizontal displacements of vertically aligned points at the top and bottom of a story at the target displacement. The subscript $i = 1, 2, ..., N_s$ denotes the i^{th} story, where N_s is the total number of stories.

In pushover analysis, the actions caused by seismic forces are considered under the following combination of gravity loads (Q_G), as specified in ASCE 41-13, Section 7.2:

$$Q_G = Q_D + Q_L \tag{15}$$

where Q_D and Q_L are the actions caused by dead loads and live loads, respectively. Q_L is taken equal to 25% of the unreduced live load in accordance with ASCE 7-10 but cannot be taken less than the actual live load.

3. EBB-BC Optimization Algorithm

The steps of exponential big bang-big crunch optimization algorithm as implemented in this study can be summarized as follows:

1. Create a random initial population which is uniformly distributed over the entire design space (population size, μ, is set to 50 in this study). Set iteration number to 1.
2. Calculate the penalized weight of each individual design in the population.
3. Determine the fittest individual based on minimum penalized weight. Record it as the elite design if it is the first iteration or fittest individual is better than the current elite design. Also, record the best feasible design found so far if any.
4. Generate the new population around the elite design (current design) found in the previous step by using the following equation for each design variable, I_i of each candidate solution:

$$I_i^{new} = I_i^c \pm round\left[\alpha \cdot E(\lambda = 1)_i^3 \frac{\left(I_i^{max} - I_i^{min}\right)}{k}\right] \tag{16}$$

In this equation, $E(\lambda)$ is an exponentially distributed random number, I^{max} and I^{min} are the maximum and minimum sequence numbers for steel sections from a section pool, respectively; k is the iteration number, α and is a constant, which is taken as 1.0 in this study.

5. Check bounds on design variables of each individual solution and apply lower and upper bounds when there is a violation.

6. If a generated candidate solution is the same as the current design, regenerate the candidate design such that it is at least one variable different than the current design.

7. Increase the iteration number by 1 and return to Step 2 until a termination criterion is activated. The termination criterion in this study is determined as one of the following conditions, whichever is satisfied first: (a) the maximum number of iterations ($iter^{max} = 1000$) is reached, (b) if no improvement is achieved during a predefined number of iterations ($iter^{ni} = 100$).

4. Numerical example: 54-member planar OMRSF

The example problem is a planar 6-story, 4-span OMRSF consisting of 54 structural elements grouped into 15 independent sizing variables (Fig. 2). All structural members are oriented such that their strong axes coincide with the major axis of bending.

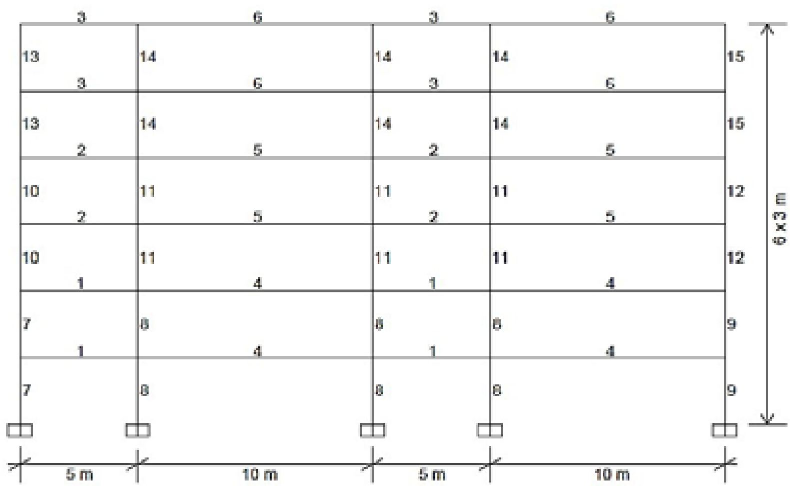


Fig. 2. 54-member planar OMRSF and grouping of structural elements

In order to size the member groups, a section pool consisting of 170 AISC standard wide-flange steel sections is used. These sections are selected from the complete AISC standard wide-flange profile list (297 sections) based on their relatively low cross-sectional area and high moment of inertia properties. The material properties of the steel are set as follows: modulus of elasticity (E) = 200 GPa, yield strength (F_y) = 344.74 MPa, and unit weight (γ) = 76.97 kN/m^3. Design loads and seismic coefficients are provided in Table 1.

Table 1. Design loads and seismic coefficients

Gravity Loads			
Dead Load		24 kN/m	
Live Load		12 kN/m	
Seismic Coefficients (for BSE-1N as 2/3 MCE_R)			
S_s (g)	**S_1 (g)**	**T_L (sec)**	**Site Class**
2.29	0.869	8	D
F_a	**F_v**	**$S_{DS} = 2/3F_aS_s$**	**$S_{D1} = 2/3F_vS_1$**
1	1.5	1.5267	0.869

For seismic performance evaluation of FBD and PBD, the basic performance objective equivalent to new building standards (BPON) given in ASCE 41-13 guidelines is adopted. BPON foresees a Life Safety (LS) structural performance level and Position Retention non-structural performance level under BSE-1N hazard level; and Collapse Prevention (CP) structural performance level under BSE-2N hazard level while non-structural performance is not considered. BSE-2N corresponds to Targeted Maximum Considered Earthquake (MCE_R) and can be thought of as a seismic hazard with 2% probability of exceedance in 50 years. (2%/50-year). BSE-1N, on the other hand, corresponds to design earthquake level and can be taken as 2/3 of MCE_R (ASCE 41-13).

For a sound comparison between optimum PBD and FBD of the frame, 10 independent runs are performed for the optimizations by each design approach. Only the best run results are presented here and used for comparison purposes. In order to accelerate the optimization process, two different approaches have been adopted: (1) Upper bound strategy (UBS) [13] is adopted to provide computational savings by skipping candidate designs (without analyzing) which has no chance to improve the current design since they have higher design weights than the penalized weight of the elite design (2) For PBD optimization runs, in order to make UBS work more efficiently in earlier stages, at the start of first iteration, optimum FBD is adopted as initial elite design after its penalized weight is calculated in accordance with PBD methodology.

The optimization algorithm is coded in MATLAB R2019a, while structural analysis and design of structural models are carried out via SAP2000 v21.02 through open application programming interface (OAPI) and in conjunction with MS-Excel to export hinge results (only for PBD). The optimization runs are performed on a PC with Intel i5-6500, 4-core, 3.2 GHz processor, and 16 GB DDR4 Ram operating at 2133 MHz frequency.

Optimum FBD and PBD are presented in Table 2 with sectional designations attained for each member group. A summary for seismic performance of these designs is provided in Table 3 by tabulating the number of structural members satisfying various structural performance levels under BSE-1N and BSE-2N hazard levels in accordance with the BPON requirements.

Table 2. Optimum PBD and FBD for 54-member planar OMRSF

Group	PBD	FBD	Group	PBD	FBD
1	W16x31	W24x68	9	W27x102	W30x90
2	W21x50	W24x55	10	W18x50	W16x45
3	W8x24	W10x26	11	W27x84	W27x84
4	W27x84	W27x84	12	W27x84	W24x84
5	W24x76	W24x84	13	W14x30	W8x24
6	W24x84	W24x84	14	W27x84	W24x84
7	W18x55	W24x84	15	W24x83	W24x84
8	W30x116	W27x94	**W (kN)**	**283.95**	**296.03**

Table 3. Seismic performance summary

Hazard	BSE-1N		BSE-2N	
Design / P. Level	PBD	FBD	PBD	FBD
IO	46	45	32	28
LS	8	6	19	22
CP	0	0	3	0
> CP	0	3	0	4
IDR_{max} (%)	2.00	1.83	-	-

The results show that optimal PBD produced for this frame, while satisfying all BPON requirements, is about 4.08% lighter than its optimal FBD counterpart, which fails to satisfy BPON requirements as multiple structural members deforms beyond CP performance level at target displacements corresponding to BSE-1N and BSE-2N hazard levels. On the other hand, it is observed that both designs are satisfying IDR constraint, which is limited to 2% for position retention nonstructural performance level. The unsatisfactory seismic performance of FBD may be attributed to the fact that the same strength reduction factor (R) is applied to all members based on the rough assumption that all members will yield simultaneously [14]. This may result in conservative sectional designations for some members, and unconservative for some others. This is also reflected by pushover capacity curves (Fig. 3). Despite having higher base shear capacity than PBD (and similar ductility), FBD demonstrates inferior seismic performance.

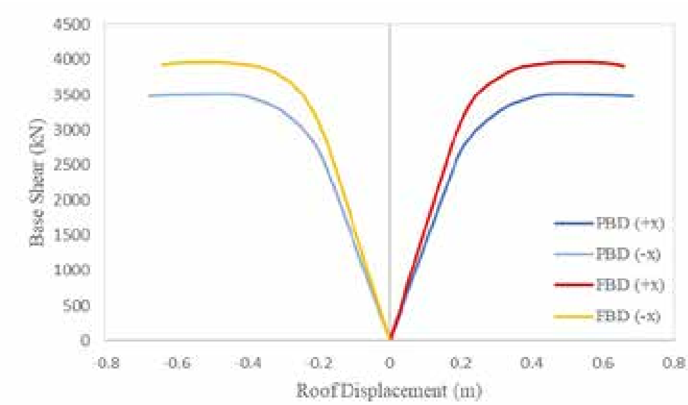


Fig. 3. Pushover curves for optimum PBD and FBD

It is important to note that to obtain corresponding optimum designs even for such a small frame, the total number of structural analyses required is on the scale of several thousands (Fig. 4) despite the strategies adopted in this study for increasing computational efficiency of the applied metaheuristic algorithm. Thus, it is a serious challenge, especially for PBD optimization problems, solution of which requires utilization of time-consuming nonlinear analysis methods throughout the optimization process. Metaheuristics would be clearly impractical for solving PBD optimization problems of large structures due to their relatively slow convergence rates.

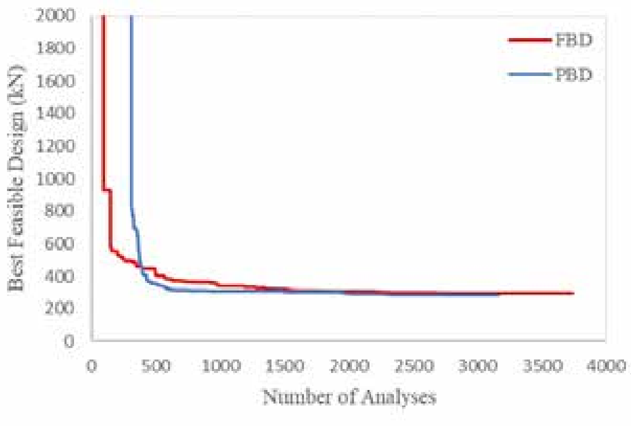


Fig. 4. Optimization histories for PBD and FBD of 54-member planar OMRSF

5. Conclusion

In this study, optimal designs produced for a planar ordinary moment frame under PBD and FBD methodologies are compared in terms of design weight and seismic performance. The results suggest that it is possible by PBD approach to produce more economical and inherently safer designs since seismic performance constraints are imposed in optimization process.

References

[1] ASCE/SEI 41-13, Seismic evaluation and retrofit of existing buildings, American Society of Civil Engineers, 2014.

[2] M. Fragiadakis, N. D. Lagaros, and M. Papadrakakis, "Performance-based multiobjective optimum design of steel structures considering life-cycle cost," Struct. Multidiscip. Optim., vol. 32, no. 1, pp. 1–11, 2006, doi: 10.1007/s00158-006-0009-y.

[3] A. Kaveh, K. Laknejadi, and B. Alinejad, "Performance-based multi-objective optimization of large steel structures," Acta Mech., vol. 223, no. 2, pp. 355–369, 2012, doi: 10.1007/s00707-011-0564-1.

[4] S. Gholizadeh and R. Moghadas, "Performance-based optimum design of steel frames by an improved quantum particle swarm optimization," Adv. Struct. Eng., vol. 17, no. 2, pp. 143–156, 2014, doi: 10.1260/1369-4332.17.2.143.

[5] J. C. Liang, L. J. Li, and J. N. He, "Performance-based multi-objective optimization design for steel structures with intelligence algorithms," Int. J. Optim. Civ. Eng., vol. 5, no. 1, pp. 79–101, 2015, [Online]. Available: http://ijoce.iust.ac.ir/article-1-200-en.pdf.

[6] S. Gholizadeh and A. Milany, "Optimal performance-based design of steel frames using advanced metaheuristics," Asian J. Civ. Engineering, vol. 17, no. 5, pp. 607–623, 2016.

[7] S. O. Degertekin, H. Tutar, and L. Lamberti, "School-based optimization for performance-based optimum seismic design of steel frames," Eng. Comput., no. 0123456789, 2020, doi: 10.1007/s00366-020-00993-1.

[8] FEMA-350, Recommended seismic design criteria for new steel moment-frame buildings, American Society of Civil Engineers, 2000.

[9] T. S. Eom, H. G. Park, and C. H. Lee, "Simplified method for estimation of beam plastic rotation demand in special moment-resisting steel-frame structures," J. Struct. Eng. (United States), vol. 139, no. 11, pp. 1906–1916, 2013, doi: 10.1061/(ASCE)ST.1943-541X.0000768.

[10] O. Hasançebi and S. Kazemzadeh Azad, "An exponential big bang-big crunch algorithm for discrete design optimization of steel frames," Comput. Struct., vol. 110–111, pp. 167–179, 2012, doi: 10.1016/j.compstruc.2012.07.014.

[11] ANSI/AISC 360-10, Specification for structural steel buildings, American Institute of Steel Construction, 2010.

[12] ASCE/SEI 7-10, Minimum design loads for buildings and other structures, American Society of Civil Engineers, 2013.

[13] S. Kazemzadeh Azad, O. Hasançebi, and S. Kazemzadeh Azad, "Upper bound strategy for metaheuristic-based design optimization of steel frames," Adv. Eng. Softw., vol. 57, pp. 19–32, 2013, doi: 10.1016/j.advengsoft.2012.11.016.

[14] M. J. N. Priestley, G. M. Calvi, and M. J. Kowalsky, Displacement-based seismic design of structures, Pavia Italy: IUSS Press, 2007, pp.30.

Artificial Intelligence Applications in Crude Oil Price Forecasting

Fetih KAYA[1], Can BALKAYA[2]

[1]Department of Big Data Analytics and Management, Faculty of Engineering and Natural Sciences, Bahcesehir University, Istanbul, Turkey, fetih.kaya@bahcesehir.edu.tr

[2]Department of Civil Engineering, Faculty of Engineering, Istanbul Aydın University, Istanbul, Turkey, canbalkaya@aydin.edu.tr

Abstract

Because of the crude oil is a common energy source for nearly all commercial sectors, its price forecasting activities have always been an important issue for both governments and commercial firms to make better decisions and investments. In this study, both the history of the crude oil price forecasting and used artificial intelligence methods on forecasting were investigated. In early stages of crude oil price forecasting, traditional statistical and mathematical models were used, while afterwards computer based artificial intelligence models became more popular. Because, these models were more appropriate to the non-linear, volatile and complex structure of oil prices. Artificial intelligence gave chance to evaluate the situation in many aspects at the same time with the help of the computers' power. For instance, news are significant information sources for oil prices and so that text mining can produce sensible outcomes from them to feed models. Evaluation of these produced outcomes together with other variables such as historical prices, weather condition, political situations etc. gave much better forecasting results for crude oil. The development of the artificial intelligence applications and most used artificial intelligence models in crude oil price forecasting were discussed in the study.

Keywords: Artificial intelligence, crude oil price, price forecasting, computer based models.

1. Introduction

Crude oil is a key component of World's energy supply by being a common energy source for almost every commercial sectors. This means, any movement in the oil market can easily affect all other markets too. So, prediction of crude oil prices is an important issue for both governments and commercial firms to make better decisions and investments by knowing its future movements [1].

The forecasting activities of crude oil prices dates back to 1970s, the Yom Kippur War between Arab countries and Israel. Arab oil producers decided to boycott and impose an embargo to America and threat the west in response to support Israel in the War [2]. As a result, crude oil prices began to increase rapidly and economic activities were interrupted. So, the significance of crude oil price prediction was understood and professional prediction activities were triggered to start.

In early stages of oil price forecasting, econometric, statistical and mathematical models were popular due to the absence of machine-made algorithms [2]. Historical price movements, new oil field openings, production rates and economic situations of oil producing countries were the used independent variables of forecasting [2]. In short, time lagged supply and demand balance based approximations were the underlying decision parameters of forecasting for traditional methods. However, these methods were able to solve only linear or near-linear problems with the limited human skills and abilities [3]. In case of the non-linear, volatile and complex structure of oil prices which is sensitive to many exogenous variables such as political decisions, weather conditions etc., the traditional models become ineffective and cannot achieve accurate forecasts [1,3,4].

Development of computer based forecasting methods in early 1990s, helped for generating more complex and accurate prediction models. The forecasting performance of these methods was enough, even more, to compete with traditional methods. In addition, the algorithms perform better in discovering the relationship between every given individual parameter on target variable than traditional methods by conducting many iterations and learning the pattern that cannot be achieved by traditional methods. This computer based algorithms began to substitute traditional algorithms to produce more improved results and "artificial intelligence" (AI) term became popular to express these complex machine-made algorithms in early 2000s. The forecasting accuracies of these computer based artificial intelligence models were much better than traditional models because these models can conduct more complex algorithms for forecasting. In this study, development of artificial intelligence models on crude oil price forecasting and possible future works were discussed.

2. Artificial Intelligence Methods

The fundamental concepts of some computer based artificial intelligence methods used in the scope of crude oil price forecasting are introduced shortly to give a better understanding while following the paper.

2.1. Neural Networks

The structure of artificial neural network (ANN) algorithm is a product of inspiration of human brains neural system. The algorithm basically consists input, hidden and output layers in order respectively, as shown in Fig.1. The input and output layers of the algorithm are filled with known data from a given data set and algorithm generates hidden layer(s) that models the problem with the best way. In short, algorithm tries to learn the followed path according to given input and output data and builds a model for future predictions. The algorithm conducts a deep process to search relations between every given individual input and this makes it successful on forecasting. The number of hidden layers are assigned by the algorithm itself according to the complexity level of the data. In addition, there are many types of neural network such as feed-forward neural networks (FFNN) and generalized-regression neural network (GRNN) which are differentiated from each other based on their processing systems.

2.2. Genetic Algorithms

Genetic algorithms (GA) are searching methods which are based on Darwin's theory of natural selection. The process begins with an initial random population that is thought it has the optimal solutions. After that selection, crossover and mutation steps are applied to this initial population to get a better describing model. During the selection stage, the best performing chromosomes are kept and the others are eliminated from the population. Then, the remaining chromosomes are paired on the crossover stage and adapted on mutation stage. This loop is repeated until the best representing population of the data set is reached. The working principle of the GA is similar to the ANN algorithm. They both go over each input to discover the best explanation of the relationships. The whole process of the GA is represented in Fig.2.

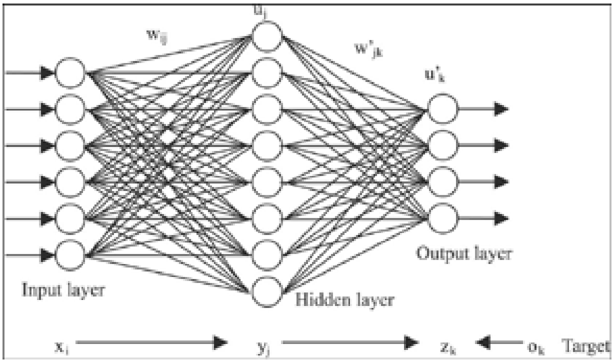


Fig. 1. Neural network structure.

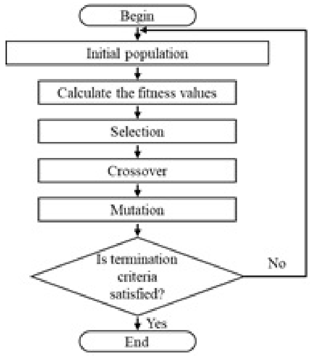


Fig. 2. Genetic algorithm structure.

2.3. Expert Systems

Expert systems are a group of systems that are database, user interface, explanation mechanism, inference engine and knowledge base, as shown in Fig.3. The base knowledge which is crude oil prices for our case, is the most significant part of problem solving. The knowledge is represented in a set of rules which determine conditions and actions to perform according to defined rules. The database holds the facts that paired with the defined rules consisted in knowledge base. The inference engine serves as an interface between the database and knowledge base to provide communication and make intelligently produced results. The explanation mechanism produces expressions for produced results. Expert systems provide a better understanding of data sets by acknowledging people about explanations of taken steps. In short, expert systems produce a set of rules by using knowledge base which is filled with approved information by experts and perform modelling on the given data set to reach the best explaining model of the data set. In addition, this process is conducted interactively with people.

2.4. Text Mining

Text mining is a process of reinterpreting any given text such as news, books etc. The text, first of all, cleaned from any punctuations and divided into words which is called tokenization. Then, the stop words which do not carry important meanings such as "the", "a" etc. are removed from the text. After that, lemmatization which is decreasing the words to its original phase such as "goodness" and "best" turn to "good", is conducted over the text. Finally, these words are vectorised and put in a table for comparison or searching purposes. A representative flow of text mining is illustrated in Fig.4.

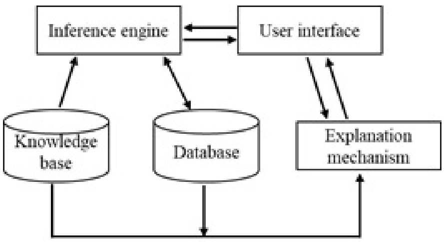


Fig. 3. Expert system structure.

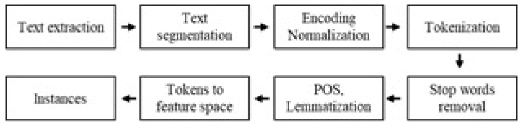


Fig. 4. Text mining process.

2.5. Support Vector Machine

Support vector machine (SVM) is a machine learning algorithm which is used for both classification and regression types of analysis. The algorithm trains a given set of examples and then constructs a prediction model for assigning new coming examples to their categories. The SVM draws lines to differentiate each group from each other by trying to achieve assigning maximum distances between given group examples in space. The working system of SVM is shown and it tries to draw a line which is equally distant from each groups elements in Fig.5.

3. Artificial Intelligence Applications on Forecasting

Application of artificial intelligence algorithms on crude oil price forecasting dates back to early 2000s. In the beginning, individual algorithms were implemented into the models to forecast oil prices. Afterwards, hybrid models, two or more algorithms applied together, were constructed to achieve better predictions. The results of the models showed that; in general, computer based algorithms outperformed over traditional ones for individual algorithm applied models, while hybrid models were the best for forecasting of crude oil prices. Thus, hybrid models have become more and more popular over years and many studies have been conducted with hybrid models.

In 2001, Kaboudan used two artificial intelligence (AI) algorithms individually, genetic programming (GP) and artificial neural network (ANN), to predict crude oil prices for a given period of time. The performance of these two algorithms were compared with random walk algorithm's (RW) results which is a conventional mathematical based algorithm that has been used for a long time. The GP algorithm outperformed over other algorithms while the ANN algorithm had the worst results. The study showed two things actually; AI based algorithms get better result in appropriate conditions, but not always [2]. On the other hand, the price data of crude oil over the period of Gulf War was excluded from the input dataset of the study, because the price volatility was very high and thus hard to integrate a price pattern over that period [2]. This means, AI based algorithms can outperform over conventional methods, but they still suffer from being not able to integrate effects of unexpected events into forecasts.

In the same year with Kaboudan, Rast conducted a hybrid model on oil price forecasting with fuzzy logic and ANN algorithms. The integrated model works in a way that; ANN makes predictions but priorly, fuzzy logic supports the ANN on modelling by giving some predetermined criteria to plot a route such as if weak demand and high inventory situations come together price goes down [3]. In the study, hybrid model was compared with individual

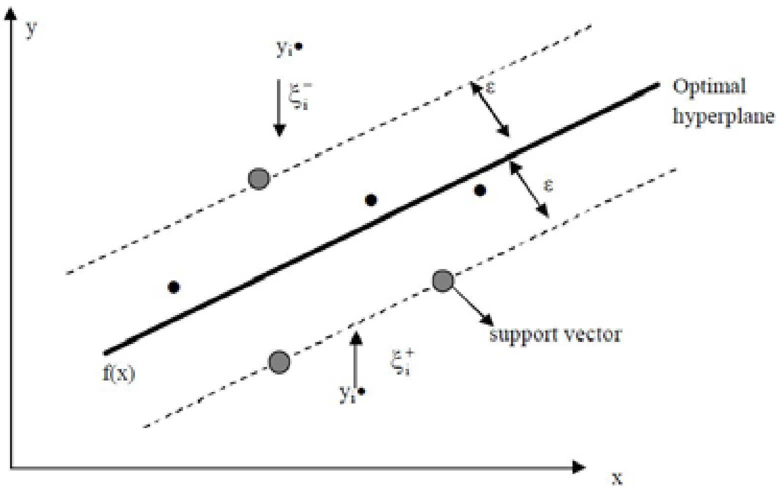


Fig. 5. Support vector machine system [5].

Table 1. Some important AI based studies on crude oil price forecasting with their used AI based methods

	Researchers	Year	Used Artificial Based Methods
1	Kaboudan	2001	Genetic programming and FFNN
2	Rast	2001	Fuzzy logic and FFNN
3	Wang et al.	2004	FFNN, expert systems
4	Yu et al.	2005	Knowledge based systems
5	Fernández	2006	FFNN and support vector machine
6	Yu et al.	2008	FFNN and adaptive linear neural network
7	Mehdi	2009	Fuzzy logic, expert systems and FFNN
8	Khashman and Nwulu	2011	Support vector machine and back-propagation FFNN
9	He et al.	2012	Wavelet analysis
10	Ahmed and Shabri	2014	Support vector machine
11	Yu et al.	2016	Least-squares support vector regression and GA
12	Bon and Isah	2019	Hidden Markov
13	Herrera et al.	2019	ANN and Random forest
14	Wang et al.	2019	Text mining and convolutional neural network
15	Lu et al.	2020	Dynamic-bayesian time series

ANN algorithm. The result; hybrid model outperformed over individual ANN and, the result forms a base for performance enhancement with hybrid models [6].

An important study was conducted by Wang et.al.(2004) to build a model that considers the effects of unexpected events on crude oil prices. The study puts together web-based text mining (WTM), rule-based expert system (RES) and ANN algorithms to generate a hybrid model. The logic behind the study is; searching the crude oil relevant texts in the internet such as news, articles etc. and digging them into to conclude with words that match with oil price movements. The study revealed that the word “war”, for example, affects prices in an upward direction within a range of 5 - 45% and, conversely “OPEC raise production” phrase affects prices downward direction within a range of 5 – 6%. After defining the words or phrases with their effects on oil prices, this information is loaded into RES for processing as a rule over predictions. In final stage, ANN algorithm is applied to historical crude oil dataset by considering determined rules. As it seen in Table 2, the root mean squared error (RMSE) values of hybrid model for each case less than only ANN applied model and also direction change statistics (D_{stat}) are higher. So that, the hybrid system gave better results rather than individual ANN algorithm, especially in the period of many important events which have effect on prices, happened. The text mining algorithm played a significant role to integrate the effects of unexpected events on oil price forecasting [7].

Fernandez (2006) applied support vector machine (SVM) algorithm which is an AI based classification technique, to forecast crude oil prices. She built both individual and hybrid models with SVM, ANN and traditional methods. The comparison of results showed that; apparent difference between applied techniques was originated from time interval. Applied traditional method gave the best prediction result in short term (couple days) while SVM and ANN gave better in long term (up to 20 days). Moreover, the forecasts of linear combination of SVM and ANN were more accurate than individuals. The study shows that, hybrid models perform better, in addition, short term forecasts of traditional models are better than AI based models [8].

Furthermore, Yu et al. (2008) applied a neural network ensemble learning paradigm which is based empirical mode decomposition (EMD), to conclude with a more accurate crude oil price forecast over worldwide [9]. Khashman and Nwulu (2011) conducted a study that compares the back-propagation neural network and SVM for crude oil price forecasting [10]. He et al. (2012) performed a wavelet analysis study to get a better explanation of dynamic movement of oil prices and make precise price forecasting [11]. Ahmed and Shabri (2014) suggested a SVM-based model to predict crude oil prices [12]. Yu et al. (2015) proposed a four-step model; decomposition, reconstruction, prediction and ensemble, on the basis divide and conquer with data-characteristic-driven reconstruction to predict oil prices [13]. Again Yu et al. (2016) proposed a new hybrid model which combines least-squares support vector regression (LSSVR) and a hybrid optimized parameter searching approach with grid method and genetic algorithm (GA) to forecast oil prices [14]. Yu et al. (2017) applied a different approach to its early study in 2016, the model works many times with randomly selected parameters and their average was accepted as actual price forecast [15]. Bon and Isah (2019) developed a hidden Markov-based model to forecast monthly crude oil prices and declared that performance of the model is better than the regular Markov model [16]. Herrera et al. (2019) carried a study by using neural network and random forest algorithms to compare with econometric methods, the result showed that AI based algorithms outperform over traditional ones and they are better at predicting price turning points [17].

Since statistical data such as oil price, supply, demand etc. cannot achieve a perfect prediction accuracy by alone on forecasting of crude oil prices, in 2019 a text-based prediction method was again become a current issue for better predictions. After the study of Wang et al. in 2004, Das and Chan in 2007 and Nguyen, Shirai and Velcin in 2015 have conducted text-based prediction studies too, however, this new one which is conducted by again Wang et al., works on an unaddressed way [18]. The study works on the effects of articles topics on oil prices because they consist more compact information and less words and repetition according to whole article [19]. The model can be summarized like; the headings of articles, news or blogs in online sources which are related with crude oil, are collected and unnecessary parts of them like 'the', 'in', 'on', 'at' etc. are cleaned to get core words. Then, the next day's oil price was compared with today's for assigning these words' effects and coefficient values. For example, if the oil prices decrease by 1.5% after the used words on heading "Iran", "supply" a day before, it is coded that these words together decrease oil prices by 1.5%. This step, of course, is made many times to conclude with a word-effect pair. Latent Dirichlet Allocation (LDA) method is used for grouping the words in the topics according to their effects, increasing or decreasing. After generating a whole set of words and their effects on crude oil prices, convolutional neural network (CNN) algorithm which is a type of ANN algorithm, is used to predict oil prices. The model performance is much better than the conventional methods and also many ANN algorithms used models. In Table 3, comparison of random forest, support vector regression

Table 2. Comparison of ANN method and used hybrid AI method by Wang et. al. in 2004 [7].

Evaluation	Full period (2000-2002)	Sub-period I (2000)	Sub-period II (2001)	Sub-period III (2002)
ANN method				
RMSE	3.413	3.405	3.020	3.324
D_{stat}(%)	61.11	50.00	66.67	66.67
Hybrid AI method				
RMSE	2.369	3.000	2.040	1.916
D_{stat}(%)	80.56	75.00	83.33	91.67

Table 3. Forecasting performances of the random forest, SVR and linear regression models by used hybrid system [18].

	Text features (1)	Financial features (2)	Combination: (1) + (2)	Percentage improvement from (2) to (1) + (2)
MAE				
Random forest	0.0785	0.0082	0.0073	12.32%
SVR	0.0252	0.0032	0.0030	6.67%
Linear regression	0.0854	0.0035	0.0045	-22.22%
RMSE				
Random forest	0.0883	0.0092	0.0088	4.55%
SVR	0.0325	0.0041	0.0040	2.50%
Linear regression	0.0953	0.0044	0.0056	-21.42%
*MAE: mean absolute error; RMSE:root mean squared error; SVR: support vector regression.				

(SVR) and linear regression algorithms were made according to used parameters such as only text parameter, only financial parameters and their combination. The study proves that again, texts that consists news about oil market carry a meaningful information which statistical data do not, about forecasting oil prices [18].

Additionally, Tang et al. (2020) applied a multi-scale model with the data collected from search engines by matching oil price driven factors in search engines [20]. Lu et al. (2020) conducted a dynamic-bayesian structural time series model by using 415 exploratory variables commonly from Google trend search data [21].

4. Conclusion

Since early 2000s, many models have been developed on the crude oil price forecasting by using both AI-based and traditional algorithms. Among them, in general, AI based models outperform over traditional ones, while hybrid models were the best for forecasting of crude oil prices. Fuzzy logic, ANN, GA, SVM, expert systems, text-mining algorithms and their sub-versions were the frequently used AI based algorithms in the models. However, a remarkable amount of models was constructed by using ANN algorithm and its sub-versions like feed-forward neural network (FFNN), recurrent neural network (RNN). Because, the most appropriate algorithm for the non-linear and complex structure of oil prices is ANN with its layered structure which makes it possible to relate many parameters with target variable a detailed way. On the other hand, the sensitive structure of oil prices to exogenous variables such as political, geographical, weather conditions etc. needs a model that can involve wide range parameters in many aspects. To achieve that, text-mining approach on articles, news, reports etc. was commonly used on crude oil price forecasting throughout years.

It is quite obvious that, ANN algorithms has the most appropriate working principle for forecasting the complex and sensitive structure of crude oil prices, because ANN assigns a coefficient number to all parameters according to their predictive power and then combine them all to reach the best. In addition, because of the oil price is very sensitive, usage of text-mining algorithms in order to explore a massive amount of texts in the internet and determine the effects of each word or phrases on oil prices is the most helpful way to feed ANN algorithm. So, changing the type of applied ANN algorithm to its other sub-types or widening the text search extent may help to reach a more developed model. On the other hand, decision tree algorithm could be a possible appropriate way to construct a model for crude oil price forecasting, because its working principle is similar with ANN. For further works, artificial bee colony (ABC) algorithm can be explored, because it is seen that the ABC algorithm have not been taken into account for crude oil price forecasting despite its promising performance. Also, because the hybrid models usually give better results, its combination with other algorithms such as text-mining or the most used one ANN, could improve the prediction results.

References:

[1] Amin-Naseri, M.R., Gharacheh, E.A.: A hybrid artificial intelligence approach to monthly forecasting of crude oil price time series. In: The Proceedings of the 10th International Conference on Engineering Applications of Neural Networks, CEUR-WS284, pp. 160–167 (2007)

[2] Kaboudan M. A. (2001). Compumetric forecasting of crude oil prices. Proceedings of the 2001 Congress on Evolutionary Computation (IEEE Cat. No.01TH8546*)*, Seoul, South Korea, pp. 283-287 vol. 1. Doi: 10.1109/CEC.2001.934402.

[3] Chiroma, H., Abdul-Kereem, S., Shukri, A., Mohd Noor, A.S., Abubakar, A., Safa, N., Shuib, L., Hamza, M., Gital, A., Herawan, T. (2015). A Review on Artificial Intelligence Methodologies for the Forecasting of Crude Oil Price. Intelligent Automation and Soft Computing.

[4] Sehgal, N., Pandey, K.K. (2015). Artificial intelligence methods for oil price forecasting: a review and evaluation. Energy Syst 6, 479–506. Doi: https://doi.org/10.1007/s12667-015-0151-y

[5] Fernández, V. (2006). Forecasting crude oil and natural gas spot prices by classification methods (No. 229).

[6] Rast, M. (2001, July). Fuzzy neural networks for modelling commodity markets. In IFSA World Congress and 20th NAFIPS International Conference, 2001. Joint 9th (Vol. 2, pp. 952–955). IEEE. Doi: 10.1109/NAFIPS.2001.944733.

[7] Wang S., Yu L., Lai K.K. (2005) A Novel Hybrid AI System Framework for Crude Oil Price Forecasting. In: Shi Y., Xu W., Chen Z. (eds) Data Mining and Knowledge Management. CASDMKM 2004. Lecture Notes in Computer Science, vol 3327. Springer, Berlin, Heidelberg. https://doi.org/10.1007/978-3-540-30537-8_26

[8] Fernandez, V (2006). Forecasting commodity prices by classification methods: The cases of crude oil and natural gas spot prices. Documentos de Trabajo from Centro de Economía Aplicada.

[9] Yu, L., Wang, S., & Lai, K. K. (2008). Forecasting crude oil price with an EMD-based neural network ensemble learning paradigm. Energy Economics, 30(5), 2623-2635.

[10] Khashman, A., & Nwulu, N. I. (2011, May). Support vector machines versus back propagation algorithm for oil price prediction. In International Symposium

on Neural Networks (pp. 530-538). Springer, Berlin, Heidelberg.

[11] He, K., Yu, L., & Lai, K. K. (2012). Crude oil price analysis and forecasting using wavelet decomposed ensemble model. Energy, 46(1), 564-574.

[12] Ahmed, R. A., & Shabri, A. B. (2014). Daily crude oil price forecasting model using arima, generalized autoregressive conditional heteroscedastic and support vector machines. American Journal of Applied Sciences, 11(3), 425.

[13] Yu, L., Wang, Z., & Tang, L. (2015). A decomposition–ensemble model with data-characteristic-driven reconstruction for crude oil price forecasting. Applied Energy, 156, 251-267.

[14] Yu, L., Dai, W., Tang, L., & Wu, J. (2016). A hybrid grid-GA-based LSSVR learning paradigm for crude oil price forecasting. Neural computing and applications, 27(8), 2193-2215.

[15] Yu, L., Xu, H., & Tang, L. (2017). LSSVR ensemble learning with uncertain parameters for crude oil price forecasting. Applied Soft Computing, 56, 692-701.

[16] Bon, A. T., & Isah, N. Application of Markov Model and Artificial Neural Network in Crude Oil Price Forecasting for PETRONAS Malaysia.

[17] Herrera, G. P., Constantino, M., Tabak, B. M., Pistori, H., Su, J. J., & Naranpanawa, A. (2019). Long-term forecast of energy commodities price using machine learning. Energy, 179, 214-221.

[18] Li, X., Shang, W., & Wang, S. (2019). Text-based crude oil price forecasting: A deep learning approach. International Journal of Forecasting, 35(4), 1548-1560.

[19] Nassirtoussi, A. K., Aghabozorgi, S., Wah, T. Y., & Ngo, D. C. L. (2015). Text mining of news-headlines for FOREX market prediction: A multi-layer dimension reduction algorithm with semantics and sentiment. Expert Systems with Applications, 42(1), 306–324

[20] Tang, L., Zhang, C., Li, L., & Wang, S. (2020). A multi-scale method for forecasting oil price with multi-factor search engine data. Applied Energy, 257, 114033.

[21] Lu, Q., Li, Y., Chai, J., & Wang, S. (2020). Crude oil price analysis and forecasting: A perspective of “new triangle”. Energy Economics, 104721.

AI And Big Data Applications: Productivity And Crop Yield In Agriculture

İrem Dağdemir[1], Prof.Dr. Zeynep Dilek Heperkan[2], Prof.Dr. Can Balkaya[3]

[1]Faculty of Science, Big Data Analytics and Management,
Bahcesehir University, No:4-634353 Besiktas, Istanbul, Turkey
irem.dagdemir@bahcesehir.edu.tr

[2]Faculty of Engineering, Department of Food Engineering, Aydın University,
Istanbul, Turkey
dilekheperkan@aydin.edu.tr

[3]Faculty of Engineering, Department of Civil Engineering,
Aydın University, Istanbul, Turkey
canbalkaya@aydin.edu.tr

Abstract

The application of Artificial Intelligence (AI) has an obvious effect on agricultural sector recently. According to continuously growing population on earth, farmers need to work harder to meet with the people's needs. Thus, in the agriculture, farmers face numerous challenges to increase productivity and maximize crop yield. Crop diseases, pest and weed infestations, inaccurate soil and crop treatment, inadequate harvest and knowledge gap between farmers and the technology are just some of the difficulties that farmers face. These technologies include big data analytics, robotics, IoT, sensors, cameras, drone technology and GIS applications for large farm areas. With the help of AI applications, decision systems with the integration of real-time monitoring of soil conditions, weathering, watering, and weeds are possible. AI and big data usage results in agricultural processes that meet with flexibility, high performance, accuracy and cost-effectiveness solutions as well as reduces risks in the process and water consumption by using smart irrigation systems. This paper presents how AI applications effect agricultural processes and the ways for utilizing systems for higher productivity.

Keywords: Artificial Intelligence, agriculture, productivity, crop yield.

1. Introduction

Agriculture is one of the most important industries in the world. It provides the most important life source for the living things, “Food”. Beside food, it also helps to provide fabrics, cotton, leathers etc. As the world population grows continuously, the necessity for these products are increasing exponentially. Therefore, the traditional methods which were used by farmers, were not sufficient enough to fulfil these demands. Thus, farmers need to come up with innovative ideas to become more productive to be able to satisfy demands and increase the crop yield. Based on these innovative ideas, AI applications are one of the effective solutions for farmers. AI applications have an important role on numerous operations on the farmed acres and improvement in the agriculture-related tasks in the food supply chain. These technological applications help healthier crops to grow, pest control on the field, monitoring soils, monitoring watering systems, monitoring weather conditions, detection of weeds etc. Besides, there are significant information that can be detect by gathering data and have directly major role in farming. Data can be consist of the information of weather conditions, climate changes, temperature, water usage, soil conditions etc. and this significant information have an important role in making better decisions by using artificial intelligence applications. With the usage of AI applications farmers can plan their land optimally and utilize the resources due to the gathered big data. Also, it increases predictability, therefore it provides traceability and under-control production.

Farming consists of varied steps that require processing and these processes link to various external factors like weather, temperature, climate etc. So, introducing AI technologies to the agriculture will be enabled by other technological advances which will be used in various areas in the agricultural operations. These technologies include big data analytics, robotics, IoT, sensors, cameras, drone technology and GIS applications for large farm areas. So, AI systems can provide predictive analysis for crop type to plant, optimal time for to sow and harvest, determination of irrigation times and use of fertilizers or pesticides etc. Thus, AI applied processes provide farmers easy and efficient farming. Also, AI helps processes become more automated and automation reduces risks in the process. So, the AI provides process flexibility, high performance, accuracy, and cost-effectiveness.

AI-driven technologies are emerging to help improve efficiency and to address challenges facing the industry [1]. This paper presents the applications of AI and big data in agriculture. Focusing on which AI and big data methods are used and how effective they are on productivity and crop yield.

2. Applications

There are various major external factors that are related with farming processes directly. These various external factors which are linked to agriculture contain uncertainties and they are climate, temperature, weather, pollution, weeds, pest attacks, wind speed and direction etc. Also, according to season conditions these factors differs and farmers need to take precautions. However, it may not be easy for farmers to keep track of these changes and take necessary actions on time. This situation can put the productivity and crop quality into danger, this requires traceability and rapidly increasing population and demands prevent this situation being easy for farmers. At this point, AI and big data applications move in and almost at every step of farming which are preparation of soil, sowing of seeds, adding fertilizers, irrigation, weed production and harvesting [2] these technologies can be developed and applied.

2.1. Soil and Crop Health Monitoring System

Soil is one of the most important main part that used in agriculture. It contains nutrients and water inside that help high quality and healthy crops to grow. Therefore, determining the right soil type is a substantial issue that directly affects the crops and if selected correctly, it enhances the crop yield and productivity. For the correct soil type, nutrition levels and quality of soils should be analysed carefully before sowing. Also, after sowing seeds, monitoring the nutrient levels and deficiencies, detecting plant diseases and pests in the soil have a major role for the crops growth. They are the factors that can affect the processes and even can cause crops to die if not intervened in time. In such kind of situations AI-based applications can be applied and the factors which may cause negative effects on crops can be identified easily.

A German-based company called PEAT has developed an AI-based application called Plantix [3]. This application offers a solution to the farmers for agricultural problems according to a photo. So, with this send photo Plantix diagnoses infected crops and offers treatments for any pest, disease or nutrient deficiency problems. It also can get a weekly action plan for farmer based on sown crops and their conditions.

Another company called Trace Genomics helps farmers with analysing their soils by using high-speed, cost-efficient data analysis and machine learning algorithms [4]. According to gathered data from the soil and analysis results of the algorithms, company provides recommendations to the farmers so that they can take the best necessary actions for their field. Also, because company have the data for tracking, farmers can monitor soil and crop's health conditions and help their conditions to be better. Thus, the productivity and crop yield will increase.

Figure 1. Data generated by agricultural drones collected on the field

Data generated by sensors or agricultural drones collected on the field is shown in the Fig. 1. SkySquirrel Technologies Inc. is a Canadian start-up company which brings drone technology. Company develops drone-based imaging technology [5] for monitoring crop health, with a primary focus on improving crop yields and reducing risk of loss from diseases. SkySquirrel Technologies acquired by another company called VineView. The company provides the data insights that optimize the vineyard productivity especially in France and California [6]. A drone is used for capturing the data from the vine yard and gathering data transfers into a computer to be analysed by experts. Captured images and videos are analysed with using algorithms and provides a detailed artificial intelligence report for the farmers that shows general situation of the crops and soils. So, farmers can take necessary actions against the crop diseases, pest infestations etc. Using drones in the agricultural processes makes farmer's work easier and help to save time as the land is a big area to follow.

2.2. Agricultural Robotics and Digital Workforce

AI-enabled companies are developing agricultural robots that can easily perform multiple tasks in farming lands.

In traditional agricultural methods, farms need many seasonal workers to produce and harvest crops. But, having unstable and lack in workforce is a big challenge for farming. AI is a fabulous solution as it reduces the large number of needed workers. Agricultural AI robots can harvest crops at a higher volume and work faster than the human workers [7]. They can also detect the weeds and make the necessary spraying operations for the harmful herbs. Agricultural robots can also check the quality of crops and make packing of the crops at the same time. They are capable to fight with challenges quicker than a human force when an agricultural force labour is faced. Robots and AI-based tools can also detect water usage on the land by controlling irrigation/watering systems. According to data gathered about water usage, smart irrigation systems and robots are developed for all over the land. These systems have knowledge about when to water the crops or which crops should be watered according to the weather conditions. This smart irrigation systems and robots ensure that only required amount of water is used in agricultural processes. Thus, water scarcity is prevented and it helps to save water on the earth.

Figure 2. Autonomous machine performing spraying operation to harmful herbs

Autonomous machine performing spraying operation to prevent the harmful herbs is shown in the Fig.2. Blue River Technology is a company that changes agriculture by creating and iterating on computer vision, machine learning and robotics to create intelligent machinery [8]. The company announced a technology called See&Spray which is an application for spraying operations to detect the weeds and make the necessary spraying operations for the harmful herbs.

Harvest CROO Robotics is a company for harvesting services of strawberry farming [9]. The company has developed a robot to help the challenges in the labour force which is lack of laborers. These robots help farmers to pick and pack their strawberry crops. So, the claim for these robotics is that they are much faster than the human force.

Autonomous vehicles can be very beneficial for the agricultural operations. Multiple various tasks can be made by the autonomous tractors. Self-driving tractors can be programed for different multiple tasks with an avoiding obstacle on the land system like irrigation objects, humans and animals while performing various tasks [10]. The CNH industrial autonomous tractor concept will provide farmers more food more sustainability and better farming [11].

2.3. IoT Powered Data Analytics & Forecasting

In every step of farming, there are a lot of data to gather. The collected data from the technologies like images from drones, analysis of crops, machines used in the farming operations have a huge volume for processing. AI applications help to analyse the data, thus enhance the value derived from these data sources [7]. Therefore, farmers can analyse weather conditions, temperature, climate, soil conditions, water usage etc. and with these information farmers are able to make right decisions, chose the right soil and crop type or decide when to sow and harvest according to season conditions. Also, these technologies help to prevent the water scarcity problem that the world is struggling, with smart irrigation system. As they have the real-time data for crops and soil, smart irrigation systems are able to know when to water or how much water to use. Thus, these systems prevent the waste of water.

2.4. Precision Farming and Predictive Analytics

In the farming field, there are many important factors that determine the quality of harvest and crop yield. These factors need to be constantly controlled by farmers to achieve a productive harvest. Controlling all of these factors is a considerably big difficulty for farmers. Therefore, AI implemented applications and tools are developed to help farmers for controlling the situations on the land and they are guidance to farmers for processes.

Precision farming is a management concept that provides and improves harvest quality by using AI technology. The aim is to define a decision support system for farmers as a guidance which also helps to optimize resources. Farmers need to have information about water management, crop rotation, harvesting time, crop type that should be planted according to climate conditions or soil type, pest attacks, nutrition management, crop diseases etc. Using machine learning algorithms, images that captured by satellites or drones, AI-enabled sensors that are landed all over the farming acre are the several solutions that helps farmers to have this information. These solutions can provide weather condition prediction, analysis of crop conditions, detection of pests, diseases and weeds, controlling nutrition level with the usage of data that is collected from the field. Data can include information about temperature, wind speed and direction, radiation level, pollution etc. FarmShots is a start-up company which deals with crop health and sustainability. With the integration of satellite, aircraft and drone imagery analysis, diseases, pests and poor plant nutrition are detected and farmers can act quickly to prevent any harm to crops [12]. So, the software that company is using can inform farmer when to use fertilizer and thus, they are able to reduce the amount of used fertilizer.

Predictive analytics creates predictive models to help farmers with agricultural process. Agricultural accuracy can be improved with the help of these models. The factors which have serious effects on farming vary season to season. They are like climate change, pollution, temperature etc. Thus, ideal planting and sowing times can change or crop type should be chose accordingly to be able to live in this condition. With the use of predictive models, seasonal forecasting can be made and farmers can spend the season in the most efficient way according to the results of models. Even prediction of months ahead can be provided with the gathered data and farm management can be optimized. Where is a company that delivers weather-based agricultural information to the farmers to enable data-driven decision on adapting to increased weather variability [13]. With the help of machine learning algorithms connected with satellites, company offers accurate agriculture-relevant daily observed weather data which can help farmers to make real-time agricultural decisions. Gathered weather-based agriculture-relevant information is shown in the Fig.3.

Figure 3. Weather based agriculture-relevant information to the farmers

Consequently, farmers need to have information about the season they will face, before starting to plant process and they should determine the path they will follow accordingly. Also, they should be able to implement the applications that they have planned and will be used during the path, easily and at the right time. Therefore, precision farming and predictive analytics are the tools to facilitate and assist farmers.

3. Advantages of AI and Big Data Applications on Agriculture

The use of AI and big data applications in agriculture helps farmers to control current situation on large farming land. As the population is growing, farmers need to meet with increasing demands. Therefore, farmers need to find a way to increase productivity. At this point, AI can potentially change the way we see agriculture, enabling farmers to achieve more results with less effort while bringing many other benefits [14]. These AI and big data applications help to analyse and come up with efficient AI and big data applications provide efficient and easy to apply ways to sow, produce and harvest crops.

- AI and big data applications in agriculture helps to check on defective crops and to improve the healthy crop production.

- AI and big data applications also include automated machines for the processes like weather forecasting and disease or pest identification on the land.

- AI and big data applications can help crop and soil management. According to the crops, the right soil type should be choosing.

- AI and big data applications can solve the problems with the help of gathered data on the land that farmers face during the processes like climate change infestation of pests and weeds that reduces yields.

- Using AI-based applications and tools in agriculture helps to find solutions quickly by monitoring the information. AI-based technologies improve the results with a minimal environmental cost.

- AI and big data applications prevent waste. For example, using smart irrigation systems ensures that only required amount of water is used during the processes. Thus, water scarcity is prevented and water on the earth is saved.

4. Future of AI in Agriculture

For the future AI applications in Agriculture, farm control and management information system with integration monitor system database and GIS and decision system with artificial intelligence and big data for large farm areas for cities or countries can be effectively use in the future. This will increase food productivity and reduce the waste of water as well as automation including smart irrigation systems and robots reduce the number of workers.

As the population increase, climate change and the more we feel the effects of global warming, the importance of water saving will increase. With the smart irrigation systems, AI applications have been helping to save water already, therefore, world may need more help in the future. Additionally, it is a smart choice to avoid from pollution in the long term. As the usage areas increase, it will help people to get insight about economic, environmental, social impacts on activities which will support and increase sustainability.

5. Conclusion

In conclusion, as population is increasing tremendously, requirements for food or any other agricultural products are also increasing. With the help of digital workforce, processes like sowing and harvesting, can be completed much faster and automatically. AI and big data applications in agriculture not only help farmers to automate the farming processes. By going beyond the usual agricultural applications, it also shifts to precise cultivation for higher crop yield and better quality for crops while using less resource. Because, the quality of crops is really important, AI applications reduce the risk of poor quality products which can be occurred by the weeds, unpredicted weather conditions or choosing wrong soil. AI based applications and tools help to improve efficiency in the agricultural processes. It also helps to address the challenges like crop yield, soil health and pest infestation that farmers face during the processes. Because the farming area is very big area, using agricultural robots become a highly valued application of AI in agriculture. Data can be gathered easily with the help of AI and big data. This data can provide daily, weekly, monthly information to the farmers which provide ability to predict changes and identify opportunities as well as warning systems for critical situations. Thus, AI and big data applications in agricultural processes become more efficient and optimized.

References

[1] Ngozi Clara Eli-Chukwu, "Applications of Artificial Intelligence in Agriculture: A Review", Engineering, Technology & Applied Science Research, Alex Ekwueme Federal University Ndufu Alike, Ebonyi, Nigeria, 2019.

[2] www.analyticsvidhya.com/blog/2020/11/artificial-intelligence-in-agriculture-using-modern-day-ai-to-solve-traditional-farming-problems/, Lifecycle of Agriculture, 15 March 2021.

[3] www.plantix.net/en/, Business, Plantix Analytics, 31 March 2021.

[4] www.tracegenomics.com/what-we-do, What We Do, Decode, 31 March 2021.

[5] www.f6s.com/skysquirreltechnologiesinc/, About, 31 March 2021.

[6] www.vineview.com/, About, Who We Are, 31 March 2021.

[7] www.searchenterpriseai.techtarget.com/feature/Agricultural-AI-yields-better-crops-through-data-analytics/, Agricultural Robotics and Digital Workforce, 9 April 2021.

[8] www.bluerivertechnology.com/, Our Methods, 31 March 2021.

[9] www.harvestcroo.com/, About, Why Harvest CROO Robotics? 31 March 2021.

[10] www.medium.com/vsinghbisen/how-ai-can-help-in-agriculture-five-applications-and-use-cases-f09c3dc326c9/, Autonomous Tractors, 31 March 2021

[11] www.cnhindustrial.com/en-us/Pages/homepage.aspx/, Innovation, 31 March 2021.

[12] www.syngenta-us.com/agriedge/technology/, Technology, Integrated Innovation, 9 April 2021.

[13] www.awhere.com/about/, About, 9 April 2021.

[14] www.intellias.com/artificial-intelligence-in-agriculture/, Artificial Intelligence in Agriculture: Rooting Out the Seed of Doubt, 9 April 2021.

Artificial Intelligence Applications in Big Data, Fraud Detection

Mehmet Emre ÖZENGEN[1], Ali OKATAN[2], Can BALKAYA[3]

[1]Student at Big Data Analytics and Management Master Program,
Bahcesehir University, Istanbul,
mehmetemre.ozengen@bahcesehir.edu.tr

[2]Professor at Department of Software Engineering,
Istanbul Aydın University, Istanbul,
aliokatan@aydin.edu.tr

[3]Professor at Department of Civil Engineering,
Istanbul Aydın University, Istanbul,
canbalkaya@aydin.edu.tr

Abstract
Artificial intelligence is used for many purposes nowadays. With the developments in technology, the fraudsters develop their methods. On the other hand, artificial intelligence methods are used in fraud detection for increasing the efficiency of corporations. AI and big data plays an important role in real time data enrichment, deep learning integration and decisions. There are ten artificial intelligence methods explained which are used for fraud detection. Each method has its unique bases and it can not be said that there is only one optimal method. In this research, the methods are briefly explained and a comparison is done for accuracy of methods. Supervised machine learning, unsupervised machine learning or semi-supervised machine learning as well as adaptive machine learning techniques against adaptive attacks with the advantage of big data and artificial intelligence are discussed with effectiveness usage for the future applications.

Keywords: Artificial intelligence, big data, fraud detection, supervised machine learning, unsupervised machine learning, adaptive machine learning.

1. Introduction

Fraud can be defined as illegally obtaining services, goods, or money belonging to other people or organizations and it is one of the greatest challenges for business and organizations. All the systems containing financial transactions are subject to fraud and most of its different forms are determined as a kind of crime in laws. Rapidly increasing volume of e-commerce is attracting the fraudsters equipped with new technologies and the technological defense tools against fraud become more crucial every day. Preventing and detecting fraud is becoming more important more than ever.

Fraud detection is always both related to and fed by data mining and text mining even before the emergence of 'Big Data' phenomenon. However, before the Big Data research techniques developed, there were limited set of ways to develop algorithms to analyze huge amounts of data.

In this research paper, it is aimed to present main ideas of the papers which reviewed the most common and most practiced artificial intelligence techniques of fraud detection. Machine learning models for fraud detection as supervised machine learning models (SMLM), unsupervised machine learning models (UMLM) and semi-supervised machine learning models (SSMLM) as well as adaptive machine learning techniques against adaptive attacks are also discussed with the advantage of using big data and artificial intelligence for the future applications.

2. Fraud Types and Detection Methods

2.1. Fraud Types

Fraud is a broad term that include many different types. Submitting fake documents while applying for a job can be defined as a fraud. On the other extent, making manipulations in the financial tables of a big multinational corporation can also be defined as a fraud. In this research, we determined the limit of fraud as financial frauds. Financial fraud can be determined with the help of two factors: The financial gain and an illegal method implementation. Limiting the framework to financial fraud gives the advantage of using fiscal terms and scales.

Financial fraud has several subsets. There are mainly three industries vulnerable to never ending fraud attempts. The first is banking. The instruments generally attached to fraud cases are listed as the credit cards, the mortgages and complex cash transactions involving money laundering. Fraud can occur in the appliance or distribution phases of credit transactions. The second sector is insurance. The most probable fraud attacks may be on the healthcare and auto insurance instruments. The last industry involving the greatest risk of fraud is the telecommunication industry. There are mainly two areas in that sector. Namely, subscription fraud in which the fraudsters obtain telecommunication accounts without paying and the other is superimposed fraud, in which the legally registered customers pay the fraudsters' expenses.

Detection of fraud is basically a classification problem and if it is not done efficiently, it may be costly for the firms. Because of that, many artificial intelligence techniques aim to increase efficiency in classification [1]. In AI terms, classification can be defined as predicting of a result with the use of inputs. To implement modelling to classify, there must be a training dataset. By benefitting the training dataset, the success of the model can be measured in a test dataset. There are types of classification namely, binary classification, multi-class classification, multi-label classification and imbalanced classification. Artificial intelligence fraud detection techniques generally use imbalanced classification in which most of the training dataset belong to normal values and the minority is labelled as abnormal. The reason of that is in real life, the fraud cases make up a very small proportion of the whole cases.

Early fraud detection works were mainly built on statistical methods such as logistic regressions and neural networks. After than that, data mining techniques were implied. Finally, the hybrid methods are the main way of fraud detection. Therefore, it is very natural that the techniques are evolving and will be improved in the future.

There are mainly two drawbacks to make fraud detection research more challenging for the researchers. The first is fraud detection techniques are mostly specialized for every different companies. The second is that accessing the real-world data is not so easy because of the privacy issues. Therefore, many scholars attempt to make research on different sectors and put weight on comprehensive analysis.

2.2. Artificial Intelligence Techniques Used for Fraud Detection

In this section, the techniques are summarized to maintain a theoretical base.

Bayesian Belief Networks: A Bayesian belief network uses a classifier to calculate for all possible classes and inserts the value X into the class with the highest probability. In this way the network is shown to classify each sample into a class that it is most likely to belong to [2].

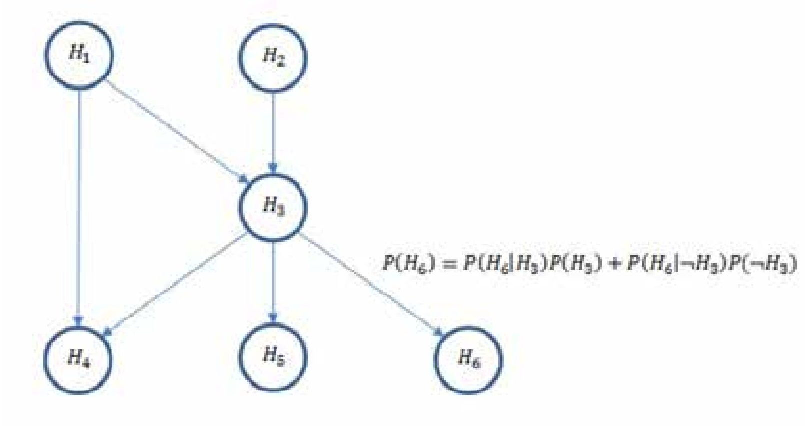


Fig. 1. Bayesian Belief Networks Method [3].

Bayesian Belief Networks: A Bayesian belief network uses a classifier to calculate for all possible classes and inserts the value X into the class with the highest probability. In this way the network is shown to classify each sample into a class that it is most likely to belong to [2].

Logistic Regression: It is statistical method of classifying binary data by using a linear model. It is generally used for predicting of the probability of a case is whether fraudulent or not [4].

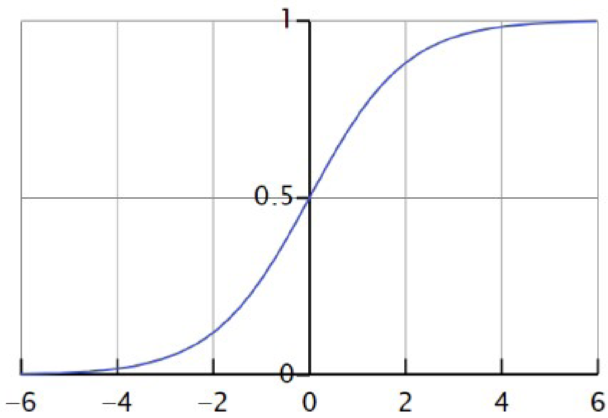


Fig. 2. Logistic Regression Method [5].

Neural Network: An artificial network of neurons or nodes is used in this method. The connections of the neurons are modeled as weights. Positive values of weights represent excitatory connection, and the negative values represent inhibitory connections. After all the weights are aggregated, a function controls the amplitude of the output. This technique is known for its relevance to the predictive modelling which is also used to predict the cases are fraudulent or not.

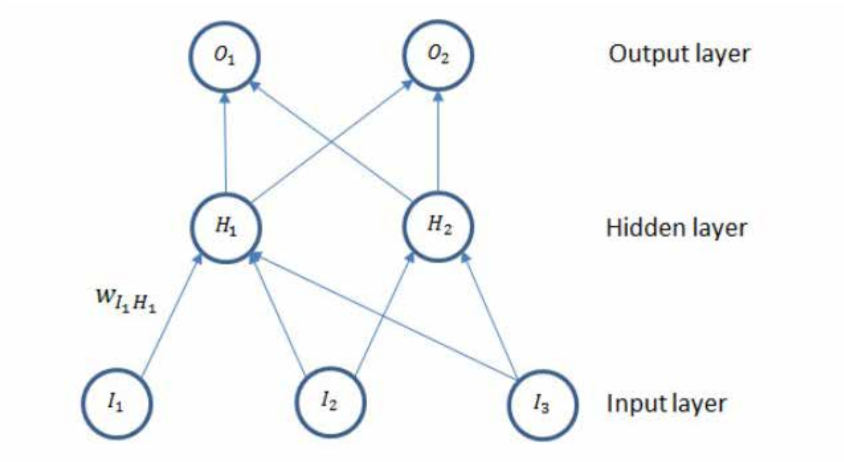


Fig. 3. Neural Network Method [3].

Support Vector Machine: This method is a kind of machine learning algorithms and used for both classification and regression analysis. This method enables complicated non-linear problems to be solved by linear classification without increasing the demand of computational complexity [3].

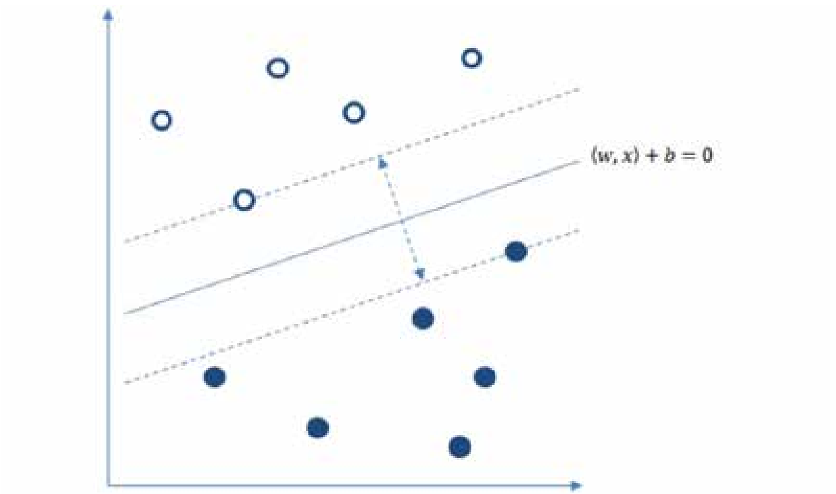


Fig. 4. Support Vector Machine Method [3].

Genetic Algorithms and Programming: This method has the aim of solving problems by evolving an initially random set of possible solutions, through the application of operators inspired by natural genetics and natural selection, such that in time the best solutions would last only. In other words, it is the search of the most optimal program among other algorithms. It selects different parts of programs and produces new generation of programs while combining them. Genetic algorithms are like neural networks in that they require no prior knowledge of the problem domain and are capable of detecting underlying relationships between the samples [6].

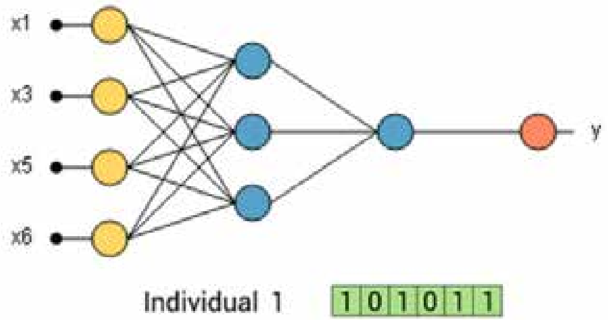


Fig. 5. Genetic Algorithms and Programming Method [7].

Decision Trees, Forest: Decision trees are a technique used to make classifications or predictions on data. It uses a tree with internal nodes representing binary choices on attributes and branches representing the outcome of that choice. The nodes are created by artificial intelligence by using the dataset and it makes decision branches until it is eventually sorted into a mutually exclusive subgroup [2].

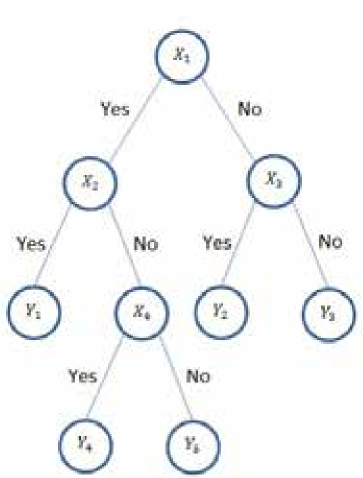


Fig. 6. Decision Tree Method [3].

Group Method of Data Handling: It is a deep learning data mining algorithm that calculating optimal solutions through a series models and increase the accuracy of the model. It has an inductive nature unlike the other deductive methods. It aims to minimize the coders' influence on result of modelling through a set of several algorithms including parametric, clusterization, analogues complexing, rebinarization and probability algorithms. It finds interpretable relations in dataset and selects effective features.

Text Mining: It is a kind of data mining based on plain text data. It filters out the stop words like 'the', 'is', or 'a'. After that, it reduces the derived forms of words into their roots. Finally, it analyzes the data according to the frequencies of words. It basically transforms the text-based data into a quantitative dataset.

Self-Organizing Map: This method is a kind of artificial neural network method which is built on a single matrix of neurons. A high-dimensional data is reduced into a 2-dimensional matrix form. The difference between self-organizing maps and artificial neural networks is that the first applies competitive learning, whereas the latter applies error-correcting learning.

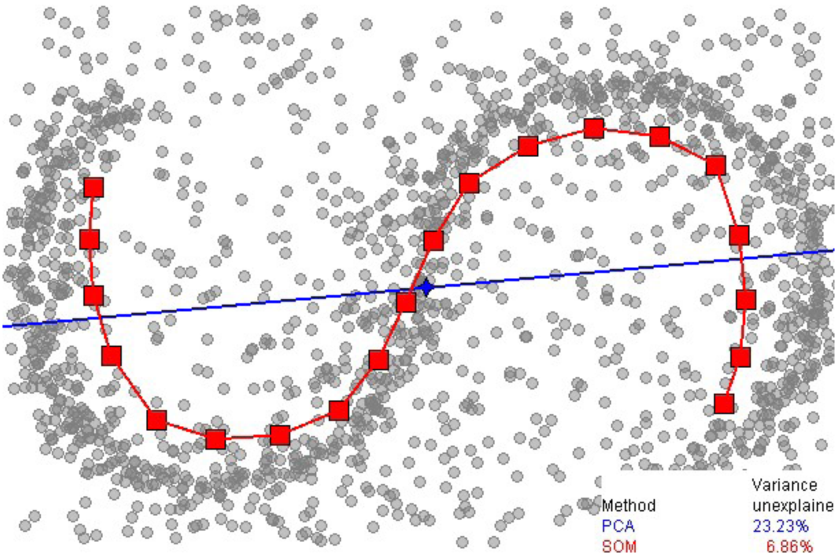


Fig. 7. Self-Organizing Map Method [8].

Process Mining: The goal of process mining is to turn event data into insights and actions. Process mining is an integral part of data science, fueled by the availability of data and the desire to improve processes. Process mining techniques use event data to show what people, machines, and organizations are really doing. Process mining uses these event data to answer a variety of process-related questions. Process mining techniques such as process discovery, conformance checking, model enhancement, and operational support can be used to improve performance and compliance [9].

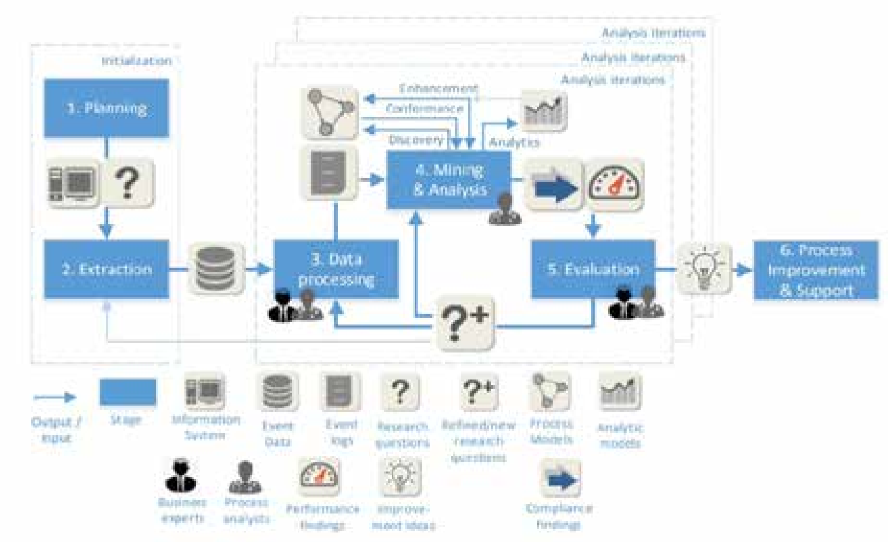


Fig. 8. Process Mining Method [10].

Artificial Immune System: It is a class of artificially intelligent, rule-based machine learning systems inspired by the principles and processes of an immune system of clever creatures like humans. The algorithms of this system are modeled very similarly to an immune system and it was inspired by the concept that learning and memory for use in problem-solving. There are four techniques that can be classified as artificial immune system: Clonal selection algorithm, negative selection algorithm, immune network algorithm, dendritic cell algorithm. The first is used for optimization, the second for anomaly detection, third for clustering and visualization, the fourth for multi-scale processing.

Hybrid Methods: They are combinations of the afore mentioned methods. They can be constructed in a number of ways. The first model's outputs may be applied to another method as an input. One method may be applied as a pre-processing method, while the other makes the essential part. They are constructed to make tailored and specifically targeted solutions.

2.3. Overviewing the Accuracy of Methods

It is researched that the success of methods in different studies. However, the most comprehensive study was done by West et al. and it shows the comparison of different methods [3]. Interpreting the values in the table, the most accurate methods differ from the datasets. In a study of credit card fraud, the accuricies are very similar. However, when inspecting a dataset of financial statements of manufacturing firms, Bayesian belief networks are more accurate than decision trees and neural networks. In other datasets, the best methods differ also.

Table 1. The comparison made on the accuracies of the methods [3].

Fraud Investigated	Method Investigated	Accuracy
Credit card transaction fraud from a real-world example	Logistic model (regression) Support vector machines Random forests	96.6-99.4% 95.5-99.6% 97.8-99.6%
Financial statement fraud from a selection of Greek manufacturing firms	Decision trees Neural networks Bayesian belief networks	73.6% 80% 90.3%
Financial statement fraud with financial items from a selection of public Chinese companies	Support vector machine Genetic programming Neural network (feed forward) Group method of data handling Logistic model (regression) Neural network (probabilistic)	70.41-73.41% 89.27-94.14% 75.32-78.77% 88.14-93.00% 66.86-70.86% 95.64-98.09%
Financial statement fraud with managerial statements for US companies	Text mining	95.65%
Financial statement fraud with managerial statements for US companies	Text mining Text mining and support vector machine hybrid	45.08-75.41% 50.00-81.97%
Financial statement fraud with managerial statements for US companies	Text mining and decision tree hybrid Text mining and Bayesian belief network hybrid Text mining and support vector machine hybrid	67.3% 67.3% 65.8%
Financial statement fraud with financial items from a selection of public Chinese companies	CDA CART Neural network (exhaustive pruning)	71.37% 72.38% 77.14%

3. Machine Learning Models for Fraud Detection

Machine learning can be used with more effectiveness as 1. Supervised Machine Learning Models (SMLM), 2. Unsupervised Machine Learning models (UMLM) and 3. Semi-Supervised Machine Learning Models (SSMLM) against adaptive attacks. Machine learning models need to collect big data to detect fraud. The model analyzes all the input data gathered and extracts the required features. The machine learning model that receives training sets that teach it to predict the probability of fraud. Then, it creates fraud detection machine learning models.

In case of supervised machine learning an algorithm that learns to perform a task from known examples as training data. A supervised learning model is based on predictive data analysis and the accuracy depends on the training set provided for it. SMLM needs large amount of labeled data and has difficulties in detection unknown data. Labeled data to train the models are the important and quantity of data and quality of the data is the biggest limitation in the supervised machine learning. In a supervised learning model, all input information has to be labeled as good or bad.

On the other hand, unsupervised machine learning will be the future of machine learning for detection of unknown attacks. UMLM has an algorithm that learns to identify linkages and patterns in the data without prior knowledge of what to look for and does not require labeled training data using auto-label and auto-rules generation. Generation large set of features, performing correlation analysis, graph analysis to link fraudulent clusters, identifying attack rings and assigning confidence score and categorizing are the basic steps in UMLM. An unsupervised learning model continuously processes and analyzes new data and updates its models. It learns to notice patterns and decide whether they're parts of legitimate or fraudulent operations. Deep learning in fraud detection is usually associated with unsupervised learning algorithms.

Semi-supervised learning models are somewhere between supervised and unsupervised learning models. SSLM works for cases where labeling information is either impossible or too expensive and requires the labor of human experts.

Effectiveness is increased in SMLM and UMLM with the increasing time however with the advantage of big data, computational time and with the advantage of decision systems in artificial intelligence these systems can be used effectively for fraud detections.

Adaptive machine learning techniques can also be effective solution for the analysis of datasets and supervised and unsupervised or even semi supervised machine learning.

4. Conclusion

Fraud detection is a very challenging and important subject to explore for increasing efficiency in some industries hence the number of fraud cases increasing with the technology. In this research, the artificial intelligence methods for fraud detection are categorized and reviewed. Some of the methods have statistical approaches, however some of them have computational approaches. These techniques can be used alone or in a combination. Even though the performances of methods may differ according to datasets, all of them have unique implementations. Every firm can make a hybrid of these methods according to their own needs. Supervised machine learning, unsupervised machine learning or semi-supervised machine learning as well as adaptive machine learning techniques against fraud detection and adaptive attacks with the advantage of big data and artificial intelligence can be used effectively for the future applications.

Acknowledgements

I wish to show my appreciation to my respectable instructer Prof. Dr. Can Balkaya for giving the opportunity and his directions to make a research project with respect to academical research methods and ethical aspects. Secondly, I would like to thank my wife Elif Kevser Özengen for her inexhaustible support while writing this research project paper.

References

[1] E. Duman, and M.H. Ozcelik, "Detecting credit card fraud by genetic algorithm and scatter search.", Expert Systems with Applications 38, 2011, pp 57-63.

[2] E. Kirkos, C. Spathis, and Y. Manolopoulos, "Data mining techniques for the detection of fraudulent financial statements.", Expert Systems with Applications 32, 2007, pp 995-1003.

[3] J. West, and M. Bhattacharya, Intelligent financial fraud detection: a comprehensive review, Computers & Security, 2015, pp 47-66.

[4] E. Ngai, Y. Hu, Y. Wong, Y.Chen, and X. Sun, "The application of data mining techniques in financial fraud detection: A classification framework and an academic review of literature", Decision Support Systems 50, 2011, pp 559-69.

[5] Logistic regression. (2021, January 03). Retrieved January 28, 2021, from https://en.wikipedia.org/wiki/Logistic_regression

[6] P. Ravisankar, V. Ravi, G. Raghava, and I. Bose, Detection of financial statement fraud and feature selection using data mining techniques. Decision Support Systems 50, 2011, pp 491-500.

[7] Genetic algorithms for feature selection. (n.d.). Retrieved January 28, 2021, from https://www.neuraldesigner.com/blog/genetic_algorithms_for_feature_selection

[8] Self-organizing map. (2021, January 03). Retrieved January 28, 2021, from https://en.wikipedia.org/wiki/Self-organizing_map

[9] Van der Aalst, W., Process Mining: Data Science in Action, 2016, p 77.

[10] M. L. van Eck, X. Lu, S.J.J. Leemans, and Wil M.P. van der Aalst, "PM^2: A Process Mining Project Methodology.", Advanced Information Systems Engineering, 2015, pp 297-313.

AI Applications in Smart Cities Waste Management & Recycling

Dr. İ. Haluk ÇERİBAŞI[1], Görkem ÖZDENER[2]

[1]ENCON Environmental Consultancy Co., Ankara, Turkey,
haluk@encon.com.tr

[2]Bahçeşehir University, Big Data Analytics and Management, Beşiktaş, Istanbul, Turkey,
gorkem.ozdener@bahcesehir.edu.tr

Abstract

The rates of urbanization have been very rapid in the last 50 years and by 2050 about two thirds of world population will be living in cities. In the meantime, cities are becoming smart like everything else with the rapid development of technology. These developments are accompanied by environmental pressures on the urban infrastructure. Among these pressures one of the most important ones is generation of wastes. Waste generation rates in urban environment/cities are increasing rapidly and use of technology in waste management, and especially recycling, has to be more closely considered. It is now possible to increase the rate of recycling and better waste management with applications of artificial intelligence (AI) and internet of things. In this article, AI applications in the field of waste management and recycling around the world will be discussed. The proposed model for this purpose applies an algorithm used for solving The Minimum Linear Arrangement Problem which is NP-hard. Some examples of applications used in the waste management sector are examined and conclusions, which could be useful for starting a new transformation about recycling and waste management, are reached.

Keywords: Waste, Waste Management, Recycling, Location Technology, Robotic, Smart City, Artificial Intelligence, Sustainability.

1. Introduction

In this research, artificial intelligence applications applied to waste management and recycling are examined. Any substance or object which the holder discards or intends/required to discard is called waste according to Article 3 of the EU Waste Framework Directive (2008/98/EC). According to this definition waste covers all refuse that could be recycled (referred as waste hereon) and could not be recycled (referred as garbage hereon). The residual material that remains after separating the materials such as paper, cardboard, glass, metal, plastic and which cannot be recycled is called as garbage, which is mostly decomposable food waste or yard waste. In this research, garbage is considered as something that needs to be disposed of, while waste is something that can be reused by separating the substances that make up waste.

Recycling aims to prevent unneeded use of resources and reducing the amount of waste to be disposed by sorting the waste at its source. Recycling and re-use of materials such as iron, steel, copper, lead, paper, plastic, rubber, glass, and electronic will block the depletion of natural resources. This approach will reduce the amount of foreign exchange paid for imported scrap material and save a great deal of energy used by many countries.

Recycling is also an important indicator that shows the development status of nations. While the population and expectations for everyday comforts on the planet are expanding, there is an inescapable expansion in utilization of resources and this puts pressure on our public/common assets (natural resources and environmental quality). Under these conditions, importance of the productive utilization of common assets turns out to be much more obvious. Zero waste policies/legislation and applications have been developed in line with these needs and are on the agenda of Turkey and the world.

2. Some AI Applications in Waste Management

Today, a significant portion of the wastes generated in big cities can be recycled. Recycling and reuse methods that can reduce environmental problems also contribute to the national economies. With the increasing world population, the existence of techniques and models that help people separate recyclable wastes have become imperative for the efficient disposal of waste materials. At this point, Artificial Intelligence technology emerges. If the collection of unwanted materials/refuse is punctual, it enhances a resident's satisfaction. In this model heavy equipment is used by the collecting company's staff for collection of wastes. If waste collection service is satisfactory, residents promote it through the word of mouth [1].

When we examine the artificial intelligence applications that we can address in waste management and recycling, we will see different methods. There are different methods of predicting solid waste generation, which can be broadly classified into five main groups: descriptive statistical models, regression analysis, material flow method, time series analysis and artificial intelligence methods [2].

There are different studies carried out in different regions all over the world by different companies. One of these companies is ZenRobotics. They use robots to separate waste as a recycler/robotic sorting station and the flowchart and characteristics of this station can be seen in Figure 1. In this system, disorganized and mixed data creates a difficult environment for the robot to make predictions and to operate Therefore, specific arrangements have to be made with regard to homogenizing the waste flows. In addition, controls get harder due to external factors such as working environment temperature changes, dust and dirt, so controls on these are also needed.

Another robotic transforming company is AMP Robotics [3]. The working principles of their system can be summarized as follows. They create a database of millions of images. Then, neurons apply a deep learning algorithm to different product groups such as glass, plastic and metal according to their color, size and brand. Furthermore, the algorithm contextualizes the data in order to improve the classification of each material it collects. The collected data also provides transparency to recyclers so that they can optimize their operations and increase recycling rates.

Fig. 1. Robotic Sorting Station [4].

There are two basic approaches in using AI in waste management. The first one is the separation of wastes in different parts of the city at dumping places (trash bins). In this approach, waste is created by households, who dump their waste into a smart trash bin, where the smart trash bin collects waste by separating it. This method stands out with the fact that it reduces the cost and does not require an additional process.

The second approach, on the other end of artificial intelligence applications of waste management, is collecting the garbage as a whole and to separate it by going through a process with the help of robotic automation. This method is more difficult to do and requires a high level of artificial intelligence software and engineering.

Currently a fast development is being experienced in Smart Cities, where engineers, urban planners, architects and city managers are joining forces with the goal of boosting up the efficiency of municipal services and increasing benefits and convenience to their communities [5].

Engineering disciplines, municipal access networks, smart roads provide us some important information to improve the governance of cities. Information is transferred to the relevant units via wireless connection such as smart garbage bins or a prototype with sensor that measures the volume in containers. This data will be used to optimize waste collection and management strategies. AI sensors are a vital improvement over traditional optic option. They can identify the type of material and take into account special considerations. A manufacturer may buy a certain kind of recyclable plastic, for example, but only if it is clean from contaminants. AI sensors can separate out plastics used with chemicals from those that are clean, even if they're the same material [6].

3. Main Waste Management Processes and AI

Waste management processes cover various steps starting from collection of waste and ending by final disposal. In this context a stepwise summary of the process is provided in Table 1 below.

Table 1. Main Processes in Waste Management

Waste Management and Recycling Steps		
Steps	**Details**	**Avg. Time**
Collecting	Separate collection of different types of wastes (e.g. plastic wastes, metals, garbage, etc.) at the source is achieved using various bins/cans such as indoor boxes, cages, containers and piggy banks.	1-3 Days
Classification	This process will allow classification of recyclable waste materials collected separately at the source to be categorized on the basis of glass, metal, plastic and paper. Waste materials under this classification will be ideally delivered to recycling facilities separately.	1 Day
Sorting	The mixed recyclable wastes that arrived with collection vehicles are separated in the separation band. They are separated as paper, metal, glass and plastic wastes. Plastic wastes can be divided into 5 different material types (PET, PE, PP, PS, PVC) and transferred to the recycling industry. Separated wastes form the raw materials for recycling.	6-12 Hours
Pressing	The classified and separated plastic wastes are then turned into bales in the press machine so that they do not take up more space in volume and sent to the storing area.	2-4 Hours
Storing	After the separation process, paper, plastic (including pressed bales), metal and glass wastes are stored separately in the stock area reserved for them and made ready for transfer. They are sent to the recycling facility regularly by the dedicated trucks.	2-3 Weeks

In the context of waste management, main discussion here will be on location-based technologies. In a nutshell, the proposed waste collection system is based on waste level data from trashcans in a metropolitan area. The data collected by sensors is sent over the Internet to a server where it is stored and processed [7].

The new trashcans, which are fully equipped by AI and some physical hardware, have some behavioral property. They have some sensors to determine the distance between top and bottom of the can. Thus, it sends information to server about how full it is. The optimization of these cycles is a combinatorial optimization problem.

When the objective function of this optimization is to minimize the driving distance (equivalent to minimizing the length of the cycles), the problem is the same as the well known The Traveling Salesman Problem and closely related to the Minimum Linear Arrangement Problem, which are NP-hard [9]. The NP algorithm could be followed from Figure 2.

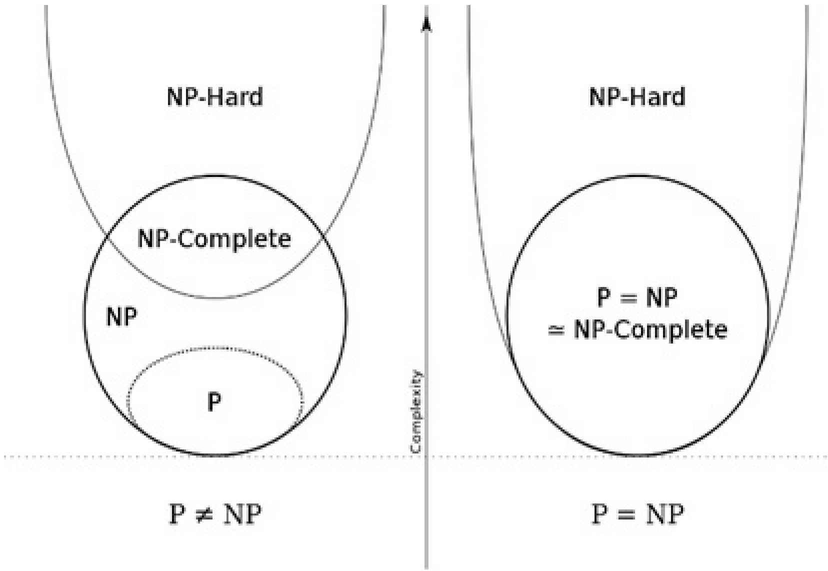


Fig. 2. NP Algorithm Chart [10].

The main point in this part is using collected datasets for optimizing daily routines of trashcans collection. The key feature of this system is that it is designed to learn from experience and to make decisions not only on the daily waste level status, but also on future state forecast, traffic congestion, balanced cost-efficiency functions, and other affecting factors that cannot be foreseen without relevant data [8]. Due to the high number of route optimizations required to carry out the iterations, it was decided to use GA, which are relatively fast in providing near-optimal solutions. We understand that it is of great importance for the waste collection crews, who conduct this duty daily, use these technologies and optimize their work.

Last, but not least there are some Artificial Intelligence applications on waste management in Turkey as well. For example, Antalya Municipality uses smart containers. Sensors in these containers measure the filling rate of each container and instantly transfer the data to related person(s). This system also optimizes for unloading containers so that vehicles can use the shortest distance. Figure 3 shows an intelligent recycling terminal in China, which is a good example of use of AI in waste management and encouraging citizens for recycling.

Fig. 3. Intelligent Recycling Terminals [12].

Some of the best AI applications used for waste management and recycling in the world are summarized in Table 2 below.

Table 2. Best Applications about Waste Management and Recycling [13]

Waste Management and Recycling Applications		
Applications	**Details**	**Operating System**
IRecycle	Earth 911, which is an organization determined to spread the word about the benefit of recycling waste, launched the iRecycle App. The app helps users locate the nearest recycle centers in their vicinity. Also, the app teaches users the different do it yourself (DIY) ways to recycle wastes.	iOS & Android
Gimme 5	There are different types of plastics; some are not recyclable, while some are known to be more profitable when recycled. The plastic type that falls under the latter category is the number 5 plastic which is polypropylene. Gimme 5 is an app that helps users to identify the right recyclable plastics.	iOS
Waste Management App	The Waste Management App is one of the most comprehensive apps for managing waste. It functions as a waste payment app as well as a hub to access local information and policies on waste management. In addition, users can use this app to monitor and track waste pickups.	iOS
RecycleNation	RecycleNation is a mobile app that helps improve recycling rates in the US. Also, the app helps users locate recycling centers on the map as well as their contact information. Furthermore, RecycleNation comes with a news portal that provides users with environmental trends and news.	iOS
My Waste	The My Waste mobile app partners with municipalities all over the world to provide its users up-to-date waste management and recycling news, policies and information. Another feature of the app is the waste collection and recycling dates reminder.	iOS & Android

With the introduction of artificial intelligence applications at every point of our lives, we are stepping into a more livable world by saving money, time and energy on this and other similar practices. Continuity is a solid part of an operation in working up any process. Whatever is done well today can be done better tomorrow. In these environments, collaborative robots equipped with artificial intelligence make purposeful contributions to process improvement from the day they start functioning. With artificial intelligence-based systems that can be integrated into co-conspirator robots, robotic systems do much more than machines for nasty, boring and hazardous jobs.

4. Conclusion

With the rapid development in technology, many innovations are applied in our daily life. While some of these developments do not contribute directly to sustainability of natural resources and development, some of them show great importance for human welfare and sustainable development. Artificial intelligence applications in the field of waste management and recycling are are among the technological developments that serve for environmental sustainability and human welfare in the long run.

The problems related to climate change/global warming and degradation of environmental quality are mainly due to rapid population increase and associated need for economic growth without giving necessary attention to environmental issues. In this context, use of technology was also focused solely on economic development and missed the need for conservation of natural resources and ecosystem. However, with the new millennium more strict policies and targets were put down by nations and international organizations to stop environmental degradation and to achieve better environmental conditions. In addition, awareness of the public has been exponentially increasing through improved communication technology and municipalities are acting more responsibly concerned with environmental impacts of wastes.

In this regard, AI technologies are being more broadly used especially in developed countries in the field of environmental and specifically waste management. In management of solid wastes recycling, reuse and recovery are becoming more and more important and technological developments are supporting better application of these techniques. Thus, improvement of AI applications in this field and their more widely use (with decreasing of associated costs) would significantly contribute to environmental protection and would be key factors in sustainable development.

References

[1] Zeithaml, V.A., 2000. Service quality, profitability, and the economic worth of customers: what we know and what we need to learn. J. Acad. Mark. Sci. 28, 67–85.

[2] Abbasi M, El Hanandeh A (2016) Forecasting municipal solid waste generation using artificial intelligence modeling approaches. Waste Management 56:13–22.

[3] https://zenrobotics.com/ , Online, Web Page

[4] https://wastelessfuture.com/revolutionizing-recycling-pick-by-pick/

[5] Kunzmann K.R., Smart Cities: A New Paradigm of Urban Development. Crios, 1/2014, pp. 9-20, doi: 10.7373/77140

[6] https://datafloq.com/read/how-ai-is-assisting-with-waste-management/7305#:~:text=AI%20sensors%20are%20a%20vital,only%20it's%20clean%20from%20contaminants.

[7] Jose M. Gutierrez, Smart Waste Collection System Based on Location Intelligence, Conference Organized by Missouri University Science and Technology 2015-San Jose, CA. Procedia Computer Science 61 (2015) 120-127

[8] Michael Jensen, Smart Waste Collection System Based on Location Intelligence, Conference Organized by Missouri University Science and Technology 2015-San Jose, CA. Procedia Computer Science 61 (2015) 120-127

[9] Petit, J., Experiments on the minimum linear arrangement problem, Systems Informatics, 2001, vol. 8, pp. 112–128, 2001.

[10] https://commons.wikimedia.org/wiki/File:P_np_np-complete_np-hard.svg

[11] Achillas, C., Aidonis, D., Vlachokostas, C., Moussiopoulos, N., Banias, G., Triantafillou, D., 2012. A multi-objective decision-making model to select waste electrical and electronic equipment transportation media. Resour. Conserv. Recycl. 66

[12] http://shekoudaily.com/blog/smart-recycling-machines-in-nanshan-offer-money-back.html

[13] https://medium.com/@swachhcoin/top-5-innovative-and-informative-waste-management-and-recycling-apps-b93f103c3e7e

AI and Big Data in Transportation Infrastructure for Condition and Risk Assessment

Can Balkaya[1], Erol Tutumluer [2]

[1]Professor, Department of Civil Engineering, Istanbul Aydın University,
Florya Campus, Kucukcekmece, Istanbul, Turkey
canbalkaya@aydin.edu.tr

[2]Professor, Department of CEE, University of Illinois at Urbana-Champaign,
1205 Newmark, Illinois, USA
tutumlue@illinois.edu

Abstract

AI and big data is changing future developments in transportation systems. Real-time transportation infrastructure of highway and railways for rapid condition assessment maintenance and rehabilitation by the help of AI and big data will reduce the accidents and improve the traffic safety. Transportation infrastructure and structures for condition assessment can be done by integration of the data obtained by using sensors, GIS, image process, vehicle cameras, drones, IoT and mobile phones. Even visual inspection or by using devices for maintenance is easy however it is difficult to observe structural condition assessment as settlement or other structural damages because of landslides, flooding or earthquake disaster that structure collapse can cause big tragedies. Artificial intelligence and big data transportation infrastructure management systems can be developed and improved by integration of condition and risk assessment data obtained previous applications. Geographic information systems using smart maps and integration with AI can be a solution for road infrastructure monitoring and maintenance to predict wear and tear. A deep learning model can be used for predicting road conditions as road roughness, crack damage levels and maintenance plans. On the other hand, for condition assessment of transportation structures, bridge structural information systems, upgrade deficiencies of bridge structures (functional and structural), high speed maglev and future transportation infrastructure on elevated guideway structures, prevention and mitigation of disasters in high speed trains are given special attention. A case study for the condition and risk assessment of a rail bridge is given for potential land sliding risk and settlement problem.

Keywords: Transportation infrastructure, condition assessment, risk assessment, artificial intelligence, maintenance.

1. Introduction

AI and big data is changing future developments in transportation systems. Artificial intelligence can reduce the human errors by using smart autonomous transportation in highways and railways. On the hand there are many accidents due to road or railway infrastructure imperfections. Thus, real-time monitoring transportation infrastructure of highway and railways for rapid condition assessment, maintenance and rehabilitation by the help of AI and big data will reduce the accidents and improve the traffic safety. High speed trains generally used elevated guideway structures. Thus, road and railway maintenance by using real-time monitoring to control of the infrastructure including special passes like lakes/bridge structures and tunnels are important for rapid condition and risk assessments. Transportation infrastructure and structures for condition assessment can be done by integration of the data obtained by using sensors, GIS, image process, vehicle cameras, drones, IoT, and mobile phones for information and communication.

Transportation infrastructure condition assessment can be done by visual inspection or using devices on site for maintenance of damaged guardrails, rails and corrosions as easily observed during investigation. However, it is difficult to observe structural condition assessment as settlement or other structural damages in superstructure as bridges/culverts because of landslides, flooding or earthquake disasters that can be cause big tragedies and lots of deaths. Earthquake will destroy and damage severely a large portion of transportation infrastructure systems that can be also important after the disaster for crisis management. By the advantage of AI and big data transportation infrastructure management systems can be developed and improved by integration of condition and risk assessment data obtained previous applications. This information can be stored for maintenance, rehabilitation planning as well as disaster managements.

2.Condition and Risk Assessment of Transportation Infrastructure

2.1. Condition Assessment for Maintenance of Transportation Infrastructure

Pavement condition assessment data for asphalt or concrete road can be determined by using vehicle video, drones, GIS with applications of image process techniques to determine pavement cracks, road bumps, potholes with the damage levels low-medium-high. Condition assessments steps are: 1. Data collection 2. Condition rating and 3. Quality management for transportation infrastructure. Pavement data collection is the most expensive and time-consuming component in pavement condition assessment. Among the automated and manual pavement data collection methods, automated data collection method is much safer, objective and faster. In the data collection, types of equipment, their calibrations and random field tests are important. Pavement condition assessment data systems are: 1. distress and profile data: the types, extent and severities of distresses on the pavement surface 2. structural data: load carrying capacity of the pavement 3. skid resistance: force developed wheel sliding along a pavement surface preventing from rotating and depending on the texture of pavements.

2.2 Artificial Intelligence and Big Data Solutions for Road Infrastructure Maintenance

Geographic information system using smart maps and integration with can be solution for road infrastructure monitoring and maintenance. Road maintenance can be efficiently planned in time and location depend on highway conditions. The Bavarian State Ministry of Housing, Building and Transport create a proof-of-concept model in 2019 by using road traffic density and road condition data collected between 2009 and 2017. This model [1] predicts wear and tear on the Bundesautobahn 70 (BAB 70 or A70) in southern Germany. A deep learning model is used for predicting road conditions as road roughness, crack damage levels as well as maintenance plans.

Fig.1. Dataset (4,839 road sections, each 100 meters long) describing road conditions and traffic density on the A70 road plan [1].

ArcGIS Pro and Python frameworks was utilized to access, verify, analyze, and visualize data and results. The road conditions and traffic data were presented in

Fig.1 in the form of feature classes, approximately 4800 road sections with each feature100-meter-long road section for every lane in both directions in the part of A70 highway. A collection of deep learning models predicts successfully highway wear and tear focusing on learning from measurements in both location and time.

2.3 Condition Assessment of Transportation Structures

2.3.1. Bridge Structural Information System

Bridge Structural Information system database will include: location, date of construction, types of superstructure (suspension, cable stay bridge, arch, etc.) material reinforced concrete, steel, composite, prestressed concrete, etc.), and geometry of superstructure (straight, curved or skewed in plan and/or elevation, or irregular) and expansion joints and bearings, number of spans, length of span and height of column/piers, abutments, materials and geometry of substructures (caissons, piles, etc.), construction details (cross-section dimensions, reinforcement percentage, transverse reinforcement and confinement, splices information, fixity conditions, seismicity information (earthquake region, fault locations, seismic ground acceleration, etc.) soil information (soil types, soil periods, settlements, allowable soil pressure, soil investigation report (if any)), foundation types (including structural drawings available for plan and section, number of piles vertical, inclined, type of piles, bearing capacity of piles, etc.) material or member deterioration or damages, traffic loading and vehicle information etc, environmental effects/loads information (potential of scouring, etc.), functional needs for current and future usage.

2.4 Upgrade Deficiencies of Transportation Structures

After a natural disaster or catastrophic event, damage assessment is used to collect initial damage information and other impacts immediately to prevent collapse depending on damage levels. A rapid damage assessment helps government and emergency management agencies to stabilize critical infrastructure, appropriate response and recovery activities.

Reduce the risk and damage before the disaster as resilient and sustainable studies, artificial intelligence can be used for the evaluation of condition and risk assessment by using damage index in ranking. The need for upgrading deficiencies of bridge structures (functional and structural) if found in risk group, such critical structures to meet functional and structural safety requirements needs to be retrofitting according to current national Turkish Bridge Earthquake Code [2] (effective October 6, 2021) or international code American Association of State Highway and Transportation Officials (AASHTO) [3] requirements. The basic steps are: integration of real time monitoring control information (settlements, etc) and bridge structures information systems with including damage information (if any) for retrofitting, performance level requirements, conducting seismic hazard analysis for the site and performing vulnerability to detect deficiencies and the level of damage, damage index, developing retrofitting methods with cost estimation for economic and safe solutions. For a country or a city lots of bridges structures can be evaluated for risk assessment in a rapid way by using artificial intelligence and big data applications. Previous retrofitting cases, solutions and cost information can be integrated and transfer to databases.

Most bridge structures do not show evidence of weakness and most of them need retrofitting for functionality, increase in traffic flow and loadings, change in seismicity data and seismic performance evaluation in the current code requirements with higher force and details than the constructed year code requirements. Most of the existing old transportations structures will not satisfy connections, foundations, piers. For these reasons, rapid speed trains will be more critical and real-time monitoring is necessary to prevent failures of structures or accidents or tragedies not only bridges but also for all structures culvert, tunnel, etc. and part of the transportation infrastructure that can be affected from floods, earthquakes and landslides.

2.5. High Speed and Future Transportation Infrastructure on Guideways

High-speed train infrastructure was mainly constructed on continuous guideway structures with fast progress of train technology as magnetic levitation or maglev for ground transportation system. Shanghai Maglev shown in Fig. 2 began operations in 2004 and connected Shanghai Pudong International Airport Station on the suburbs of the city [4] with a maximum speed of 430km/h.

Author Balkaya worked on conceptual guideway structural design for maglev high-speed ground transportation system in US Army Construction Research Lab. (USACERL) at Champaign with Prof. W. J. Hall [5] and Don Plotkin [6]. Artificial intelligence and big data will be the solution for high speed and future ground transportation systems by using rapid and real-time evaluation for transportation safety. Parameters used in the guideway design approach [5] can also be considered in the condition assessment and reduce the risk for magnetic levitation (maglev) or other future high-speed transportation systems on elevated guideway structures.

Fig. 2. Shanghai Maglev Train on Guideways [4]

Future rapid-speed transportation system can also use mainly superstructure as infrastructure on elevated guideways as well as evacuated tubes on superstructure. Spain-based Zeleros is evaluating composites and other materials in the development of its Hyperloop system [7]. For that reason, continuous superstructures as a transportation infrastructure will plays an important role for the future transportation system. The hyperloops uses the technology of magnetic levitation. Hyperloops advantages [8] are less polluting, air and noise pollution minimum, advantages against climate-change. Hyperloops could reduce major traffic jams on the city roads. The hyperloop tubes structure will include rooftop photovoltaic solar panels, resulting in a clean environment and a self-sustaining electric system. The hyperloop would send passengers between cities at speeds of more than 970km/h in capsules that float in partial vacuum tubes on mostly elevated guideways. Future transportation systems will include additional monitoring and condition assessment parameters with the advantage of artificial intelligence applications for transportation safety.

3. Prevention and Mitigation of Disaster in High Speed Trains

3.1. Disasters in the Railway Infrastructure

Types of disasters in the railway infrastructure are generally in two groups: natural and other manmade disasters. Natural disasters are: earthquakes, floods, tornadoes, landslides, tsunami, other extreme weather conditions, etc., train collisions (with a huge number of casualties), train derailments at a bridge over a river, and derailment of a train with explosives/inflammable materials, tunnel collapse on a train, fire or explosion in trains, and other miscellaneous cases etc.

Manmade disasters are: terrorism and sabotage as fire to a train, railway installations etc., bomb blast at station/train, chemical/biological and nuclear disaster. Prevention and mitigation of disaster plays an important role to reduce the loss of human life and properties by controlling/monitoring the systems for disaster resistance and disaster management planning.

Accidents on high-speed trains are much rarer compared to traditional ground transportation networks. These are mainly due to train derailments, signaling failures, extreme weather conditions, faulty systems during natural disasters especially in earthquake regions by the advantage of artificial intelligence, technological developments and big data, learning from past mistakes and real time monitoring and rapid decisions will decrease the risk in the future usage. Fatal accidents resulted in a derailment or stopping of the train are human error of technicians in case of wrong displaying of tracks. Also, high-speed rail drivers and inspection methods for condition assessments are critical for safety of operations. Rail drivers are trained on a simulator before they use the real vehicle. Non-destructive test (NDT) equipment's allows for continuous inspection of railway track up to 40km/h manufactured with ultrasonic, eddy current and combined for high-speed inspection of rails.

3.2. Prevention and Mitigation of Disaster in High Speed Trains

For prevention and mitigation of disaster at high speed trains the following items can be considered in their design [5]:

1. Route locations might be along the interstate highways or other corridors, for construction, retrofitting, maintenance and emergency management.

2. Special attention to the crossing of wetlands, rivers and soft ground, in the sense of long-term performance (settlement and stability) of the piers/piles (and guideway) for condition assessment and in the event of emergencies.

3. Vehicles tilting will be critical on curves.

4. Columns/guideways (shape selected for aesthetics) will be structurally monitoring for long term stabilities and settlement problems, soil-structures interaction, load and traffic flow increase in future/current usage and seismicity.

5. The matter of loss of power and measures to handle emergencies loss of power would shut down the system and may require the need for quick exit.

6. Safety of personnel (passengers and operating crew) in an emergency stop.

7. For different climatic conditions, the system would necessarily operate, from warm through cold climates, usage of air-condition systems. The route will be for its ability to accommodate the varying types of weather and environments.

8. Environmental effects as earthquakes (seismic effects), wind (straight winds up through tornadoes), rain, ice, snow, lightning, dirt, debris, and flooding, and their affiliated effects, including slope stability and liquefaction will be controlled by using sensors, cameras or GIS applications. Real time identification of disaster areas by an open-access vision-based tools [9] can also be used.

9. High winds, earthquakes irrespective of whether their source or faults is near or far can pose a major problem, not only vibration of the bridge/guideway system and the operating vehicle, but also rotational motion of the aboveground supports (pillars), differential ground settlement arising from slope instability or liquefaction.

10. Debris and dirt, can have a major effect upon the operational characteristics of the system. The debris could originate from environmental effects (i.e. wind) or be placed by people (sabotage). Some types of monitoring/control systems will be necessary for condition and risk assessments.

11. Vibration arising from operation of these high-speed systems can be a factor, especially in urban areas. The matter of derailment might occur at switches, and more unlikely (but possible) in the event of ground motions arising from unanticipated settlement, slides, or possibly earthquake effects.

12. Electromagnetic effects in maglev might affect passengers, much less those staff and maintenance workers located near the vehicles at stations.

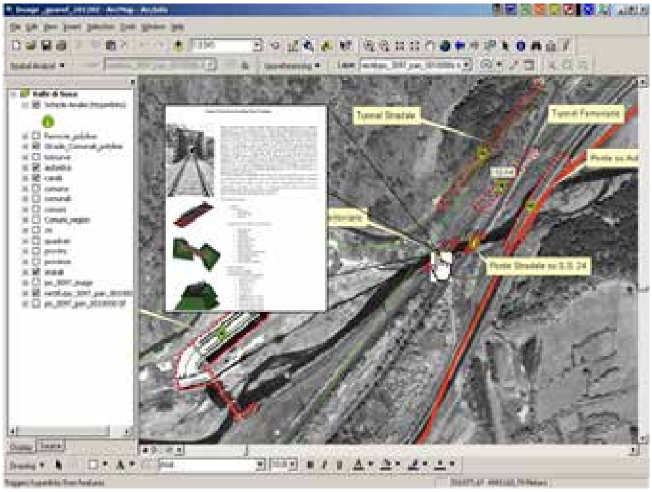
13. Flawed battery likely causes of Shanghai Maglev fire shown in Fig. 3 [10]. A faulty onboard storage battery has been identified as the possible cause of the fire that broke out on Shanghai’s magnetic levitation (maglev) train on August 11, 2006.

Fig. 3 Shanghai Maglev fire [10]

4. Case Study for the Condition and Risk Assessment of Transportation Infrastructure

An application for the condition and risk assessment of transport network infrastructure [11] is given for a node identified in the tunnel, bridge and land sliding site located in Pont-Ventoux in the Val Susa Valley (Province of Torino-Italy). Real time information as control assessment is integrated with structural information for risk assessment evaluation.

For the risk assessment of railway bridge in the Val Susa Valley in Italy, the steps are as follows: 1. GIS based satellite image is converted to a smeared 3-D model of ground 2. structural model and analysis of railway bridge are performed and for risk assessment vulnerability curves are obtained 3. results are compared with GIS based real-time measurements for risk assessment pre-disaster case and crisis management after disaster cases. To obtain geo-referenced satellite image that is an image where the coordinates of each pixel are exactly the real coordinates of the point with a reference system. After the geo-referential process, the image can be used as the support for the database by drawing, photos and technical information texts for transportation infrastructure especially in the high environmental risk regions because of land sliding, hydro-geological or seismicity/fault locations. A database for the case of railway bridge in Val Susa Valley was prepared and shown Fig.4a. The real rail bridge photo is shown in Fig.4b. The bridge has a span of 48.75m and a width of 5.10m. 3D structural analysis was done by using SAP2000.



a) Database b) Railway Bridge

Fig. 4. A database for railway bridge in Val Susa Valley [11]

a) Val Susa Valley Area Satellite Image b) Transportation Network Level Curves

Fig. 5. Satellite image and transportation network level curves for Val-Susa Valley

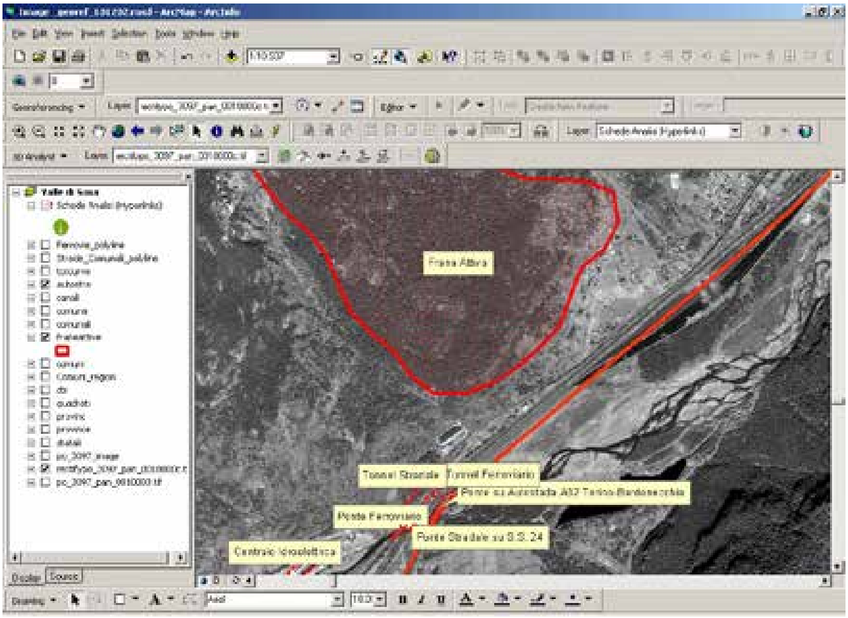


Fig. 6. Potential landslide on the Val Susa's site.

Satellite image and transportation network level curves for Val-Susa Valley are shown in Fig.5. Potential landslide was derived by using associated hazard function details given in reference [12] in Fig. 6. The results obtained for the case in post catastrophic event are reported in the Table 1 expressing the relationship between the complement to 1 of the safety factor and the support settlement. The derived vulnerability curves for support settlements are shown in Fig. 7, regarding the ordinate as the damage in equation of general risk, the likely uncertainty for the design variables converts each spike for given intensity and associated damage into a curve with in the ordinate the probability of that damage for the intensity in the abscissa.

Table 1. Relationship between complement to 1 of the Safety factor and support *Settlement.*

Settlement	0.00	0.05	0.08	0.15	0.17	0.19
Complement to 1 of the Safety Factor	0.00	0.15	0.35	0.64	0.68	0.73

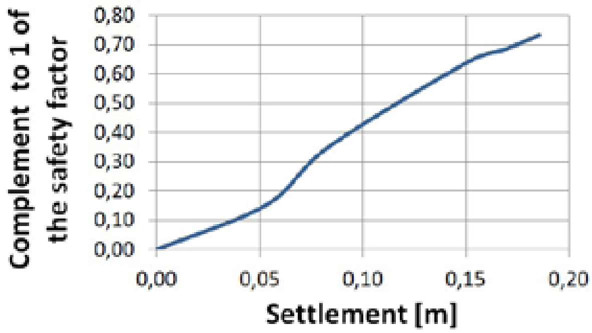


Fig. 7. Support settlement from derived the vulnerability curve.

Exact result would have required an elasto-plastic analysis for the entire family of critical nodes of the infrastructure is considered. A function in general risk equation linking the damage with the value of a mid-span displacement to be compared with its ultimate threshold. This is achieved by the pushover analysis of an equivalent structure for an estimate of the mid-span displacement given in Table 2 and a way to upload the global reliability of the bridge shown as the result of the pushover (blue curve) on an equivalent structure leading to an estimation of function in Risk's General Equations (red curve) in Fig.8.

Table 2. Relationship between displacement and force.

Displacement [m]	0.00	0.05	0.08	0.19	0.22	1.01	1.00
Force [kN]	0.0	342.3	426.5	474.2	486.5	486.5	486.5
Displacement [m]	0.00	0.05	0.08	0.10	0.13	0.14	0.15
Force "pushover analysis" [kN]	0.0	342.3	404.4	528.3	958.9	1078.7	1279.7

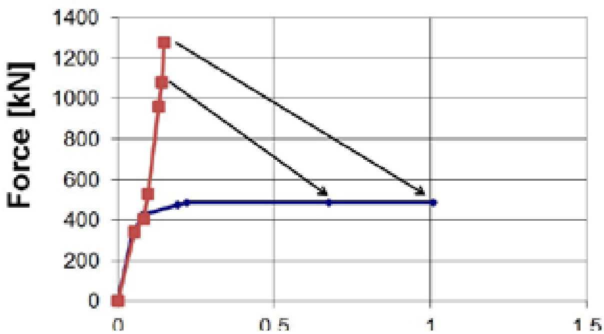


Fig. 8. Result of the pushover (blue curve) on an equivalent structure leading to an estimation of function in Risk's General Equations (red curve).

Artificial intelligence can be used for many transportation structures in national-wide for rapid decision/evaluation and big data to store all information databases from condition assessment as well as adopting the previous national and international case studies and information databases in deep learning expertise system to prevent disaster in risk assessment and to use in disaster management.

5. Conclusion

Artificial intelligence and big data will be the solution for high speed and future ground transportation systems mainly on elevated guideway structures as magnetic levitation maglev by using rapid and real-time evaluation for transportation safety.

Prevention and mitigation of disaster in high speed trains will be important in the design stage given this study with interdisciplinary works and with the help of artificial intelligence evaluation and decision using control and risk assessment learning from previous cases for both maintenance and strengthening as well as disasters for transportation safety.

Geographic information systems using smart maps and integration with AI can be a solution for road infrastructure monitoring and maintenance to predict wear and tear. A deep learning model can be used for predicting road conditions as road roughness, crack damage levels and maintenance plans.

An application for the condition and risk assessment of transport network infrastructure is given for node identified in the tunnel, bridge and land sliding site located in Pont-Ventoux in the Val Susa Valley (Province of Torino-Italy). Real time GIS, sensors, monitoring systems can be integrated structural information and vulnerability curves or damaged index for the structural control and risk assessment especially in high seismic regions, flooding regions, potential landslide region or support settlements because of transportation train loads, soil-structure effects or liquefaction. Results are compared with based on real-time measurements for risk assessment pre-disaster case and crisis management after disaster cases.

Integration of structural information database as bridge structural information system and real-time sensors data to transfer database system and previously collected infrastructures data for existing bridge or infrastructure systems adopted and transfer to big data for a country or a city or a region will result in rapid decisions by using artificial intelligence applications to prevent disasters and to save the lives for transportation infrastructure system safety.

Acknowledgements

The author thanks to University of Pavia, Department of Civil Engineering, Professor A. Nieto for Val-Susa Valley, Italy case study applications.

References

[1] Gimmler L. and Giner N., (2021) Analytic, https:www.esri.com/arcgis-pro/analytics/paving-the-road-ahead-with-arcgis-and-ai/

[2] Turkish Bridge Earthquake Code (2021), Ministry of Transportation and Infrastructure.

[3] LRFD Bridge Design Specifications, 9 th Edition, American Association of State Highway and Transportation Officials, AASHTO, 2020

[4] W. Li, D. Li, X. Zhang, J. Cao, "Status and research progress of the linear rail transit system in China", Transportation Systems and Technology 2(1), 16-41, 2016.

[5] C. Balkaya and W.J. Hall, "Conceptual guideway structural design for MAGLEV high-speed ground transportation system", International Journal of Engineering Technologies, 3(2), 61-71, June 2017.

[6] Don Plotkin and Simon Kim, "Maglev guideway design issues", Journal of Transportation Engineering, May/June 1997, 189-191, 1997.

[7] https://www.compositesworld.com/news/zeleros-raises-7-million-for-european-hyperloop-development- Edited by Grace Nehls, June, 5, 2020.

[8] https://www.engineeringpassion.com/hyperloop-technology-the-future-of-transportation/, Aimal Khan, August 6, 2020.

[9] C. Balkaya, F. Casciati, S. Casciati, L. Faravelli, M. Vece**,** "Real-time identification of disaster areas by an open-access vision-based tool"**,** Advances in Engineering Software*,* Vol. 88, October 2015.

[10] Flawed Battery likely cause of Shanghai Maglev fire (Xinhua) Updated: 2006-08-21 10:31, China Daily, http://www.chinadaily.com.cn/china/2006-08/21/content_669554.htm

[11] C. Balkaya, F. Casciati, D. Bortoluzzi, "System architecture for an interactive database to structural control of transport network infrastructure: A case study of railroad bridge", Proceedings of the 5th European Conference on Structural Control EACS2012, Genoa, Italy, 2012.

[12] Losi P., Risk analysis in the Val Susa Valley Area: Vulnerability Assessment, Master's degree thesis, Department of Structural Mechanics, University of Pavia, 2003.

On the Exploitation and Exploration Search Behaviors of Interactive Search Algorithm

Ali Mortazavi[1], Mahsa Moloodpoor [2]

[1]Graduate School of Natural and Applied sciences, Ege University, Izmir, Turkey,
ali.mortazavi.phd@gmail.com

[2]Mechanical Engineering Department, Ege University, Izmir, Turkey,
mahsa.moloodpoor@gmail.com

Abstract

There are several studies in the literature that prove the successful application of the metaheuristic algorithms in the solution of different mathematic and engineering optimization problems. A more detailed investigation about their search mechanism can provide deeper insight into their working structure. Attained information can help researchers to select/develop metaheuristic technique(s) which is more compatible with their desired problem. In this regard, the current investigation deals with empirical analysis of the search behavior of these methods. In this regard the Interactive Search Algorithm (ISA) is selected as the pilot method and its performance is compared with five other metaheuristic techniques. To this end, predefined parameters based on normalized diversity concept are utilized to monitor the search behavior and performance of selected methods. The results are provided through illustrative diagrams and tables. Attained outcomes reveal that, adaptability and coherence of the agents during the entire optimization process play a key role in the search performance of a population-based metaheuristic algorithm.

Keywords: Metaheuristic methods, Exploration and exploitation, Diversity concept

1. Introduction

In many engineering systems, reaching an optimal state is the final goal of the design process. In this regard, optimization methods can be divided into two main categories as deterministic and non-deterministic methods. The deterministic methods require a continuous objective functions and gradient information to determine the search direction and step sizes [1]. Since in many engineering problems finding such objective functions is very difficult (or even impossible), their application in complex optimization problems is highly challenging. Due to this reason, along with achieved developments in the computer technology, the non-deterministic techniques have received a lot of attention from researchers to solve complex problems from different fields of engineering and sciences. These methods do not require any continuous objective function and/or its gradient information to handle optimization problems [2].

One of the most important group of these techniques is the metaheuristic algorithms [3]. These methods are usually inspired from natural phenomena (e.g., physical rules and social behaviors and etc.). One can list some these methods chronologically as: Differential Evolution (DE) [4], Ant Colony Optimization (ACO) [5], Gravitational Search Algorithm (GSA) [6], Symbiotic Organisms Search (SOS) [7], Drosophila Food-Search Optimization (DFO) [8], Search group algorithm (SGA) [9] Water Wave Optimization (WWO) [10].

These methods do not demand continuous objective function and/or its gradient information. Also, they are efficient and ease to implementation. These specifications make them suitable techniques to solve complex engineering optimization problems in which finding a continuous objective function is very hard or even impossible. Therefore, in past two decades, these techniques are widely employed to solve different types of optimization problems [11-16]. However, the investigations addressed in the relevant literature designate that the different metaheuristic methods demonstrate various performance in handling the optimization problems [17-20]. Most of the related studies assessed the search capability of these methods numerically (i.e., quantitatively). However, still, it can provide illustrative information about why they show varied performances if their behaviors are evaluated qualitatively.

So, in the current work, a concept of population diversity is applied to trace and evaluate the search behavior of metaheuristic algorithms. In this regard, the current study aims to investigate the search behavior of Interactive Search Algorithm (ISA) [21] which is selected as sample metaheuristic method. To provide a more insight into the working mechanism of this method, its search behavior is assessed based on its different internal parameter settings. Also, to provide more generalizable data, the ISA's search behavior is, also, compared with three other metaheuristic algorithms. The selected methods are Butterfly Optimization Algorithm (BOA) [22], Jaya algorithm [23], and Firefly Algorithm (FA) [24]. The relation between search behavior and performance of the ISA algorithm and other selected methods are tested on solving four mathematical functions with different properties.

2. Interactive Search Algorithm (ISA)

The Interactive Search Algorithm (ISA) is a population-based metaheuristic search technique [21]. This method applies *tracking* and *interacting* search pattern to perform exploration and exploitation search behaviors, respectively. the ISA uses tendency factor to adjust usage of these two patterns [21]. Based on this information the ISA is mathematically formulated as below:

$$Tracking\ pattern : (\tau \geq 0.3)$$

$$^{t+1}\boldsymbol{V}_i = \omega \cdot {}^{t}\boldsymbol{V}_i + \boldsymbol{\varphi}_1 \odot ({}^{t}\mathbf{X}_i^P - {}^{t}\mathbf{X}_i) + \boldsymbol{\varphi}_2 \odot ({}^{t}\mathbf{X}^G - {}^{t}\mathbf{X}_i^P) + \boldsymbol{\varphi}_3 \odot ({}^{t}\mathbf{X}^W - {}^{t}\mathbf{X}_i)$$

$$Interacting\ pattern : (\tau < 0.3)$$

$$^{t+1}\boldsymbol{V}_i = \boldsymbol{\varphi}_4 \odot ({}^{t}\mathbf{X}_i - {}^{t}\mathbf{X}_j) \qquad if \qquad f(\mathbf{X}_i) \leq f(\mathbf{X}_j)$$

$$^{t+1}\boldsymbol{V}_i = \boldsymbol{\varphi}_4 \odot ({}^{t}\mathbf{X}_j - {}^{t}\mathbf{X}_i) \qquad if \qquad f(\mathbf{X}_i) > f(\mathbf{X}_j)$$

$$Updating\ formulation:$$

$$^{t+1}\boldsymbol{X}_i = {}^{t}\boldsymbol{X}_i + {}^{t+1}\boldsymbol{V}_i$$

where, □ indicates Hadamard product, and τ is the tendency factor which is selected randomly from the [0,1] interval; $\mathbf{X}^P$ is the memory matrix which records the previous best locations of agents; ${}^{t}\mathbf{X}_i^P$ and ${}^{t}\mathbf{X}_j^P$ respectively designate the current agent and an arbitrarily selected agent from $\mathbf{X}^P$ matrix while i≠j; $\mathbf{r}_j^T$, $\boldsymbol{\varphi}_1, \boldsymbol{\varphi}_2, \boldsymbol{\varphi}_3$ and $\boldsymbol{\varphi}_4$ are random vectors which their components are selected from [0,1] interval [21]; ${}^{t}\mathbf{X}^W$ is the weighted agent, and it is formulated as follows:

$$\mathbf{X}^W = \sum_{s=1}^{M} \bar{c}_s^w \mathbf{X}_s^P$$

in which,

$$\bar{c}_s^w = \left[\hat{c}_s^w \Big/ \sum_{s=1}^{M} \hat{c}_s^w \right]$$

$$\hat{c}_s^w = \frac{\max\limits_{1\leq k\leq M}\left(f(\mathbf{X}_k^P)\right) - f(\mathbf{X}_s^P)}{\max\limits_{1\leq k\leq M}\left(f(\mathbf{X}_k^P)\right) - \min\limits_{1\leq k\leq M}\left(f(\mathbf{X}_k^P)\right) + \varepsilon} \qquad , s = 1,2,\dots,M$$

where, M is the size of population, is the parameter presents the effect of ith agent on the weighted agent. f(.) returns the objective function value for the current agnt, $\max\limits_{1\leq k\leq M}\left(f(\mathbf{X}_k^P)\right)$ and $\min\limits_{1\leq k\leq M}\left(f(\mathbf{X}_k^P)\right)$ are the worst and best agents' objective function values, respectively. ε is defined as a small positive number to prevent from a probable division by zero case, if any.

3. Applied Behavioral and Performance Indexes

Reporting the numerical results may present a global view about the search capability of metaheuristic methods. However, defining qualitative indexes can provide a deeper insight into the working mechanism of these techniques. in this regard, to trace exploration and exploitation search behaviors of the selected algorithms two indexes based on the diversity level of the population are employed as below [25]:

$$\mathrm{Div} = \left(\frac{1}{\mathrm{nD}}\right) \sum_{J=1}^{D} \sum_{i=1}^{n} median(x^j) - x_i^j$$

$$Exploration\% = \left(\frac{\mathrm{Div}}{\mathrm{Div}_{max}}\right) \times 100$$

$$Exploitation\% = \left(\frac{|\mathrm{Div} - \mathrm{Div}_{max}|}{\mathrm{Div}_{max}}\right) \times 100$$

In which, Div indicates the diversity of the population and measured by first norm of the current component of the jth dimension with the median of the all population over the jth dimension; reflects the maximum recorded level of diversity; n and D, are respectively show the number of population and dimension of the problem. The argument behind these definitions is that, when the algorithm starts from arbitrary locations in the search domain the diversity of the population is high and gradually when the promising location of the domain is spotted the diversity of the population is increased to search the vicinity of the detected location in more detail. Also, in the current work, to provide a comparative index to compare the search performance of different methods (or single method with different settings). This index uses the exploration search behavior and object function value and defined as follows:

$$Performance\% = Exploitation\% \times (OFE_{max}/OFE)$$

where, shows the maximum objective function value (i.e., the worst objective function in a minimization problem). The logic behind this index is that, when the algorithm spots the promising location of the search space tries to reduce the diversity of the population and also the found optimal solution shows the accuracy of the located optimal point. So, by combining these two concepts we can achieve an index to qualitatively evaluate the performance of the algorithm.

4. Numerical Tests

In the current section, the performance and search behavior of the ISA methods is assessed based on its different internal tendency parameter settings. Also, to provide more details about its efficiency, it is compared with Butterfly Optimization Algorithm (BOA) [22], Jaya algorithm [23], and Firefly Algorithm (FA) [24]. The setting for these methods is provided in the following table.

Table 1 Parameter setting for the applied algorithms

Algorithm	Year	Parameter
FA [24]	2009	$\alpha = 0.5, \beta_{min} = 0.2, \gamma = 1$
Jaya [23]	2016	Parameter-free
BOA [22]	2019	P=0.8, c=0.01, a=0.1

For case study, four functions from CEC 2017 database [26] are selected. To provide more clarity, their 3D schematics are given below:

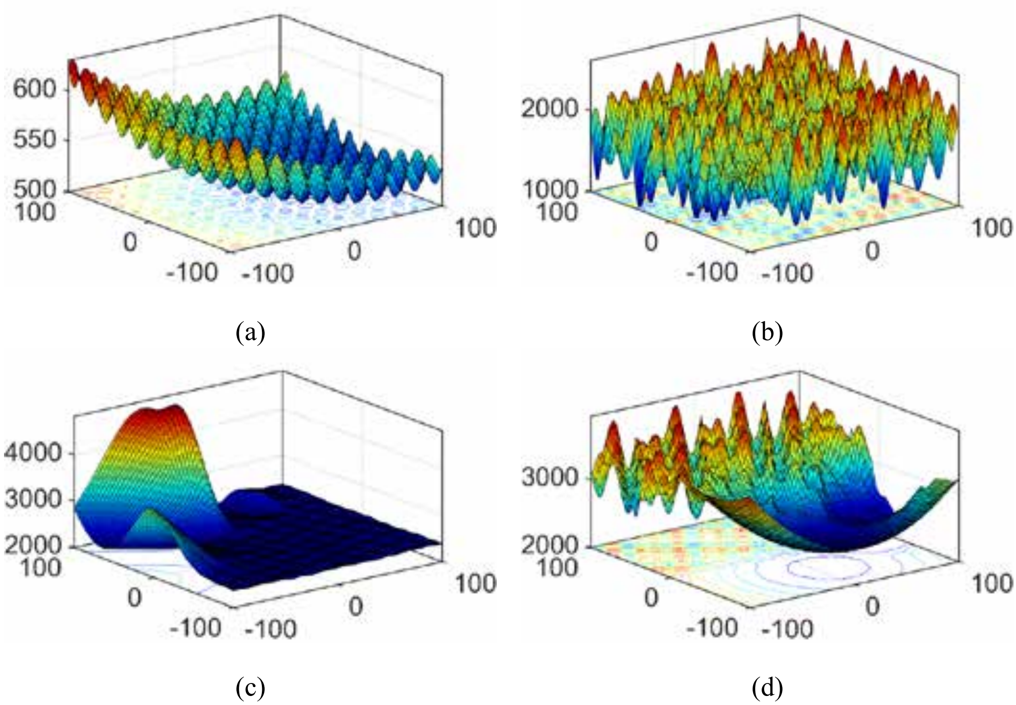


Fig. 1 3D schematics for (a) F1, (b) F5, (c) F20, and (d) F22 of CEC 2017

4.1. Effect of tendency factor on performance of the ISA

In this section using defined indexes the performance of the ISA methods is evaluated based on the its tendency factor (i.e., its internal adjusting parameter). The performance diagrams (using *Performance%* factor) is provided in Fig. 2. Based on the given diagrams, when the τ=0.3 (i.e., 70% exploration and 30% exploitation) the algorithm has higher performance on solving the F5, F10 and F22 functions, however for F20 function the higher exploitative behavior (τ=0.7) leads to better solution. The outcomes indicate that, tough for many of the metaheuristic system the internal adjusting parameter is set to a constant value (e.g., based on different sensitivity analyses), for more a higher search performance they should be tune for each specific problems exclusively.

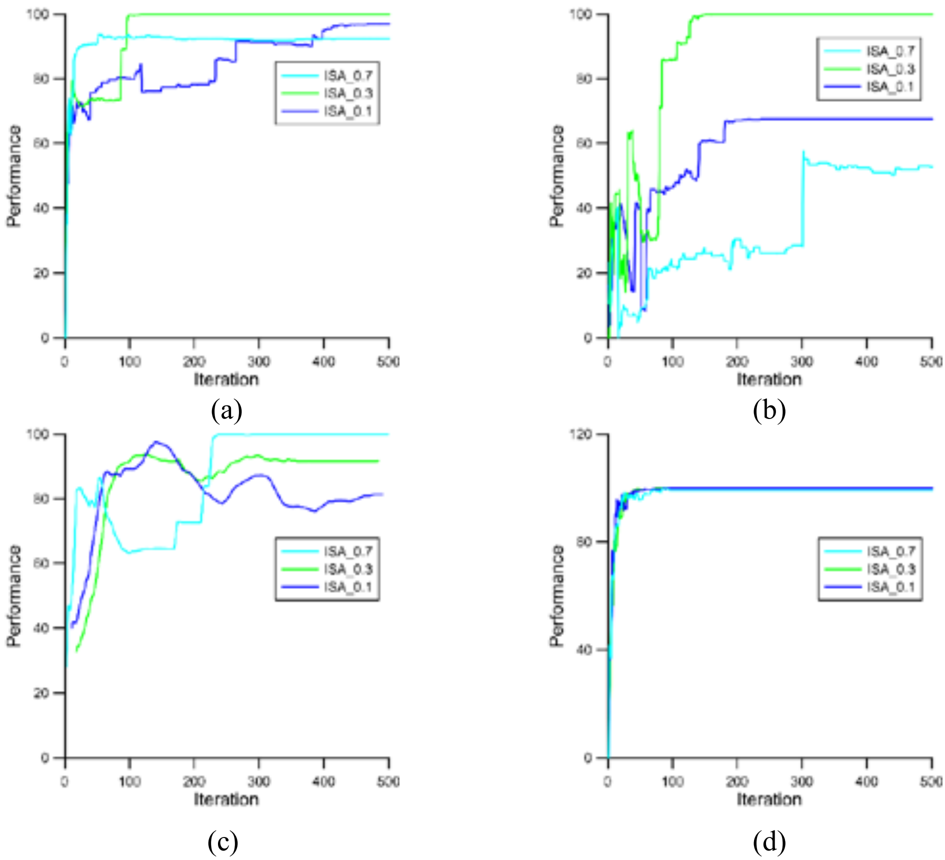


Fig. 2 The performance diagrams for (a) F5, (b) F10, (c) F20 and (d) F22 for different values of tendency factor

4.2. Assessment of ISA Compared with Other Methods

In the current section, the ISA's search behavior and performance are compared with BOA, Jaya and FA methods. In this regard, initially the behavioral diagrams based on the normalized diversity index are provided and then the performance of selected methods is compared based on the *Performance%* factor. the acquired numerical results are given in Table 2 and the behavioral and performance diagrams are respectively shown in Figs 3 and 4.

Table 2 Comparative results for F5, F10, F20 and F22 function from CEC2017 Database

Function	Value	FA	Jaya	ISA	BOA
F5	OFV	553.52	537.73	509.94	623.84
	Rank	3	2	1	4
F10	OFV	3378.25	2183.25	1736.29	2563.82
	Rank	4	2	1	3
F20	OFV	2246.65	2198.25	2118.65	2177.32
	Rank	4	3	1	2
F22	OFV	2309.82	2332.65	2300.49	2445.63
	Rank	2	3	1	4

qBased on given behavioral diagrams, the ISA and Jaya have complete behavioral diagrams. However, the BOA and FA have not a complete search behavior since the exploration and exploitation search behaviors are not even shifted during the optimization problems.

Based on the given results, the ISA could outperform other selected methods in terms of performance and accuracy. Considering F20 function, as given in Table 2, the BOA can find more accurate result than the Jaya, but based on the given performance diagrams the Jaya has higher performance level than BOA. This shows that, Jaya has a better search quality and the BOA can find better solution more stochastically. Attained more complex results for engineering problems proves this conclusion [20] .The same condition can be observed for the Jaya and FA methods. Such results indicate that, performance diagrams (as qualitative measurement tool) provide deeper insight into search capacity of a method rather than just numerical results (as quantitative measurement tool).

5. Conclusion

In current study, the search mechanism of the ISA is assessed depended on its internal adjusting parameter (tendency factor) and in comparison with Butterfly Optimization Problem (BOA), Jaya algorithm and Firefly Algorithm (FA). To meet this aim, a normalized diversity index is applied to evaluate the search behavior of the algorithms (and and also the performance index (combination of objective function value (OFV) and search behavior of algorithm) to qualitatively evaluate the algorithms. Based on the acquired results, the coherence between population agents plays important role in the success of a population-based metaheuristic. Also, to evaluate a metaheuristic method, as global optimizer tool, its more proper to evaluate them qualitatively rather than just quantitatively (i.e., based on the suite of digital data). This approach (i.e., using the behavioral and performance analyses) is also can be used in tuning the internal parameters of algorithm instead of conventional numerical sensitivity analyses.

Reference

[1] Hasançebi O., Teke T., Pekcan O., "A bat-inspired algorithm for structural optimization", Computers & Structures, doi: http://dx.doi.org/10.1016/j.compstruc.2013.07.006, Vol. 128, No. pp. 77-90. 11//2013

[2] Mortazavi A., "A new fuzzy strategy for size and topology optimization of truss structures", Applied Soft Computing, doi: https://doi.org/10.1016/j.asoc.2020.106412, Vol. 93, No. pp. 106412. 2020/08/01/2020

[3] Mortazavi A., "Interactive fuzzy search algorithm: A new self-adaptive hybrid optimization algorithm", Engineering Applications of Artificial Intelligence, doi: https://doi.org/10.1016/j.engappai.2019.03.005, Vol. 81, No. pp. 270-82. 2019/05/01/2019

[4] Storn R., Price K., "Differential Evolution – A Simple and Efficient Heuristic for global Optimization over Continuous Spaces", Journal of Global Optimization, doi: http://dx.doi.org/10.1023/a:1008202821328, Vol. 11, No. 4, pp. 341-59. 1997

[5] Dorigo M., Blum C., "Ant colony optimization theory: A survey", Theoretical Computer Science, doi: http://dx.doi.org/10.1016/j.tcs.2005.05.020, Vol. 344, No. 2, pp. 243-78. 2005/11/172005

[6] Rashedi E., Nezamabadi-pour H., Saryazdi S., "GSA: A Gravitational Search Algorithm", Information Sciences, doi: http://dx.doi.org/10.1016/j.ins.2009.03.004, Vol. 179, No. 13, pp. 2232-48. 6/13/2009

[7] Cheng M.-Y., Prayogo D., "Symbiotic Organisms Search: A new metaheuristic optimization algorithm", Computers & Structures, doi: https://doi.org/10.1016/j.compstruc.2014.03.007, Vol. 139, No. pp. 98-112. 2014/07/15/2014

[8] Das K.N., Singh T.K., "Drosophila Food-Search Optimization", Applied Mathematics and Computation, doi: http://dx.doi.org/10.1016/j.amc.2014.01.040, Vol. 231, No. pp. 566-80. 3/15/2014

[9] Gonçalves M.S., Lopez R.H., Miguel L.F.F., "Search group algorithm: A new metaheuristic method for the optimization of truss structures", Computers & Structures, doi: http://dx.doi.org/10.1016/j.compstruc.2015.03.003, Vol. 153, No. pp. 165-84. 6//2015

[10] Zheng Y.-J., "Water wave optimization: A new nature-inspired metaheuristic", Computers & Operations Research, doi: http://dx.doi.org/10.1016/j.cor.2014.10.008, Vol. 55, No. pp. 1-11. 3//2015

[11] Mortazavi A., Togan V. Metaheuristic Algorithms for Optimal Design of Truss Structures. In: Nigdeli SM, Bekdaş G, Kayabekir AE, Yucel M, editors. Advances in Structural Engineering—Optimization: Emerging Trends in Structural Optimization. Cham: Springer International Publishing; 2021. p. 199-220.

[12] Moloodpoor M., Mortazavi A., Ozbalta N., "Thermal analysis of parabolic trough collectors via a swarm intelligence optimizer", Solar Energy, doi: https://doi.org/10.1016/j.solener.2019.02.008, Vol. 181, No. pp. 264-75. 2019/03/15/2019

[13] Mortazavi A., Toğan V., Nuhoğlu A., "Weight minimization of truss structures with sizing and layout variables using integrated particle swarm optimize", Journal of Civil Engineering and Management, doi: https://doi.org/10.3846/13923730.2017.1348982, Vol. 23, No. 8, pp. 985-1001. 2017/11/172017

[14] Hasançebi O., Çarbas S., Saka M.P., "A reformulation of the ant colony optimization algorithm for large scale structural optimization", Civil-Comp Proceedings, doi: Vol. 97, No. pp. 2011

[15] Hasançebi O., Çarbaş S., "ANT COLONY SEARCH METHOD IN PRACTICAL STRUCTURAL OPTIMIZATION", International Journal of Optimization in Civil Engineering, doi: Vol. 1, No. 1, pp. 91-105. 2011

[16] Hasançebi O., Erbatur F., "Layout optimization of trusses using improved GA methodologies", ActaMechanica, doi: 10.1007/BF01178797, Vol. 146, No. 1, pp. 87-107. 2001//2001

[17] Hasançebi O., Çarbaş S., Doğan E., Erdal F., Saka M.P., "Comparison of non-deterministic search techniques in the optimum design of real size steel frames", Computers & Structures, doi: http://dx.doi.org/10.1016/j.compstruc.2010.06.006, Vol. 88, No. 17–18, pp. 1033-48. 9//2010

[18] Hasançebi O., Çarbaş S., Doğan E., Erdal F., Saka M.P., "Performance evaluation of metaheuristic search techniques in the optimum design of real size pin jointed structures", Computers & Structures, doi: http://dx.doi.org/10.1016/j.compstruc.2009.01.002, Vol. 87, No. 5–6, pp. 284-302. 3//2009

[19] Mortazavi A., "The Performance Comparison of Three Metaheuristic Algorithms On the Size, Layout and Topology Optimization of Truss Structures", Mugla Journal of Science and Technology, doi: https://doi.org/10.22531/muglajsci.593482, Vol. 5, No. 2, pp. 28-41. 2019

[20] Mortazavi A., "Comparative assessment of five metaheuristic methods on distinct problems", Dicle University Journal of Engineering, doi: https://doi.org/10.24012/dumf.585790, Vol. 10, No. 3, pp. 879. 2019/09/29/2019

[21] Mortazavi A., Toğan V., Nuhoğlu A., "Interactive search algorithm: A new hybrid metaheuristic optimization algorithm", Engineering Applications of Artificial Intelligence, doi: https://doi.org/10.1016/j.engappai.2018.03.003, Vol. 71, No. pp. 275-92. 5//2018

[22] Arora S., Singh S., "Butterfly optimization algorithm: a novel approach for global optimization", Soft Computing, doi: http://dx.doi.org/10.1007/s00500-018-3102-4, Vol. 23, No. 3, pp. 715-34. 2019/02/012019

[23] Rao R.V., "Jaya: A simple and new optimization algorithm for solving constrained and unconstrained optimization problems", International Journal of Industrial Engineering Computations, doi: Vol. 7, No. 1, pp. 19-34. 2016

[24] Yang X.-S. Firefly Algorithms for Multimodal Optimization. In: Watanabe O, Zeugmann T, editors. Stochastic Algorithms: Foundations and Applications. Berlin, Heidelberg: Springer Berlin Heidelberg; 2009. p. 169-78.

[25] Hussain K., Salleh M.N.M., Cheng S., Shi Y., "On the exploration and exploitation in popular swarm-based metaheuristic algorithms", Neural Computing and Applications, doi: 10.1007/s00521-018-3592-0, Vol. 31, No. 11, pp. 7665-83. 2019/11/012019

[26] N.H. A., M.Z. A., Suganthan P.N., Liang J.J., B.Y. Q., "Problem definitions and evaluation criteria for the CEC 2017 special session on real-parameter optimization", Technical Report, Nanyang Technological University, Singapore, School of Computer Information Systems, Jordan University of Science and Technology, Jordan, School of Electrical Engineering, Zhengzhou University, Zhengzhou, doi: http://dx.doi.org/10.1007/s10732-008-9080-4, Vol. No. pp. 2017

Interactive Search Algorithm for Solving Thermal Model of Parabolic Trough Collector Receiver

Mahsa Moloodpoor[1], Ali Mortazavi[2], Necdet Özbalta[3]

[1]Department of Mechanical Engineering, Ege University, Izmir, Turkey, mahsa.moloodpoor@gmail.com

[2]Graduate School of Natural and Applied Science, Ege University, Izmir, Turkey, ali.mortazavi.phd@gmail.com

[3]Department of Mechanical Engineering, Ege University, Izmir, Turkey, necdet.ozbalta1@gmail.com

Abstract

In the current study a one-dimensional cross-sectional heat transfer analysis of a parabolic trough collector (PTC) receiver is investigated. The heat transfer equations of the thermal model are adopted to a constrained optimization model and solved via a recently developed Interactive Search Algorithm (ISA). Since ISA is a non-deterministic optimization algorithm, it is not sensitive to the start point and step intervals of the search process. The objective function of the optimization problem is achieved by establishing a one-dimensional thermal balance between the ambient and heat transfer fluid flowing inside the receiver. On the other hand the constraint of the optimization problem is defined as the ultimate balance in the absorber temperature. The proposed optimization model is tested on SEGS LS-2 PTC module. The attained theoretical results are validated by comparing them with the practical data reported by Sandia National Laboratory (SNL). Also, the corresponding percentage errors indicate that the proposed approach provides promising results on solving the governing equations of the thermal model of PTC receiver tube.

Keywords: Interactive Search Algorithm, Parabolic Trough Collector, Thermal analysis

Nomenclature

A	cross-sectional area, m^2	Greek letters	
c_p	specific heat capacity, J/kg K	α	absorptivity
D	diameter, m	α_n	normal absorptivity
f	focal distance, m; friction factor	γ	intercept factor, shape factor
F'	collector efficiency factor	ε	emissivity
F''	collector flow factor	η_{th}	thermal efficiency
F_R	collector heat removal factor	μ	dynamic viscosity, Pa.s
g	optimization constraint	ρ	reflectance of parabolic mirror; density, kg/m^3
h	convection heat transfer coefficient, W/m^{2}K	σ	Stefan Boltzmann constant (5.67$\times$10^{-8} W/m^2 ^{o}K)
I_b	beam radiation, W/m^2	τ	transmissivity
k	thermal conductivity, W/m K		
$K_{\gamma\tau\alpha}$	incident angle modifier	Subscripts	
l_a	absorber length, m	a	absorber
m	mass flow rate, kg/s	ai	inside of absorber
n	refractive index	ao	outside of absorber
Nu	Nusselt number	amb	ambient
Pr	Prandtl number	c	cover
Q	Rate of heat transfer, W	ci	inside of cover
Q_u	Useful heat gain, W	co	outside of cover
Re	Reynolds number	env	environment
S	absorbed solar energy, W/m^2	fi	inlet fluid
T	temperature, K	sky	sky
U_L	thermal loss coefficient, W/m^2 K		
v_w	wind velocity, m/s		
w_a	parabola aperture, m		

1. Introduction

Parabolic trough collectors (PTCs) are one of the most utilized concentrated solar collectors, which are applied in high-capacity solar energy power plants. These collectors are promising technology for working in high temperatures. In a PTC there is a receiver tube, reflector surface and standing equipment. The receiver tube is located in the focal line of the reflector surface and a heat transfer fluid (HTF) flowing inside it. The sun radiance coming to the reflector surface is reflected to the receiver tube and heating up the HTF inside the tube. The heated fluid entering a Rankine cycle provides the demand energy for its proper application. The performance of the collector is depending on the optical and thermal efficiency of the system. The optical efficiency varies through the geometric properties of the reflector surface and receiver tube, the solar radiance amount and the material of the components of the system. Where, the thermal efficiency depends to process properties of the collector like the temperature of the inlet HTF, the mass flow rate of the HTF, the ambience temperature, velocity of the wind and the thermo-physical properties of the HTF. For analyzing the performance of the system and predicting its performance requires to carry out a thermal model of the receiver tube. Solving the achieved thermal model is a complicated process. In the reported technical literature there are studies focused on thermal modeling and solving of the receiver tube in PTCs. Daniel et al. (2011) investigated the effect of the outer vacuum shell around the absorber tube on the performance of the collector [1]. Padilla et al. (2011) carried out a one-dimensional heat transfer study on the PTC systems, neglecting the interactions between the surfaces close to each other [2]. Yılmaz and Söylemez (2014) investigated the heat losses and the effect of optical efficiency in the performance analysis of PTC, considering the heat losses from the receiver tube for steady-state flow in one dimensional condition [3]. Hachicha et al. (2013) implemented a numerical study on the heat transfer analysis in a PTC system [4]. In some other studies the effect of using nano-fluids as the HTF on the performance of the system is investigated. For instance, Risi et al. (2013) utilized CuO and Ni nanoparticles for providing higher absorption capability and applied optimization procedure for determining the main geometrical and operational properties [5]. Zadeh et al. (2015) developed an optimization model for thermal analysis in a PTC using nano-fluids [6]. Moloodpoor et al. (2019) solved the proposed one-dimensional optimization problem via integrated Particle Swarm Optimization (iPSO) technique.

In the current study, a one-dimensional thermal model of a PTC receiver tube is employed and its governing equations are mathematically converted into a single-objective constrained optimization problem. In this regard, a thermal balance between the HTF and ambient surrounding the receiver tube is provided. The difference between nodal heat losses in tube's one-dimensional sectional path is considered as the objective function. To solve the achieved optimization problem recently developed Interactive Search Algorithm (ISA) is utilized. ISA is a non-deterministic optimization algorithm, so, it is not sensitive to the start point and step intervals of the search process. The proposed optimization model is tested on SEGS LS-2 PTC module. The attained theoretical results are validated by comparing them with the experimental data reported by Sandia National Laboratory (SNL).

In the rest of the current study the thermal model of the receiver tube is detailed in section 2. The utilized optimization method is described in section 3. Section 4 is devoted to numerical solution of the optimization problem. Finally, the concluded points are summarized in the conclusion part.

2. Thermal modeling

The receiver tube of a PTC consist of cylindrical layers. The sectional view of these layers is presented in Fig. 1a. Heat transfer fluid is flowing in the center of this tube, inside the absorber tube. A glass cover is surrounded the absorber tube for preventing heat losses. The area between the glass cover and absorber tube is named as annulus part. This section in different conditions could be vacuumed or filled with different gases. Considering the receiver tube as a heat transfer element the thermal resistance method is applied for estimating the nodal temperatures. The heat transfer from the

HTF to the ambient is schematized with the help of thermal resistance method in Fig. 1b. Based on the temperature diversity in radial aspect, governing equations between nodal points of the resistance network are as below.
The convection heat transfer between the absorber tube and heat transfer fluid (HTF) depends on absorber and fluid temperature is calculated as below [7]:

$$\dot{Q}_{fi-a} = h_{fi}\pi D_{ai}\left(T_a - T_{fi}\right)$$

where the convection heat transfer coefficient for the heat transfer fluid, h_{fi} is given by [8]:

$$h_{fi} = Nu_{fi}\frac{k_{fi}}{D_{ai}}$$

For this aim, depending the HTF flow regime inside the absorber the Nusselt number and friction factor could be determined. Also, the fluid's thermo-physical properties such as conductivity, viscosity and specific heat capacity is available by knowing the temperature of the fluid.

The absorber tube's wall is very thin, so, the heat transfer through absorber tube via conduction is neglected [9]. The annulus part of the receiver in the current study is considered to be vacuumed. So the heat transfer between the absorber tube and glass cover inner wall is occurred via radiation.

$$\dot{Q}_{a-ci} = \frac{\pi D_{ao} l_a \sigma\left(T_a^4 - T_{ci}^4\right)}{\frac{1}{\varepsilon_a} + \frac{1-\varepsilon_c}{\varepsilon_c}\left(\frac{D_{ao}}{D_{ci}}\right)}$$

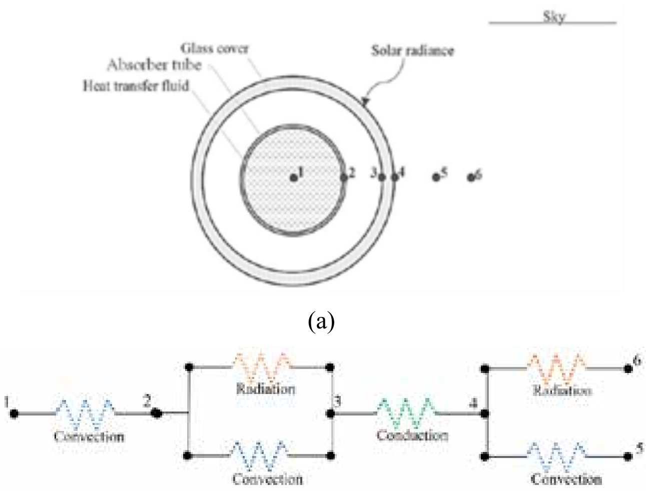


Fig.1. Receiver tube: a) cross sectional view and b) thermal resistance model

The conduction heat transfer through the glass cover's wall is given as bellow [7]:

$$\dot{Q}_{ci-co} = \frac{2\pi k_c l_a (T_{ci} - T_{co})}{\ln\left(\frac{D_{co}}{D_{ci}}\right)}$$

Heat transfer between the glass cover and environment is the sum of the radiation heat transfer between the glass cover and sky and convection heat transfer between the glass cover and ambient [7].

$$\dot{Q}_{co-env} = \left(\pi D_{co} l_a h_w (T_{co} - T_{amb})\right) + \left(\varepsilon_c \pi D_{co} l_a \sigma \left(T_{co}^4 - T_{sky}^4\right)\right)$$

The heat transfer coefficient of the air flowing around the receiver tube is depending the wind velocity and the temperature of the flowing air. The temperature of the sky is calculated as follows [10]:

$$T_{sky} = 0.0552 T_{amb}^{1.5}$$

Consequently, the energy balance equations for all nodal points in receiver tube's cross-sectional path is as below.

$$\dot{Q}_{loss} = \dot{Q}_{co-env} = \dot{Q}_{ci-co} = \dot{Q}_{a-ci} = \dot{Q}_{fi-a}$$

Thermal performance in PTCs considering the optical properties of the received radiation, gained heat and heat losses could be evaluated. Thermal efficiency of the receiver tube in PTCs is computable by [11]:

$$\eta_{th} = \frac{\dot{Q}_u}{A I_b}$$

Received beam radiation and optical properties determine the solar radiation absorbed per unit by the reflector surface.

$$S = I_b \rho (\gamma \tau \alpha)_n K_{\gamma \tau \alpha}$$

Useful gained heat, collector heat removal factor, collector efficiency factor and collector flow factor is calculated as below.

$$\dot{Q}_u = F_R A_a \left[S - \frac{A_r}{A} U_L \left(T_{fi} - T_{amb}\right)\right]$$

$$F' = \frac{1/U_L}{\frac{1}{U_L} + \frac{D_{ao}}{h_{fi} D_{ai}} + \left(\frac{D_{ao}}{2k_a} \ln \frac{D_{ao}}{D_{ai}}\right)}$$

$$F" = \frac{\dot{m} c_p}{A_r U_L F\prime} \left[1 - \exp\left(-\frac{A_r U_L F}{\dot{m} c_p}\right)\right]$$

$$F_R = F' F"$$

The average temperature drop from the the absorber to the heat transfer fluid is [7]:

$$\bar{T}_a - \bar{T}_{fi} = \dot{Q}_u \left[\frac{1}{\pi D_{ai} L h_{fi}} + \frac{\ln(D_{ao}/D_{ai})}{2\pi k_r L}\right]$$

So, temperature of the absorber under different conditions can be calculated as below.

$$T_a = \frac{2T_{fi}+\Delta T}{2} + \left(\bar{T}_a - \bar{T}_{fi}\right)$$

ΔT is the temperature difference between the inlet and outlet temperatures of the HTF.

2.1. Objective function and constraint of the optimization problem

The governing equations of the thermal model is converted to a single-objective optimization problem. In this regard, the difference between nodal points of the cross-sectional view of the receiver tube for achieving an ultimate thermal balance must converge to zero. So, the objective function of the optimization problem is defined to minimizing the difference of the heat losses of inner and outer segments of the receiver tube[9].

$$f(\mathbf{X}) = minimize\ \left(\dot{Q}_{a-ci} - \dot{Q}_{co-env}\right)$$

X holds the design variables of the optimization problem as the absorber tube (T_a) and glass cover (T_{co}) temperatures. These design variables will be selected from [200, 600] interval during the optimization process.

The proper constraint of the optimization problem is the ultimate balance in absorber's temperature and it is defined as below.

$$g = T_a - \bar{T}_a = \left[\frac{2T_{fi}+\Delta T}{2} + \left(\bar{T}_a - \bar{T}_{fi}\right)\right] - \left[\bar{T}_{fi} + \dot{Q}_u\left[\frac{1}{\pi D_{ai} L h_{fi}} + \frac{\ln\left(\frac{D_{ao}}{D_{ai}}\right)}{2\pi k_r L}\right]\right] = 0$$

The complexity of the optimization problem due to its complicated search domain, challenges the solver optimization technique and the solutions are mostly closed or even at the edge(s) of the search space [12]. Consequently, ISA optimization method with good performance in solving different optimization problems is selected and described in the next section [13-15].

3. Interactive Search Algorithm (ISA)

Interactive Search Algorithm (ISA) is a non-deterministic and swarm-based optimization technique [16]. This method uses a tendency factor (τ), for balancing one of two different search patterns (i.e. Tracking and Interacting). In the Tracking phase, the exploration search behavior of the algorithm is activated. However, in Interacting phase the exploitation aspect of the algorithm is dominant. The Tendency factor, according to a series of sensitivity analysis is determined as τ=0.3 [16]. The proposed ISA is mathematically formulated as follow:

$$Tracking\ phase: (\tau \geq 0.3)$$

$$^{t+1}\mathbf{V}_i = \omega \cdot {}^{t}\mathbf{V}_i + \boldsymbol{\varphi}_1 \odot ({}^{t}\mathbf{X}_i^P - {}^{t}\mathbf{X}_i) + \boldsymbol{\varphi}_2 \odot ({}^{t}\mathbf{X}^G - {}^{t}\mathbf{X}_i^P) + \boldsymbol{\varphi}_3 \odot ({}^{t}\mathbf{X}$$

$$Interacting\ phase: (\tau < 0.3)$$

$$^{t+1}\mathbf{V}_i = \boldsymbol{\varphi}_4 \odot \left({}^{t}\mathbf{X}_i - {}^{t}\mathbf{X}_j\right) \qquad if \qquad f(\mathbf{X}_i) \leq f\left(\mathbf{X}_j\right)$$

$$^{t+1}\mathbf{V}_i = \boldsymbol{\varphi}_4 \odot \left({}^{t}\mathbf{X}_j - {}^{t}\mathbf{X}_i\right) \qquad if \qquad f(\mathbf{X}_i) > f\left(\mathbf{X}_j\right)$$

$$Updating\ formulation:$$

$$^{t+1}\mathbf{X}_i = {}^{t}\mathbf{X}_i + {}^{t+1}\mathbf{V}_i$$

where, □ is the Hadamard product and τ is randomly chosen from the [0,1] interval; $\mathbf{X}^P$ is a matrix which holds the previous best locations of agents; ${}^{t}\mathbf{X}_i^P$ shows the ith member and ${}^{t}\mathbf{X}_j^P$ is an arbitrarily selected member from $\mathbf{X}^P$ matrix while i≠j; $\mathbf{r}_j^T$, $\boldsymbol{\varphi}_1, \boldsymbol{\varphi}_2, \boldsymbol{\varphi}_3$ and $\boldsymbol{\varphi}_4$ are the random vectors which their components are chosen form [0,1] interval [16]; ${}^{t}\mathbf{X}^W$; is the weighted agent and it is calculated as follows:

$$\mathbf{X}^W = \sum_{s=1}^{M} \bar{c}_s^W \mathbf{X}_s^P$$

in which

$$\bar{c}_s^W = \left[\hat{c}_s^W \Big/ \sum_{s=1}^{M} \hat{c}_s^W\right]$$

$$\hat{c}_s^W = \frac{\max\limits_{1\leq k\leq M}\left(f(\mathbf{X}_k^P)\right) - f(\mathbf{X}_s^P)}{\max\limits_{1\leq k\leq M}\left(f(\mathbf{X}_k^P)\right) - \min\limits_{1\leq k\leq M}\left(f(\mathbf{X}_k^P)\right) + \varepsilon} \qquad , s = 1,2,\dots,M$$

here, M is the population size, $\hat{c}_i^W$ is the parameter presents the effect of ith agent on the weighted agent. *f*(.) designates the objective function value of the problem. $\max\limits_{1\leq k\leq M}\left(f(\mathbf{X}_k^P)\right)$ and $\min\limits_{1\leq k\leq M}\left(f(\mathbf{X}_k^P)\right)$ are the worst and best agents' objective values, respectively. ε is defined as a small positive number to prevent from a probable division by zero case which is taken as 0.0001.

4. Numerical solution

The proposed optimization model is tested on SEGS LS-2 PTC module [17]. The attained theoretical results are validated by comparing them with the practical data reported by Sandia National Laboratory (SNL). The tests were performed with water and Syltherm800 oil. The properties of the considered collector is presented in Table 1.

The numerical computations are implemented using MATLAB® programming platform in a computer system equipped with an intel-i7™ CPU and 12 MB of installed RAM.

The optimal results utilizing the ISA are reported comparatively and given in Table 2. The percentage error for the calculated temperature difference and thermal efficiency is also reported.

The design variables (i.e. absorber temperature and glass cover temperature) are selected during the optimization process. Tracing the value of the design variables are given in Fig. 2. According to this figure, the optimal state is caught nearly in 15 Iterations.

The population size of the optimization process is 10. The convergence of the two components of the objective function (i.e. $Q_{con\text{-}env}$ and $Q_{a\text{-}ci}$) during the optimization process is shown in Fig. 3. As can be seen from the figure, the difference of these two components is converged to its minimum state (3E-10) nearly after 1500 objective function evaluations.

Table 1 Geometric properties of the PTC

Property	Value
Absorber length (l_a)	7.8 m
Collector width (w_a)	5 m
Focal distance (f)	1.84 m
Absorber Emittance (ε_a)	0.14
Glass cover Emittance (ε_c)	0.88
Glass cover transmittance (τ)	0.95
Absorber absorptance (α)	0.96
Reflection surface reflectivity (ρ)	0.93
Incident angle modifier (F_a)	1
Shape factor (γ)	0.92
Inclination angle (θ)	0.0
Glass cover outer diameter (D_{co})	115 mm
Glass cover inner diameter (D_{ci})	109 mm
Absorber tube outer diameter (D_{ao})	70 mm
Absorber tube inner diameter (D_{ai})	66 mm
Glass cover conductivity (k_c)(DURAN glass @ 90 ^{o}C)	1.2 W/m K

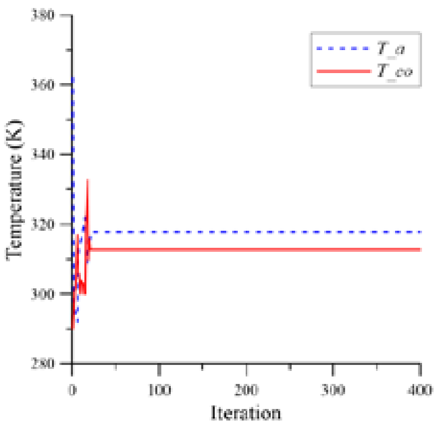


Fig. 2. Tracing the design variables during the optimization process

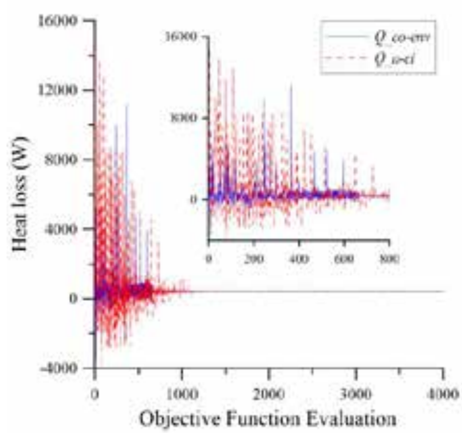


Fig. 3. Tracing the objective function component differences during the optimization process.

Table 3 Comparison of obtained optimal results and reported experimental and theoretical results

Case	Fluid	I_b (W/m^2)	Flow rate (l/min)	v_w (m/s)	T_{amb} (°C)	T_{inlet} (°C)	SNL exp. Results [17]		Garcia Model [18]		iPSO [9]		ISA Current study	
							ΔT(°C)	η_{th} (%)	ΔT(°C)	η_{th} (%)	ΔT(°C)	η_{th} (%)	ΔT(°C)	η_{th} (%)
1	Water	807.9	18.4	1.0	15.8	18.3	17.83	72.63 ± 1.91	17.79	72.39	17.76	73.11	17.75	73.02
2	Oil (syltherm800)	933.7	47.7	2.6	21.2	102.2	21.8	72.51 ± 1.95	21.25	70.69	21.49	72.41	21.51	72.39
3	Oil (syltherm800)	968.2	47.8	3.7	22.4	151.0	22.3	70.90 ± 1.92	22.94	70.13	22.21	71.94	22.22	71.52
4	Oil (syltherm800)	982.3	49.1	2.5	24.3	197.5	22.0	70.17 ± 1.81	21.67	69.32	21.95	71.39	21.95	71.39
5	Oil (syltherm800)	909.5	54.7	3.3	26.2	250.7	18.7	70.25 ± 1.90	18.19	68.26	18.59	70.16	18.62	69.91
6	Oil (syltherm800)	937.9	55.5	1.0	28.8	297.8	19.1	67.98 ± 1.86	18.85	67.40	19.01	68.72	19.01	68.69
7	Oil (syltherm800)	880.6	55.6	2.9	27.5	299.0	18.2	68.92 ± 2.06	17.61	67.08	17.89	68.71	17.95	68.72
8	Oil (syltherm800)	903.2	56.3	4.2	31.1	355.9	18.1	63.82 ± 2.36	18.40	65.19	18.32	64.52	18.27	64.02
9	Oil (syltherm800)	920.9	56.8	2.6	29.5	379.5	18.5	62.34 ± 2.41	18.83	63.84	18.63	62.63	18.70	63.05
									Average Error **1.99%**	Average Error **1.80%**	Average Error **0.79%**	Average Error **0.78%**	Average Error **0.74%**	Average Error **0.72%**
									Max Error **3.35%**	Max Error **2.74%**	Max Error **1.73%**	Max Error **1.71%**	Max Error **1.39%**	Max Error **1.71%**

5. Conclusion

In the current study a constrained optimization model is presented for one-dimensional heat transfer analysis of a receiver tube in a parabolic trough collector (PTC). For this target, the thermal balance in the receiver tube is analyzed with the help of thermal resistance method. A series of governing equations for determining of the heat losses at the nodal points are written. To solve these equation, the problem is converted to a single objective constrained optimization problem. The difference between the heat losses considered as the objective function and minimized. The temperature of the absorber tube and glass cover are assumed as the decision variables of the optimization problem. The proposed optimization problem is solved via the recently developed Interactive Search Algorithm (ISA) method. The obtained results are compared with those are experimentally carried out and reported by SNL and other theoretical models. According to given results, unlike the gradient-based methods, the problem is not depended on initial condition and process can start from any initial arbitrary condition. Since, attained outcomes are in good agreement with the experimental results of SNL, the utilized method is compatible with optimization problem. Also, based on the reported percentage errors, ISA presents more accurate numerical achievements in comparison with other models reported in the technical literature. Consequently, the proposed model and optimization technique is applicable for solving the PTC thermal model in distinct conditions.

References

[1] Daniel P., Joshi Y., Das A.K., "Numerical investigation of parabolic trough receiver performance with outer vacuum shell", Solar Energy, doi: https://doi.org/10.1016/j.solener.2011.04.032, Vol. 85, No. 9, pp. 1910-4. 2011

[2] Padilla R.V., Demirkaya G., Goswami D.Y., Stefanakos E., Rahman M.M., "Heat transfer analysis of parabolic trough solar receiver", Applied Energy, doi: https://doi.org/10.1016/j.apenergy.2011.07.012, Vol. 88, No. 12, pp. 5097-110. 2011

[3] Yılmaz İ.H., Söylemez M.S., "Thermo-mathematical modeling of parabolic trough collector", Energy Conversion and Management, doi: https://doi.org/10.1016/j.enconman.2014.09.031, Vol. 88, No. pp. 768-84. 2014

[4] Hachicha A.A., Rodríguez I., Capdevila R., Oliva A., "Heat transfer analysis and numerical simulation of a parabolic trough solar collector", Applied Energy, doi: https://doi.org/10.1016/j.apenergy.2013.04.067, Vol. 111, No. pp. 581-92. 2013

[5] de Risi A., Milanese M., Laforgia D., "Modelling and optimization of transparent parabolic trough collector based on gas-phase nanofluids", Renewable Energy, doi: https://doi.org/10.1016/j.renene.2013.03.014, Vol. 58, No. pp. 134-9. 2013

[6] Mohammad Zadeh P., Sokhansefat T., Kasaeian A.B., Kowsary F., Akbarzadeh A., "Hybrid optimization algorithm for thermal analysis in a solar parabolic trough collector based on nanofluid", Energy, doi: https://doi.org/10.1016/j.energy.2015.01.096, Vol. 82, No. pp. 857-64. 2015

[7] Duffie J.A., Beckman W.A. Solar Engineering of Thermal Processes. 4 ed. New York: Wiley; 2013.

[8] Çengel Y.A., Ghajar A. Heat and mass transfer: fundamental & applications. 5th ed. 2 Penn Plaza, New York, NY 10121: McGraw-Hill Education; 2014.

[9] Moloodpoor M., Mortazavi A., Ozbalta N., "Thermal analysis of parabolic trough collectors via a swarm intelligence optimizer", Solar Energy, doi: https://doi.org/10.1016/j.solener.2019.02.008, Vol. 181, No. pp. 264-75. 2019/03/15/2019

[10] Al-Kayiem H.H., K.V S., Gilani S.I.U.-H., "Mathematical analysis of the influence of the chimney height and collector area on the performance of a roof top solar chimney", Energy and Buildings, doi: https://doi.org/10.1016/j.enbuild.2013.09.021, Vol. 68, No. pp. 305-11. 2014

[11] Mokhtari A., Yaghoubi M., kanan P., Vadiee A., Hessami R. thermal and optical study of parabolic trough collectors of shiraz solar power plant. Third International Conference on Thermal Engineering: Theory and Applications Amman, Jordan2007. p. 65-70.

[12] Mortazavi A., Toğan V., "Simultaneous size, shape, and topology optimization of truss structures using integrated particle swarm optimizer", Structural and Multidisciplinary Optimization, doi: https://doi.org/10.1007/s00158-016-1449-7, Vol. 54, No. 4, pp. 715-36. 2016

[13] Mortazavi A., Toğan V., Moloodpoor M., "Solution of structural and mathematical optimization problems using a new hybrid swarm intelligence optimization algorithm", Advances in Engineering Software, doi: https://doi.org/10.1016/j.advengsoft.2018.11.004, Vol. 127, No. pp. 106-23. 2019/01/01/2019

[14] Mortazavi A., "Size and layout optimization of truss structures with dynamic constraints using the interactive fuzzy search algorithm", Engineering Optimization, doi: 10.1080/0305215X.2020.1726341, Vol. No. pp. 1-23. 2020

[15] Mortazavi A., Toğan V., Daloğlu A.T., Nuhoglu A., "Comparison of Two Metaheuristic Algorithms on Sizing and Topology Optimization of Trusses and Mathematical Functions", Gazi University Journal of Science, doi: Vol. 31, No. 2, pp. 416-35. 2018

[16] Mortazavi A., Toğan V., Nuhoğlu A., "Interactive search algorithm: A new hybrid metaheuristic optimization algorithm", Engineering Applications of Artificial Intelligence, doi: https://doi.org/10.1016/j.engappai.2018.03.003, Vol. 71, No. pp. 275-92. 2018/05/01/2018

[17] Dudley V.E., Kolb G.J., Mahoney A.R., Mancini T.R., Matthews C.W., Sloan M., et al. Test results: SEGS LS-2 solar collector. Sandia National Laboratories; 1994.

[18] García-Valladares O., Velázquez N., "Numerical simulation of parabolic trough solar collector: Improvement using counter flow concentric circular heat exchangers", International Journal of Heat and Mass Transfer, doi: https://doi.org/10.1016/j.ijheatmasstransfer.2008.08.004, Vol. 52, No. 3-4, pp. 597-609. 2009

Vehicle Classification and Tracking Using Convolutional Neural Network Based on Darknet Yolo with Coco Dataset

Ahmet Dogan [1], Ali Okatan [2], Ali Cetinkaya [*, 3, 4]

[1] Computer Engineering, Faculty of Engineering and Architecture, Istanbul Gelisim University, 34310, Istanbul, Turkey, ahmetdogan@outlook.in,
ORCID ID: 0000-0001-5723-9029

[2] Software Engineering, Faculty of Engineering, Istanbul Aydin University, 34295, Istanbul, Turkey, aliokatan@aydin.edu.tr,
ORCID ID: 0000-0002-8893-9711

[*, 3] Technology Transfer Office Application and Research Centre, Istanbul Gelisim University, 34310, Istanbul, Turkey, alcetinkaya@gelisim.edu.tr,
ORCID ID: 0000-0003-4535-3953

[4] Electronics and Communication Engineering, Faculty of Engineering, Kocaeli University,41380, Kocaeli, Turkey,

ORCID ID: 0000-0003-4535-3953

Abstract

In this study, vehicle classification based on convolutional neural network has been examined with fully connected layers to apply on highways. With the convolutional neural network, the classification of the vehicles in the image has been made through video images. The accuracy of the classification has been calculated. Convolutional neural networks are multilayer neural networks. Each layer consists of many layers that can be trained. This allows the network to learn better. As the number of layers gets higher the time needed to train the network increases. An open-source Yolo algorithm has been used to provide rapid vehicle detection via video streams. Thanks to this study, the classification of the vehicles present in the inner city and highways is made. Also, the vehicles in the same class are counted. In order to increase the success rate in the experiments carried out, the coco data set is used. In the study, experiments are carried out with video images of 4 vehicle classes taken over 7 different media. 78.75% success rate and 21.25% error rate in the working truck class, 78.08% success rate and 21.92% error rate in the bus class, 80.70% success rate and 19.30% error rate in the motorbike class, car class It has a success rate of 86.21 and an error rate of 13.79%.

Keywords: Convolutional Neural Network, Fully Connected Layers, Vehicle Detection, YOLO.

1. Introduction

With today's technological developments, smart transportation addresses issues such as technology, infrastructure, energy and health as a priority. Digital image processing and artificial neural networks should be used together to perform vehicle classification. The reason for using numerical data is because computer systems only work with numerical data. In this case, video images are transferred numerically to the computer environment so that artificial neural networks can work. In order for Video images to be transferred numerically, digital image processing technique must be used.

Thanks to the interfaces provided by programming languages, it supports Microsoft Windows, Linux and Mac OS mobile operating systems in desktop operating systems and Android and IOS platform in mobile operating systems [1,3]. A deep Convolutional Neural Network (CNN) system pre-trained on data is used as a feature extraction tool on the image. The study is performed well both in terms of accuracy and computational time [4]. Digital image processing is the process of obtaining information on images obtained from the image source using various techniques. Application areas; systems developed on autonomous vehicles, military defense and weapons systems, imaging devices used in health care and image processing techniques are used in satellite image retrieval operations [5]. Millions of combinations of biological nerve cells make up the nervous system. Artificial neural networks have been developed using biological nerve cells [6]. The human brain can successfully apply learning, association, classification, generalization, feature, determination and optimization [7]. In some studies, a CNN model with different architectures is created by using the learned parameters over the data set [8,10]. CNN is used in applications such as autonomous vehicle design, detection of logos found in vehicles, and detection of vehicles located in the parking lot [11,15].

Processing images is simple and easy with The Yolo algorithm. Yolo has an opening referred to as You Only Look Once. The expression translated as look only once indicates that the definition can be made after the image passes through the algorithm only once. The Yolo algorithm has an architecture that is the fastest in detecting objects [16]. The number of convolution layers, which are 24 in the Yolo algorithm, is 9 in Fast Yolo and less filters are used in the convolution layers. Except for the size of the quick version created, all training and testing parameters are the same with the normal version [17,19]. A method has been proposed to improve the accuracy of a real-time detection algorithm. In this method, it is aimed to model Yolo with the Gaussian parameter and to redesign the loss function. The developed method has been shown to be able to cope with incorrect positions in autonomous driving controls compared to a traditional Yolo. It has been shown to have a faster detection capability of 42 FPS (frame per second) with the proposed algorithm [20]. The developed real-time object detection detector Yolo has been deepened. Design changes have been made to make it better. The scope of the network included in the Yolo model has been enlarged and retrained. It has been found that the new model works three times faster than the old versions. The success of the system is measured by testing the working accuracy of the model with object recognition models such as RetinaNet [21]. Applications such as automatic license plate recognition and real-time vehicle classification are implemented with the Yolo algorithm [22,23].

2. Material and Method

2.1. Artificial Neural Network

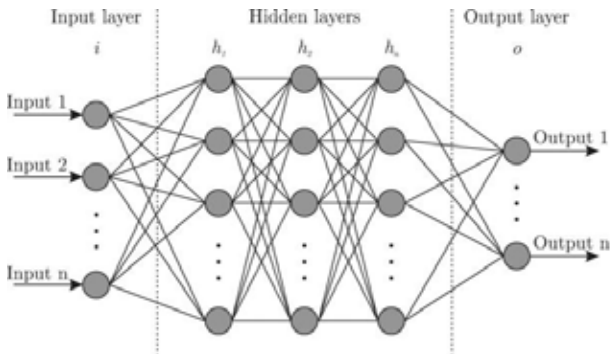


Fig. 1. Artificial Neural Network Architecture

Artificial neural networks are computer systems designed to perform the learning process, which is the characteristics of the brain that are the center of the human body and take the human brain as an example. Figure 1 shows the structure of cells located on the artificial neural network [27]. The artificial nerve cell consists of input, weight, activation function, collection function and output elements. Inputs are information necessary for the operation of the network to the artificial nerve cell. Learning the artificial neural network is determined by examples for the desired state.

$$Net = \sum_{i=0}^{n} x_i w_i$$

In the formula expressed by equation (1) above, "x" value indicates the inputs, "w" value weights, "n" value indicates the number of inputs to the artificial neural cell.

$$Net = \prod_{i=0}^{n} x_i + w_i$$

$$Net = Max(X_i W_i) \mid Net = M$$

With the equations (2) and (3) given above, different addition functions are included. In the first addition function, the net input is calculated by adding and multiplying the input and weight values. The second and third addition functions have a similar structure. Of the last values obtained by multiplying the input and weight values, the second addition function uses the largest one, and the third addition function uses the smallest value as net input. In artificial neural network models, there must be a function in which the derivative of the activation function can be obtained. The activation function commonly used in artificial neural network models is the sigmoid function [28].

$$f(Net) = \frac{1}{1+e^{-Net}}$$

The net value expressed by Equation (4) above is the net input value obtained using the sum function. Figure 2 shows the graph of the sigmoid function [6]. It sends a value between 0 and 1 to the output unit by passing the Net input value through the sigmoid function.

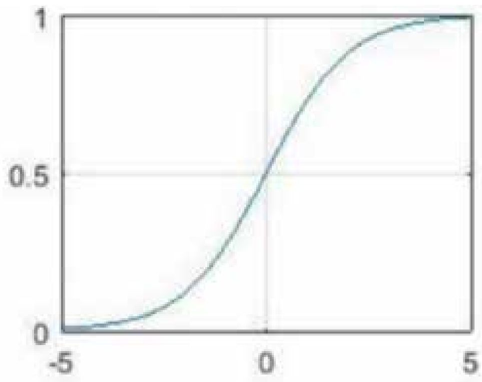


Fig. 2. Sigmoid Function Plot

2.2. Convolutional Neural Networks (CNN)

Convolutional neural networks are a feed forward neural network built on a multi-layer sensor, a model of artificial neural networks. It is used for image and video applications as it is affected by visual perception. Convolutional neural networks are systems formed by the combination of multilayer perceptron layers for convolution, pooling and classification. In order for the convolutional neural network to form, the convolution and pooling layers must have an order. Convolutional neural networks are generally in 5 layers and after each convolution layer, the transition to the pooling layer is provided by using the activation function. These methods are Average Pooling, Max Pooling and Sum Pooling. The leveling layer is located after the pooling layer.

The activation function determined for the convolutional neural network performs its operations in this layer. Relu function is used as activation function. The main reason for choosing this function is that the network performs the best during training. Equation (8) expresses this formula. The values in the property map are passed through the Relu function one by one, and the negative ones are changed to 0. Pooling Layer In convolutional neural networks, the pooling layer comes after the nonlinear layer.

$$f(n) = \max(0, n)$$

2.3. Darknet Yolo

Darknet Yolo algorithm is capable of processing images at a rate of 45 frames per second. Fast Yolo, which is a small version in terms of size, can process 155 frames per second, although it has 2 times other algorithms capable of real-time object detection. Previously developed models in object detection reuse the fully connected layer where the classification process takes place in order to realize the detection. The Yolo algorithm looks at object detection as a single regression problem, directly from the pixels that make up the image, to the coordinates of the bounding boxes and class probabilities.

$$Tensor = S\ x\ S\ x\ (B$$

Using the above given equation (9) and the tensor formula, after determining the tensor required in the output section, the convolutional neural network is formed. The S parameter on Equation 9 represents the spatial size of the input image, the B parameter the number of bounding boxes, and the C class number [19].

2.4. Description of The Data Set Used in The Study

Whether the COCO data set is statistically different in the methods evaluated for both classification and object detection, what the system is learned from images, and the methods have been analyzed [24]. The PASCAL VOC dataset contains standard image datasets for object class recognition. It enables the evaluation and comparison of different methods over a common tool set to access data sets and explanations [25]. YOLO9000, a real-time object detection system that detects more than 9000 objects in real time, has been introduced with YOLOv2 and YOLO9000. Improvements have been made in the study with the information obtained in the YOLO detection method. With the improved model, it has been found that it is better than detection algorithms such as YOLOv2, PASCAL VOC and COCO. WordTree is used by combining a common optimization technique with information from various sources to get data at the same time in ImageTree and COCO [26].

In this study, convolutional neural networks, deep learning techniques, yolo algorithm and current data sets, and convolutional neural network-based instrument classification are examined. With this system, a software that can be integrated into city systems is developed. With this software, operations are performed on video images, classification of cars, buses, trucks and motorcycles and counting of vehicles passing over that road are made. During the application, it is ensured that the tools are separated from each other by the software and their number is kept according to their type.

2.5. Flowchart of The Study

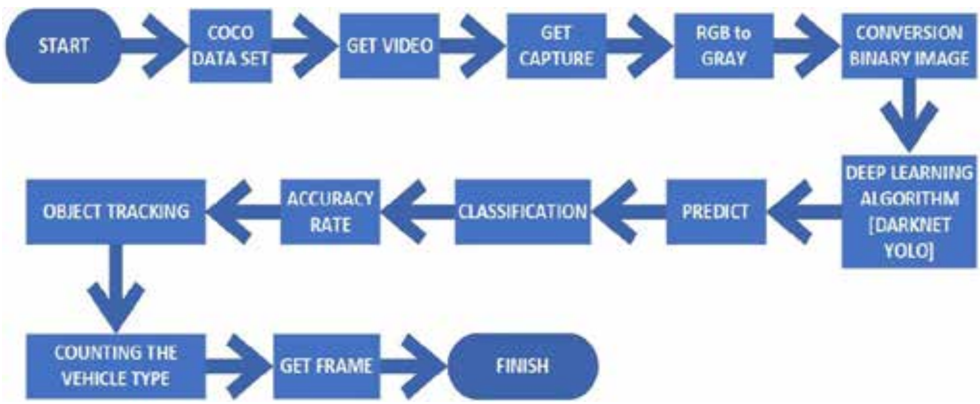


Fig. 3. Flowchart of The Study

Figure 3 shows the flow chart of the algorithm followed for the CNN-based vehicle classification application. The algorithm consists of seven main stages. These; image pre-processing, feature extraction, deep learning algorithm, COCO data set, classification, accuracy estimation, object tracking and counting. Dividing the video into frames given to the software algorithm in the image preprocessing stages (video reading, taking pictures from the video, taking parts from the pictures and turning the picture into gray tone) is converted to gray tone in order to extract the features of the objects in these frames. Then the image is made ready to be sent to the deep learning algorithm. Darknet Realizes the training of the network to determine the bounding boxes of objects on the image using YOLO. After the training is performed, it is sent to the deep learning algorithm using the found weight values and configuration settings. The COCO data set, which is ready as a reference, is used for the deep learning algorithm to prepare the classification and prediction processes of the network.

Image used in classification processes; After passing through the deep learning algorithm, the training of the network is performed and the class of the objects to be detected is determined. More than one class can be detected on an image at the same time. After the classification process takes place, the accuracy rate of the object belonging to the detected class is found and the margin of error of the algorithm and the realized software is determined.

After the margin of error is determined, the object tracking is done and counting is performed. The counting process is performed by using the center coordinate values of the bounding boxes of the objects on the image, after all results are produced regarding the objects on the frame in the deep learning algorithm.

The Darknet YOLO algorithm, developed using the CNN structure, has been shown to correctly classify the desired tools. The vehicles used as classification are car (car), motorbike (motorcycle), bus (bus) and truck (truck). Different video images taken from traffic are used to better examine the results. An attempt is made to reduce the margin of error in order to correctly identify the tools detected by the Yolo algorithm. In order to reduce the margin of error, the software has been tested using different versions of the Yolo algorithm developed by Darknet. After the vehicle classification process is performed, the vehicles in the traffic video image are taken into the bounding boxes, and an accuracy value between 0-1 and the name of the class associated with the vehicle detected on the bounding boxes is added. The original version of the traffic video image is preserved and the vehicle classification process is carried out. After the traffic video image is divided into frames and processed, a new video image is obtained by combining it again.

3. Findings and Discussion

During the application, a number is assigned to the classified vehicle and remained defined until the identified vehicle is displayed. Different numbers are defined for each vehicle on the application. In the software, the classification of the vehicles is defined with frames of different colors. The determined type of the defined vehicle and the calculated accuracy rate are located in the upper left part of the frames.

Table 1. Colors and Shades Used in The Application

Vehicle Type	Color	Red	Green	Blue
Car	Blue	88	218	245
Truck	Purple	180	154	188
Motorbike	Red	255	97	0
Bus	Orange	249	107	44
Numbers	Green	15	246	17
Writting Tone	Yellow	255	255	0
Paint Tone	Orange	255	127	39

Different colors are used in the classification of the vehicles while working with the video in the application. These colors are given in table 1 with their color codes. In the classification process, the colors that make up the colors of the frames are determined as “Blue” for the car, “Purple” for the truck, “Red” for the motorcycle and “Orange” for the bus. Tests are carried out using traffic videos obtained from different location information, and the obtained results revealed the accuracy and effect on the topics in the literature with YOLO using the structure of convolutional neural networks. The algorithm used has an open source structure and the convolutional neural network structure is flexible.

The designed system has been tested on a 6GB GTX 1060 GPU. The videos taken within the scope of the application are recorded with a 16 MP resolution camera. License plates, advertisements and human faces in the screenshots taken from the applications are painted with orange color in order to protect personal data. Applications using the Darknet YOLO algorithm have been tested in seven different locations in Istanbul. Screen shots of the application are given on the metrobus route, intersections, tunnel exits and joining roads. Figure 4- (a) and (b) 25s and 60s on first video, (c) second video 31s, (d) third video 30s, (e) fourth video 30s, (f) fifth video 30s, (g) sixth video 61sec and (h) seventh video is 28sec views.

Fig. 4. Images of Applications on Traffic Video Images

4. Conclusions and Recommendations

In this study, a study is conducted to classify vehicles in urban and extra-urban roads in city systems using the YOLO algorithm, which has a convolutional neural network structure of the Darknet library. Vehicle classification process is carried out using traffic videos. The success criteria of the sources consisting of videos with different file lengths and recording times are given in Table 2 and Table 3. In Table 2, the identification rates of the vehicles classified in the system are given in the tests. Results are given in percentage (%) with minimum and maximum accuracy values.

Table 2. Accuracy results

Video No.	Vehicle List			
	Truck	Bus	Motorbike	Car
Video 1	50%-90%	79%-91%	50%-80%	75%-99%
Video 2	66%-95%	61%-80%	57%-92%	73%-98%
Video 3	90%	64%	77%-89%	76%-95%
Video 4	69%-85%	-	-	72%-99%
Video 5	71%-89%	64%-92%	91%	88%-89%
Video 6	60%-90%	64%-98%	-	75%-99%
Video 7	-	90%	90%	70%-99%

When the accuracy estimates given in Table 2 are analyzed within the system, errors are detected in the "Bus" and "Truck" classifications on "video 3" and "Bus" classifications on "video 7". It is determined that the most important error occurred is due to the lack of "Bus" and "Truck" data, which are among the vehicle types in the data set used. To correct these errors, different bus and truck images should be tagged in the data set and the model should be retrained. The total number of vehicles classified in the application is given in table 3.

Table 3. Application results

Video No.	Vehicle List				Video Length (Seconds)
	Truck	Bus	Motorbike	Car	
Video 1	10	11	3	316	60
Video 2	9	4	2	175	31
Video 3	6	2	5	109	31
Video 4	1	0	0	60	30
Video 5	2	3	1	110	31
Video 6	2	2	0	22	61
Video 7	0	1	2	34	31

In this developed application, the attributes of the vehicle classes are used in the coco data set. In this way, the estimation rate in the vehicle classification process carried out in the study is increased. With the measurements made within the experiments using the method followed in the study, 78.75% success rate and 21.25% error rate in the truck class, 78.08% success rate and 21.92% error rate in the bus class, 80.70% success rate in the motorcycle class and 19 30% error rate, 86.21% success rate and 13.79% error rate in the car class are measured. While determining the success and error results within the scope of the study, it is calculated by taking the average of the values on table 2 and table 3.

With the method developed in our study, the number and type of vehicles passing through certain roads can be determined, and information on how often the road maintenance services of highways and municipalities will be performed. In order to increase the classification accuracy of vehicles in subsequent studies, it should be ensured that vehicles with subclasses such as minibuses and vans are included in the data set. This addition can be added to the YOLO algorithm, which enables the training of the data set to make more precise classifications. Real-time imaging systems can be used instead of pre-recorded traffic videos. In order to classify the vehicles in the images taken from the cameras, operations can be performed with high-spec computers. With the license plate recognition application that can be integrated into vehicles belonging to the specified class using real-time images, control can be provided on urban and extra-urban roads.

Acknowledgements

Author may acknowledge to any person, institution or department that supported to any part of study.

References

[1] E. Erişti, "Görüntü İşlemede Yeni Bir Soluk, OPENCV", 2010.

[2] OpenCV library document, https://opencv.org/, (Access Date: 15.02.2021).

[3] J. Howse, OpenCV computer vision with python, 2013.

[4] N. Ammour, H. Alhichri, Y. Bazi, B. Benjdira, N. Alajlan and M. Zuair, "Deep learning approach for car detection in UAV imagery", Remote Sensing, 2017.

[5] M. Y. Demirci, "Gerçek zamanlı bir görüntü işleme sisteminin model tabanlı olarak tasarlanması", MS thesis, 2018.

[6] E. Öztemel, Yapay sinir ağları, Papatya Yayıncılık, (2012).

[7] H. Ergezer, M. Dikmen, and E. Özdemir. "Yapay sinir ağları ve tanıma sistemleri", PiVOLKA 2.6, 2003.

[8] B. Hicham, A, Ahmed, and M. Mohammed, "Vehicle type classification using convolutional neural network", IEEE 5th International Congress on Information Science and Technology (CiSt), IEEE, 2018.

[9] M. Shahverdy, M. Fathy, R. Berangi, and M. Sabokrou, "Driver behavior detection and classification using deep convolutional neural networks", Expert Systems with Applications, 2020.

[10] R. Xin, J. Zhang, and Y. Shao, "Complex network classification with convolutional neural network", Tsinghua Science and technology, pp. 447-457, 2020.

[11] B. Rajesh, D. Ramakrishna, A. R. Raju, and A. Chavan, "Object Detection and Classification for Autonomous Vehicle. In Journal of Physics: Conference Series", Vol. 1817, No. 1, pp. 012004, March 2021.

[12] J. Cai, H. Chen, L. Kneip, and S. Schwertfeger, "Improving CNN-based Planar Object Detection with Geometric Prior Knowledge", 2019.

[13] E. C. Nwankpa, W. Ijomah, A. Gachagan, and S. Marshall, "Activation Functions: Comparison of Trends in Practice and Research for Deep Learning", 2018.

[14] H. Xie, Q. Wu, B. Chen, Y. Chen, and S. Hong, "Vehicle Detection in Open Parks Using a Convolutional Neural Network", Sixth International Conference

on Intelligent Systems Design and Engineering Applications, DOI 10.1109/ISDEA.2015.233, 2015

[15] Y. Kutlu and Y. Camgözlü, "Detection of coronavirus disease (COVID-19) from X-ray images using deep convolutional neural networks", Natural and Engineering Sciences, DOI: 10.28978/nesciences.868087, Vol. 6, No. 1, pp. 60-74, January 2021.

[16] Y. Akbulut and R. Khalaf, "Smart Arms Detection System Using YOLO Algorithm and OpenCV Libraries", Turkish Journal of Science and Technology, Vol. 16, No. 1, pp. 129-136, March 2021.

[17] J. Redmon, S. Divvala, R. Girshick, and A. Farhadi, "You Only Look Once Unified Real Time Object Detection", 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), DOI: 10.1109/CVPR.2016.91, 2016.

[18] M. F. İnkaya and H. Gürkan, "A YOLOv3-Based Smart City Application for Children's Playgrounds", Journal of Innovative Science and Engineering, DOI: 10.38088/jise.813664, Vol. 5, No. 1, pp. 25-40, June. 2021.

[19] S. Vasavi, N. Kanthi Priyadarshini, and K. Harsha Vardhan. "Invariant Feature based Darknet Architecture for Moving Object classification", IEEE Sensors Journal, 2020.

[20] J. Choi, D. Chun, H. Kim, and H.-J Lee, "Gaussian YOLOv3: An Accurate and Fast Object Detector Using Localization Uncertainty for Autonomous Driving", Computer Vision and Pattern Recognition, 2019.

[21] J. Redmon, and A. Farhadi, "YOLOv3: An Incremental Improvement", 2018.

[22] R. Laroca, L. A. Zanlorensi, G. R. Gonçalves, E. Todt, W. R. Schwartz, and D. Menotti, "An Efficient and Layout-Independent Automatic License Plate Recognition System Based on the YOLO detector, Computer Vision and Pattern Recognition", 2019.

[23] M. Yıldırım and A. Çınar, "Classification of 40 Different Human Movements with CNN Architectures and Comparison of Their Performance", Turkish Journal of Science and Technology, Vol. 16, No. 1, pp. 103-112, March 2021.

[24] M. Everingham, L. V. Gool, C. I. K. Williams, J. Winn and A. Zisserman, "The PASCAL Visual Object Classes (VOC) Challenge", Int J Comput Vis 88: 303–338, DOI: 10.1007/s11263-009-0275-4, 2010.

[25] The PASCAL Team, Pascal VOC data sets, http://host.robots.ox.ac.uk/pascal/VOC/ (Access Date: 15.02.2021)

[26] J. Redmon and A. Farhadi, "YOLO9000: Better, Faster, Stronger, 2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)", DOI: 10.1109/CVPR.2017.690, 2017

[27] F. Bre, J. M. Gimenez and V. D. Fachinotti, "Prediction of wind pressure coefficients on building surfaces using artificial neural networks", Energy and Buildings, 2018.

[28] M. Yıldırım and A. Çınar, "Classification of 40 Different Human Movements with CNN Architectures and Comparison of Their Performance", Turkish Journal of Science and Technology, pp. 103-112, 16 January 2021.

Application of Artificial Intelligence for the Protection of Consumers on the Cyber-Physical System Platform

G. Moskvins[1], E. Spakovica[1], M. Moskvins[1]

[1]Institute of Energetic, Faculty of Engineering,
Latvia University of Life Sciences and Technologies
logicor@llu.lv

Abstract

The scientific work contains the researches for elaboration of Cyber-Physical System for protection of interests and rights of consumers. Artificial Intelligence on a Cyber-Physical System platform are a new trend in the development in the area of intelligent systems and technologies. More recently, a lot of research activities have been dedicated to the fields of mobile ad hoc network and wireless sensor networks. The cyber physical system and Internet of Things today has emerged as a promising direction to enrich the interactions between physical and virtual worlds. For making a decision reliable on-line information is necessary. The results of researches demonstrate how Cyber-Physical System by using of mobile applications exploit the physical information collected to identify important research challenges related to elaboration of optimal intelligent methods and instruments for consumer protection. On the Artificial Intelligence platform, we have developed the new model of intelligent systems for consumer protection. In the course of the research, statistical models Markov chains and Bernoulli compliance control were developed and tested. The creation of a cloud Internet of Things cyber-physical platform is a link between the physical and virtual world when using mobile e-application in smartphones. As a result, the consumer has a complete set of operational cyber physical and legal information, including information on product safety, food additives, origin of the product, traceability etc., which will enable consumers to independently secure their legal rights to a safe existence.

Keywords: Internet of Things, Cyber - Physical Systems, consumer protection, intelligent sensors, mobile applications.

1. Introduction

A key tool to ensure the flow of information to enabling swift reaction when risks to public health are detected in the food chain is RASFF – the Rapid Alert System for Food and Feed. The *RASFF* was put in place to provide food and feed control authorities with an effective tool to exchange information about measures taken responding to serious risks detected in relation to food or feed [1-3].

Connecting to corporate data, information resources, both internal and external, is the basis for digitalization of products and services of enterprises and institutions, and connecting users interested in the service to information sources and to each other is the basis for digital transformation. Such a comprehensive view of the changes taking place in the business, in management of large organizational systems and consumer protection bodies helps to harness new ideas and opportunities for protection of consumers through the universal digital transformation of goods and services.

The focus of the computing industry is gradually shifting towards *AI*, which is becoming the centrepiece of generating new value not only for business and management, but also for protecting consumer interests. This situation creates the necessity to draw a clear boundary between hardware and software to adapt technological and computing capabilities to various data processing requirements and to integration in *IoT-CPS* and *Android App* [4].

Artificial Intelligence (AI) creates new opportunities for production, management, control and business. In the digital world of continuous transformations, literally everything is subject to change. Also, even *Internet of Things*, which covers critical production and business processes with sensors, changes the traditional understanding. The *IoT* typically includes four main functional - operational components: sensing devices, connectivity, data processing and user interface (Fig.1).

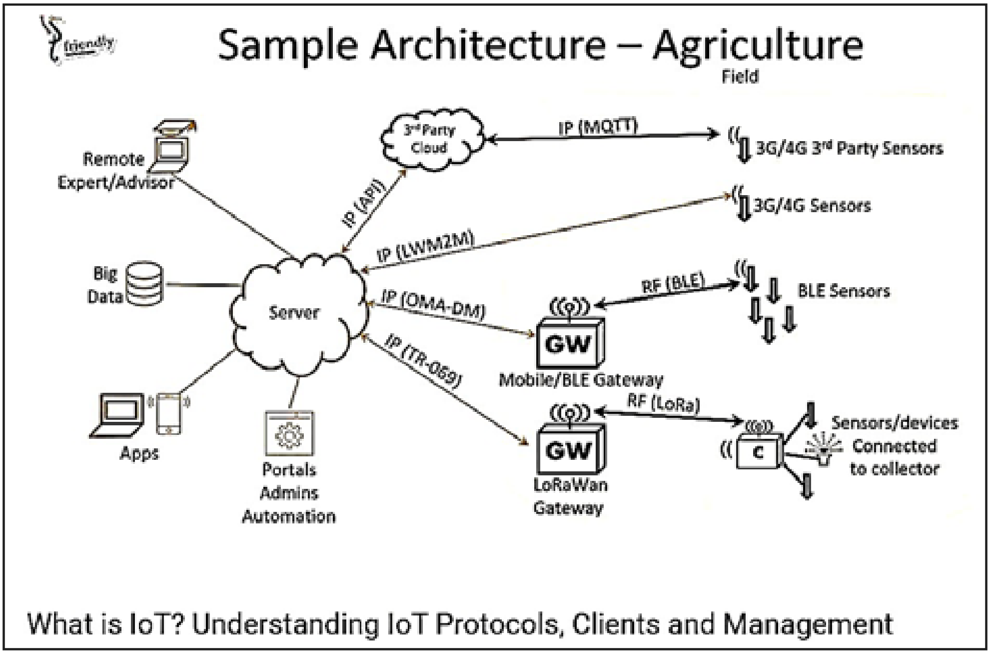
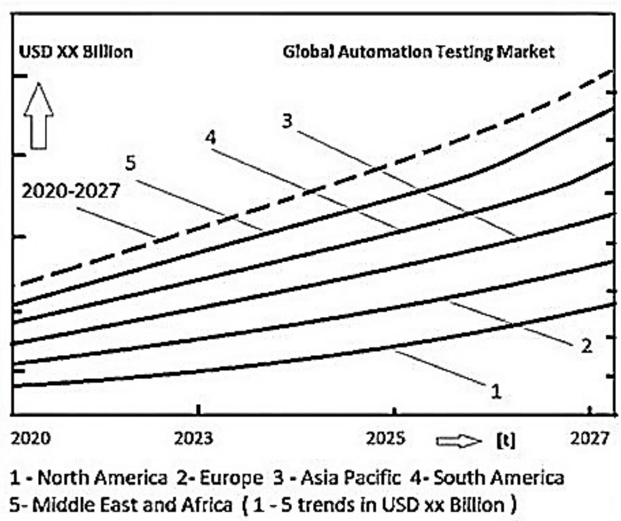


Fig. 1. The IoT scheme for the testing market [adapted from 4]

The model-based approach that we take here and taken in previous studies [5-7] represents different experimental and analytical methods for obtaining valuable perspective on the genesis and characteristics of global evolution trends in the area of consumer protection by means of *IT* technologies and *AI* testers. The primary strength of received data is its ability to precisely capture local conditions, but even the best-available experimental data have limitations, uncertainties, and potential unobserved biases (for example, due to observational procedures, instrumentation type, or siting), as well as highly incomplete spatiotemporal coverage. In contrast, reanalysis products and high-resolution regional models satisfy the need for spatiotemporal continuity and consistency and allow analysis of additional variables, but often underestimate extremes and local uncertainties factors.

The current stage of digitalization, through which the whole world passes, requires new methods for implementing digitalization. Today's universal connectivity is a fundamental thing that, in fact, changes everything in the usual *IT*. Network evolution from 1G to 5G opens a new era in which absolutely all things will be interconnected. In the future automation testing market will grow thanks to advisory and consulting service, planning and consumer service, IoT and CPS support and maintenance, documentation and training, implementation and management, Fig.2.

a)



b)

Fig. 2. The Growth of Global Automation Testing Market (a) and Logicor AFM Tester "Artificial Tongue" (b), [7].

In this work we demonstrate that efforts of better understanding the possibilities of consumers' protection would benefit from further close examination, and improved integration, of experimental data to alleviate model shortcomings - especially in real situations where primary data can vary markedly real data. Therefore, a new *IoT-CPS* modelling–based approaches are fundamentally complementary. It is of utmost importance to design new generation conformity control methods and devices with the application of fractal method, which can effectively work under changing peculiarities operation as well as to adopt themselves to consumer's intellect and to bionic stochastic processes with incompletely systemized parameters and partially formalized requirements in real conditions [7].

The product's properties and registering results of the research, which have to be fixed by means of different measuring devices, can be so significantly different that there can be no essence at all in the indications of the measuring device. The advantage of *IoT-CPS* for the protection of consumers is that instead of a set of crude information, the complex of qualitative knowledge is used. From the practical consumer experience, it is known that not deeply enough processed information rather misinforms, disorientates than informs consumers about risks or defects of the tested products.

On the *IoT-CPS* platform we have developed the new model of intelligent conformity control systems for consumer protection by means of using *AT* and mobile applications. Similar systems implemented on other physical, analytical and programmatic principles are under development, updating and practical testing. For example, EWG's Healthy Living System built consumer IoT guides to help to learn about the hidden health dangers in food, water and everyday products to make better decisions.

For example, since 2004, *EWG's Skin Deep®* cosmetic *IoT* database has helped people protect themselves from potentially toxic chemicals in personal care and beauty products. *Dirty Dozen™ IoT EWG's* 2021 Shopper's Guide Database provides insight into *Pesticides in Produce™,* Tap *Water Database* provides insight into pollutants in local water resources, [9,10], Fig.3.

Fig. 3. Mobile applications «ProdCat» and EWG's Health Living System on the IoT- CPS platform

2. Methods and Instruments

Methods of identifying the properties of products on the *IoT-CPS* platform. The functional structure of Internet of Things *(IoT)* includes sensing device, connectivity, data processing and user interface. A simple form of visualization and presenting ready-to-use knowledge about the properties of the product is very convenient for the protection of consumers. In addition, the geometry images of the conformity created based on cyber-physical platform using physic, metrical and mathematical models contain parameters and main elements of self-similarity and fractality of tested products. Therefore, an elaboration of the library of the standard geometric images automatically lets us recognize and compare different samples of the product being tested. This methodology gives the additional capabilities for automatic analysis of product quality using primary sensing devices on the *IoT-CPS* platform. As the preliminary step of identification, the physic -metrical models of a tested product can be interpreted by using the mathematical and statistical models. During the research, statistical models *Markov chains* and *Bernoulli* compliance control were developed and tested. The research of the stochastic matrix in the task of conformity control is based on the mathematical model of *Markov chain* of the set of measurements of the parameters of the tested product. During conformity control experiments, a certain matrix of critical frequencies of the product quality benchmark and the obtained product quality indicators is calculated. Such a matrix is created for each product sample. In the following, the probability of compliance of the quality parameters for each product assesses standards (sample, benchmark), risk of quality deviation or non-compliance.

The applied sensing methods and instruments for the *IoT-CPS* platform can be E-tongue - potentiometric sensors, measurements of conductivity, volt - ampere - ohm metric, optical sensors, biosensors, AFM-resonance. E-nose: conductivity sensors - MOSFET (metal-oxide-silicon field effect transistor), CP (conducting polymer). Piezoelectric sensors - QMB (quartz crystal microbalance, SAW (surface acoustic wave), optical sensors. The "Artificial Tongue" *(AT)* device is an electronic intelligent instrument, which consists of data acquisition and data analysis systems. Analysts expect that as regulations pertaining to food and agricultural products testing continue to be adopted, the shift towards rapid nonconformity assessment methods will continue. These standards are commonly referred to as commodity standards or standards of identity.

At the preliminary stage various statistical methods were used to create a library of standard images and to identify the inconsistency of the tested samples of standard metric image of the product [5]. In this work, the principal component analysis (PCA) method has been used, for recognition of unknown sample, discrimination and identification discrimination factorial analysis (DFA) method was used, for quality control against reference good product and "good-bad" modelling the SIMCA method as soft independent modelling of class analogy, as well as the least square (PLS) method has been used.

Based on the preliminary studies, the new identification method based on Markov stochastic matrix has been proposed. Markov stochastic matrix property is such that the conditional probability distribution of the state in the future, given the state of the process currently and in the past, is the same distribution as one given the current state. In other words, the past states carry no information about future states. For learning procedure *IoT-CPS* beforehand ask from expert qualitative knowledge about the main physical property of agricultural products. For use of *IoT-CPS* technology it is beforehand necessary to create a “info-quarks” library for the automatic comparison of standard geometrical images with real physical-metrical images of tested product (Fig.4).

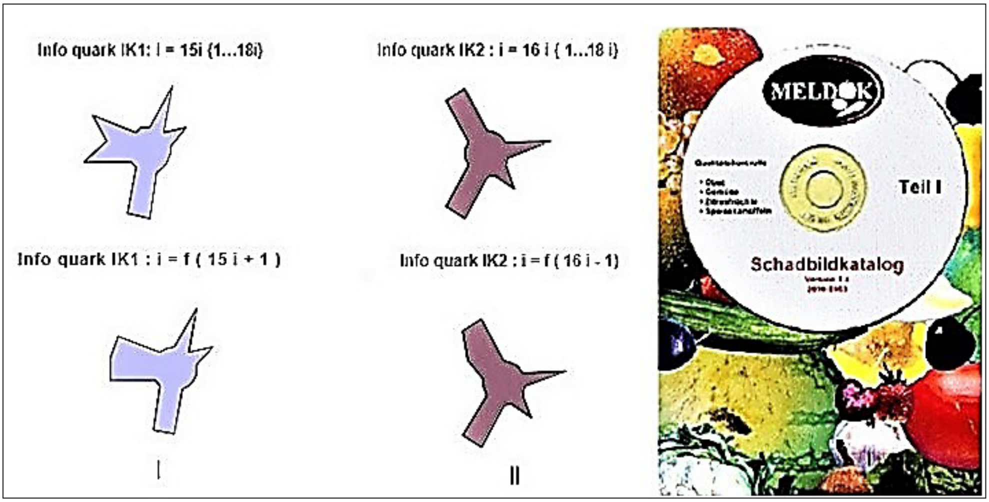


Fig.4. CD (“Schadbildkatalog”, BMVEL) and *Logicor* (LLU) fractal “info-quark” metrical images for the catalogue of *CPS-IoT* library

3. Results and Discussion

The fundamental purpose of a quality control program is to acquire dependable information on all the attributes of a product, which affect its quality. The quality of a fruit and vegetables often is not a linear combination of all measurable quality parameters. This presents a major problem as to how these measurements should be combined into quality indices and grading decisions. The quality of fruits and vegetables is a combination of numerous parameters such as: firmness, acidity, aroma, colour, colour uniformity, bruises, scars, cuts, and presence of soil, size, shape, and insects’ diseases.

The main sensible parameters are specific to the individual fruit and vegetable. Thus, the concept of this work is to develop a system that can classify fruits based upon several of these parameters by using sensor data acquisition. Despite the numerous techniques developed for non-destructive evaluation of the quality, for example, of fruits and vegetables, quality sorting is still primarily based on manual decisions and hand labour.

The experimental research proved that quality and conformity control of agricultural products can be determined by using fractal conformity method *Logicor* and intelligent sensor "artificial tongue" on the *IoT-CPS* platform. The research was mainly focused on developing the classification algorithms. The sensor system that utilizes the imaging system, impact sensor, sensor of electro-conductivity constant and alternating current, electronic chronometer for determination of relaxation time – T_{rel}, as well as the ultrasonic sensor, gauge for measuring an electrical resistance R_a, force gauge and "artificial tongue" device (Fig.5).

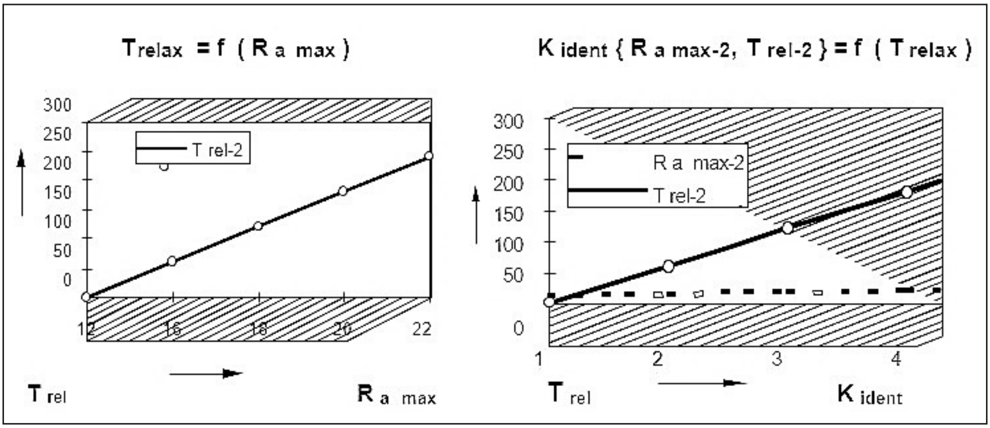


Fig.5. Results of the researches of physical properties for the CPS-IoT platform

Techniques and criteria for choosing the training sets for the classifier have also been developed. A middle training set, in total about 70 dates was required to achieve good conformity classification (93%), Fig.6. The results of calculation are shown in Table 1.

Table 1. The results of calculation the Markov' matrix

Codes name of tested samples	C_1	C_2	C_3	C_4
0	0	3	2325	612
1	0	2	1633	418
3	0	2	1112	396
4	0	2	962	381

In columns $\mathbf{c_1}$ are shown: the results of the conformity control (**0** - the conformity of the sample was established); $\mathbf{c_2}$ - the total number of files containing the sample encoded a priori data. Column c_3 contains the total number of coded metric symbols of the sample that belong to $\mathbf{F(g_{i,j})}$, that is, $\mathbf{c_3 = \Sigma_j | F(g_{i,j}) |}$.

Column $\mathbf{c_4}$ contains the number of coded metric symbols belonging to $\mathbf{F(y^i)}$, that is, $\mathbf{c_4 = | F(y^i) |}$. The row of the table column $\mathbf{c_{1,j}}$ - contains the rank of the number $\mathbf{L_j(F(y^j))}$, which is between the numbers $\mathbf{\{L_i(F(y^j)) | i = 0, 1, 2, 3\}}$. By the term "rank" we mean the number $\mathbf{L_j(F(y^j))}$ between the numbers $\mathbf{\{L_i(F(y^j)) | i = 0, 1, 2, 3\}}$, arranged in ascending order. Let's place the matrix in rows $\mathbf{j = 0, 1, 2, 3}$ by four numbers in each $\mathbf{L_j(F(y^j)) / | F(y^j) |, i = 0, 1, 2, 3.}$

The searched values of column $\mathbf{c_1}$ are located on the diagonal of the stochastic R-matrix. According to $\mathbf{t(x) = argmin_{i=0,...,n-1} L_i(x)}$ we can conclude that $\mathbf{t(F(y^j)) = j}$ only when the rank $\mathbf{L_j(F(y^j)) / | F(y^j) |}$ between the numbers $\mathbf{\{Li(F(y^j)) / | F(y^j) |, i = 0,1,2,3\}}$ is equal to **0**. The rank of $\mathbf{L_i}$ numbers can be defined for each data row:

$$\mathbf{L} = \begin{pmatrix} 2.412346 & 2.507465 & 2.4808653 & 2.4476736 \\ 2.493173 & 2.477468 & 2.5136458 & 2.4812874 \\ 2.499033 & 2.504508 & 2.4754778 & 2.4861542 \\ 2.530487 & 2.522861 & 2.5396557 & 2.5167367 \end{pmatrix} \quad \mathbf{R} = \begin{pmatrix} \mathbf{0} & 3 & 2 & 1 \\ 2 & \mathbf{0} & 3 & 1 \\ 2 & 3 & \mathbf{0} & 1 \\ 2 & 1 & 3 & \mathbf{0} \end{pmatrix}$$

So, if in any row in the column of Table 1 $\mathbf{c_1}$ occurs **0**, then the compliance check pattern is successfully identified. Table 1 shows that all product control samples were found to be compliant. On the Fig.6 we demonstrate the results of *ANN* tests by means of *BPM* and *Logicor* classification method with using correction algorithm.

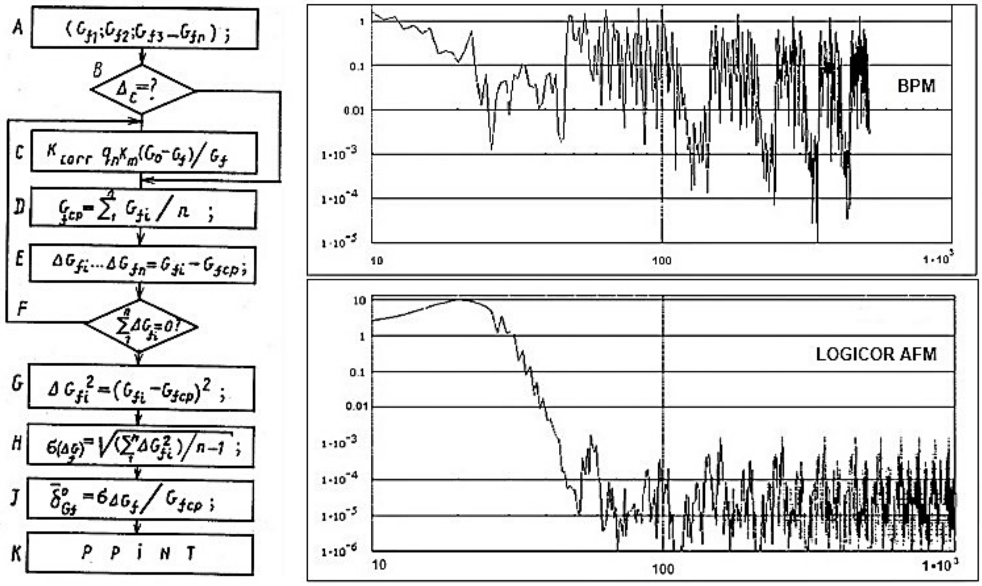


Fig 6. The results of tests BPM and Logicor-AFM classification and correction algorithms for the IoT-CPS platform

The results of experiments confirmed the advantages of the *Markov* chain mathematical model compared to the *Bernoulli* method, because the *Markov* model contains many more *CPS* parameters. The statistical modelling experiment based on the *CPS* platform are confirmed the hypothesis, in the experiments 88% vegetable and 93% fruit samples were unambiguously identified. Techniques and criteria for selecting training sets for the classifier were developed. A middle training set, in total about 200 dates, was needed to achieve good classification. A positive feature of the proposed concept is the development of a library of metric images using the Markov stochastic matrix and a more advanced classification and correction algorithm.

The EU *RASFF* system cannot be considered as a safe consumer protection instrument, as it ignores individual and regional conditions. Food has the greatest impact on human life and health, so special attention is required to address consumer protection opportunities through *CPS-IoT* and mobile application technologies. The consumers must be aware that no state controlling or supervisory institution will solve his problems. Only the consumer's own knowledge and ability to defend themselves with *IoT, CPS* and mobile application technologies help to build an intelligent relationship between the consumer, the manufacturer and the business.

On Fig.7 an example of the elaborated standard metrical images for the *IoT-CPS* platform is shown.

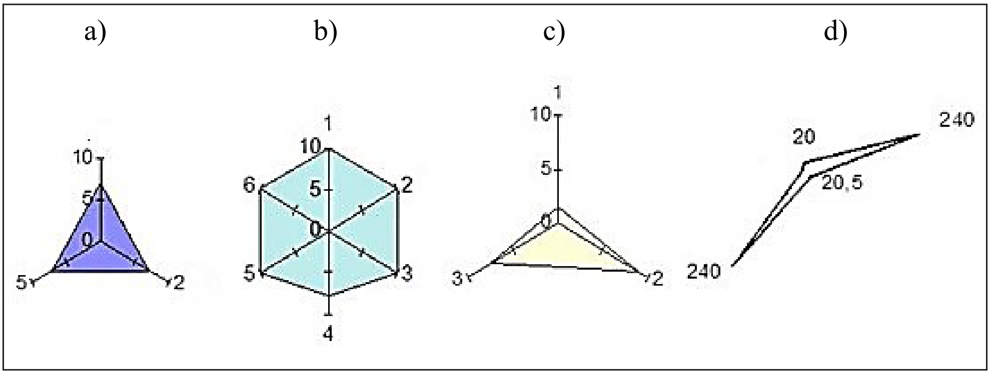


Fig.7. Standard fractal metrical images for the catalogue of library the IoT-CPS platform (example for artificial tongue *Logicor AFM*, F=5 kHz, A=0,5 dB, T=22°C, where a-potato "Mutageni", b - potato "Sarmite", c - deviation of standard image "Mutageni", d - apple "Merrigold" is shown).

4. Conclusions

The creation of a cloud *IoT* cyber-physical platform for consumer protection is a link between the physical and virtual world when using mobile e-application in smartphones. As a result, the consumer has a complete set of operational cyber physical and legal information, including information on product safety, food additives, origin of the product, traceability etc., which will enable consumers to independently secure their legal rights to a safe existence.

Market analysis of possibilities for the development of *IoT*-C*PS* has demonstrated that the automation testing market will grow thanks to advisory and consulting service, planning and consumer service, support and maintenance, documentation and training, implementation and management.

A new *IoT-CPS* consumer's protection *AI* technology has been developed based on results of experimental research and modelling "of intellect of the consumer". Interrogations show positive consumer and business opinions to the new possibility for protection of their rights and interests.

The *IoT-CPS* traceability *AI* system can also further serve as an additional essence that provides consumer and government sectors enabling both to improve its performance.

References

[1] RASFF (https://ec.europa.eu/food/safety/rasff/for_consumers_en), 10.05.2021.

[2] European Food Safety Authority (http://www.efsa.europa.eu/), 10.05.2021.

[3] The Rapid Alert System for Food and Feed - Annual Report 2019, Luxembourg: Publications Office of the European Union, 2020, © European Union, 2020, ISBN 978-92-76-17509-4 ISSN 1830-7302 doi:10.2875/233333 EW-AC-20-001-EN-C.

[4] Data Bridge. https://www.databridgemarketresearch.com, 10.05.2021.

[5] Moskvin G., Spakovica E. Intelligent Technology for the Conformity Assessment of Agricultural Products. Advances in Computer, Information, and Systems Sciences, and Engineering. Hardcover ISBN: 1-4020-5260-X, Springer Berlin-Heidelberg-New York, 2006, pp. 109-114.

[6] Moskvins G. Researches of intelligent measuring devices for consumer protection // 19th International scientific conference "Engineering for rural development": proceedings, Jelgava, Latvia, May 20 - 22, 2020, Vol. 19, 1840.-1845.lpp. ISSN 1691-5976.

[7] Moskvin G., Spakovica E., Moskvin A., Development of Intelligent Technologies and Systems in Agriculture. 7th International Scientific Conference «Engineering for Rural Development» LLU, TF, 29.-30.05.2008, Jelgava, p. 208-215.

[8] Moskvins G. Intelligent chemical and physical conformity assessment tools. ISBN 978998488533/Catalogue of the 7th International Invention and Innovation Exhibition MINOX 2018/CONNECT Latvia. Riga, 2018, p.14-15.

[9] Dirty Dozen™ EWG's 2021 Shopper's Guide to Pesticides in produce™, https://www.ewg.org/foodnews/dirty-dozen.php, 10.05.2021.

[10] ProdCat,https://play.google.com/store/apps/details?id=com.shubansoftware.prodcat&hl=en&gl=US, 10.05.2021.

Artificial Intelligence Application in Biosensor

Ayşe İpek GÖZÜTOK

Department of Biomedical Engineering, Faculty of Engineering and Natural Sciences, Biomedical Engineering Master Degree in Bahcesehir University, 34353

ayseipek.gozutok@bahcesehir.edu.tr

Abstract

Studies in the field of health have also changed with the developing technology. These changes have also affected the biosensors used in patient treatment and diagnosis. Biosensors are now common in biomedical diagnosis and a wide variety of other applications, such as point-of-care monitoring of treatment and disease development, environmental monitoring, food safety, drug discovery, forensics and biomedical science. Biosensors have been developed by paying attention to the expectations of authorized healthcare professionals in diagnosis and treatment and to be innovative for the patient. For this, the use of artificial intelligence has given the most accurate result of using biosensors. A perfect combination of increased computer processing speed, larger data collection data libraries, and a large AI talent pool has enabled rapid development of AI tools and technology, also within healthcare. Biosensors created using artificial intelligence have given the user the opportunity to use various fields. The aim of this study is to show readers both in which field biosensors can be used in artificial intelligence applications and the results obtained within these uses. In this context, we recommend the more widespread and active use of biosensors obtained by artificial intelligence applications to facilitate the lives of users and to obtain fast and accurate data in diagnosis and treatments.

Keywords: Artificial Intelligence (AI), Biosensors, Nanomaterials, Internet of Things (IoT), Cardiovascular Disease (CVD).

1. Introduction

Biosensor is the word that the biological sensor is abbreviated and written together. Definition of electronically operating biosensors. They are analytical devices that show selective properties against biological samples in the environment and convert the structure and density information of the samples into measurable and processable electrical signals. A biosensor; It is used in the detection of biological and chemical substances by converting the chemical, biological or biochemical signal into an electrical signal that can be measured and processed with the converter in its structure. Biosensors generally cover a wide range from the clinic to the agricultural field and the food industry. Basic properties of biosensors; stability, cost, precision and repeatability [1][2]. From the traditional point of view, depending on the type of transducer units, biosensors can be classified as bioelectrode sensors, semiconductor biosensors, thermal biosensors, photo biosensors, and piezoelectric crystal biosensors.

To give examples of some biosensor types and features; Bio-electrode sensors (BES) are an emerging technology for water quality monitoring. It enables tracking of the user's metabolic activity in a microbial culture. An anode / cathode pair with a resistor between them is the basic architecture of a BES device. The organic material is oxidized in the waste water by exo-electrogenic microbes on the anode and electrons are transferred through the circuit to the cathode. The protons (H+) formed at the anode migrate to the cathode, forming a full electrical circuit, to recombine with electrons passing through the circuit. Electrons are measured and registered traveling along the resistance. Due to their flexibility in chemical and electrochemical behavior, organic semiconductors can be used in the field of bioelectrochemistry.

Using organic synthesis for use in cell stimulating, renewable energy production and biosensors, these materials have been modified. Ultimately, biosensors work on the basis of a physical, optical or electrochemical alteration in the active material as the desired analyte is encountered. In a redox process, the analyte either interacts directly with the semiconductor material or undergoes a redox process with a component such as an enzyme added to the material of the semiconductor. The electrochemical signal is then expressed through the material of the semiconductor [3].

In a biochemical reaction, thermal biosensors calculate the heat energy emitted or consumed. We can detect organic pollutants in waste water by using COD. COD is abbreviated as Chemical Oxygen Demand. It is known as the amount of oxygen equivalents absorbed by strong oxidation agents during the organic oxidation of pollutants. During this oxidation, this thermal biosensor tests the heat released. Flow injection analysis is used [4].

Photoelectrochemical (PEC) under light illumination and applied potential is used to transform chemical energy into electricity. Due to its capacity to detect biomolecules by photocurrent derived from biomolecule oxidation, PEC biosensing [5] has attracted great interest.

Piezoelectric biosensors are a group of analytical instruments that function on the theory of capturing affinity interactions. Due to a mass bound on the piezoelectric crystal floor, a piezoelectric platform or piezoelectric crystal is a sensor component operating on the theory of shift in oscillations. Piezoelectricity is a physical phenomenon which refers to a material's ability to generate voltage when mechanically stressed. When the voltage alternates, the voltage administered to the surface of a piezoelectric material induces mechanical tension or oscillation [6]. Depending on the type of recognition elements, biosensors can be divided into enzyme sensors, nucleic acid sensors, microbial sensors, cell sensors, tissue sensors and immunosensors. Depending on the type of recognition elements, biosensors can be classified as biosensors, metabotropic biosensors and catalytic biosensors [7-8].

Artificial intelligence (AI) is a branch of computer science that studies the properties of intelligence by synthesizing intelligence. AI refers to systems or machines that imitate human intelligence to perform tasks and can recursively improve themselves based on the information they gather. Artificial Intelligence is about super-powered thinking and data analysis capability and process rather than any particular form or function. Artificial intelligence can be used in various fields, such as; health analysis and patient care, voice assistants, communication, information management. As one of the most important categories among them, biosensors have experienced a long development from classic electrochemical biosensors to wearable and implantable biosensors, and have been widely applied in food security, healthcare, disease diagnosis, environmental monitoring and biosafety [9-12].

The basic architecture of AI-biosensors is composed of three main elements: information collection, signal conversion and AI data processing. Information gathering refers to a group of biosensors for continuous monitoring of physical, chemical, biological, environmental or identity information. The signal conversion system converts the information from the data collection area into an electrical output signal with a defined sensitivity. Artificial intelligence data processing; can be grouped into interface, data classification, data model and analysis, and decision layer.

2. AI Biosensor Applications

Recently, biosensing has started a new phase due to the incentive and application of concepts such as big health, Internet of Things (IoT) and big data. Developing wearable biosensors aim to overcome the limitations of central, delicate healthcare by giving individuals an idea of their physical dynamics. Integration of artificial intelligence (AI) approaches with biosensors, including pattern analysis and classification algorithms, can bridge the gap between data collection and analysis and provide improved diagnostic and treatment accuracy [13].

Biosensors also make use of nanomaterials in wearable biosensors and other sensors. Nanomaterials have important properties in biosensors, by increasing the sensitivity of the biosensor, it can lead to changes in optical, thermal, magnetic and mechanical properties, which are particularly suitable for the study of biological processes in living cells and the development of major diseases [14]. At the molecular or subatomic level, nanotechnology is the study of materials. This entails particle processing smaller than 100 nanometers (one nanometer is one billionth of a meter) and the technology requires the creation of materials or structures that are invisible to the human eye and sometimes hundreds of times thinner than the width of human hair. When reduced to the nanoscale, the mechanics and chemistry of materials are drastically different; they have multiple powers, conductivity and reactivity, and leveraging this could revolutionize medicine.

A big problem in modern medicine, for instance, is that the body can not digest the full dosage of the medication offered to a patient. Scientists can ensure that medications are administered to particular locations in the body with greater specificity using nanotechnology, and the drugs can be designed such that the active ingredient permeates cell membranes faster, minimizing the dosage needed [15].

Nanocoatings, which improve biocompatibility and thus facilitate integration with the surrounding tissues of a variety of medical implants, are one of the most significant forms of nanotechnological applications. Examples are applicable in cardiology (stent coating), orthopedics (joint replacement implant coating) and dentistry (dental implants). In addition, nanomaterials' antimicrobial properties are used in coatings, as well as in wound care and surgical textiles. A multitude of uses for nanotechnology are unique to oncology. Examples include screening methods used in the early diagnosis of cancer and instruments used to recognize tumor or metastasis borders during surgical procedures. The effect of therapies such as chemotherapy or radiation therapy can either be improved by nanomaterials by locally increased temperatures, or they can destroy tumor cells directly at high temperatures [16].

Wearable biosensors are biosensors that are combined with artificial intelligence and made available to everyone in daily life. It provides various biofeedbacks to the user while doing sports, resting, eating or even sleeping on smartphones, smart watches, in other words, many devices that people use in their daily lives. Due to the growing consumer desire for health and fitness awareness, and many other factors such as the rapid development of material technology, the prevalent use of smartphones and the popularity of internet technology, wearable biosensors have attracted great attention. At the same time, portable biosensors make it possible to monitor continuously and are easy to operate and transport.

By combining artificial intelligence with various IoTs, we can benefit from the following applications; measure and monitors pulse (HR), motions sensors (accelerometer and gyroscope), the health app measures body mass index (BMI), basal metabolic rate (BMR), measures the level of stress and the peripheral oxygen saturation (SpO2), measuring and monitoring daily activity levels (counting steps, heart beats, calories consumed daily, water consumption) etc [17]. We can give an example from some research using the various applications of wearable biosensors described above.

Researchers discussed the effect of artificial intelligence on higher education teaching and learning and the benefits of new technology. They also introduced the concept and implementation of a fully mobile method of biofeedback, directed by a trained mentor, using a smartwatch for teaching personalized people while living in natural behaviors. These experiments have been used to observe the impact of wearable biosensors on consumers. Research has shown that consumers can also use the intelligent watch that can be used for Education 4.0 to acquire some info. The results of the tests are that sensor knowledge from wearable devices such as accelerometer, gyroscope, motion, pedometer, heart rate calories, NFC and Bluetooth can be used for Education 4.0 [18]. The “IoT for Education 4.0” system can evaluate students’ bio signals in real time to monitor their behavior and progress, without distracting them during learning activities. According to the results of the research using smart phones and smart watches, wearable devices can be evaluated as “easy to use”. In this context, increasing the use of wearable devices both for students with a single research group and for many research groups allows the necessary information to be quickly uploaded to the system and examined by professionals. These data can also increase the quality of life of research groups and users in the future [19].

To give another example of usage areas of wearable sensors; Google patented the use of always-on optical sensors in locations such as bathrooms in November 2016 to collect cardiovascular function data. To see if blood flow has been dispersed, the sensors detect variations in skin colors. These sensors can passively provide blood flow data any time a person walks into the bathroom to brush their teeth or wash their hands. Embedded in everyday objects, optical sensors could provide access to cardiovascular activity and help users improve health. We see that in 2019, smart home-based healthcare technology will enter the market [20].

The joint use of artificial intelligence and biosensors is also used in cardiac health care area. Cardiovascular diseases(CVDs) and stroke are at the top causing death globally. Early and quick diagnosis is crucial for succesful prognosis of CVD and stroke. Ideal biosensor can play a vital role in the timely and accurate diagnosis of CVD, to spare numerous lives, particularly for the patients enduring the herat attack. The point of care (POC), which is carried out close to the patient’s ambience for diagnostic purposes, can be summarized as a fast, cheap and effective process. The integration of biosensors with Bluetooth, Wi-Fi and GPS wireless features as shown in Fig. 1 enables communication between the professional health expert and patient care at home [21].

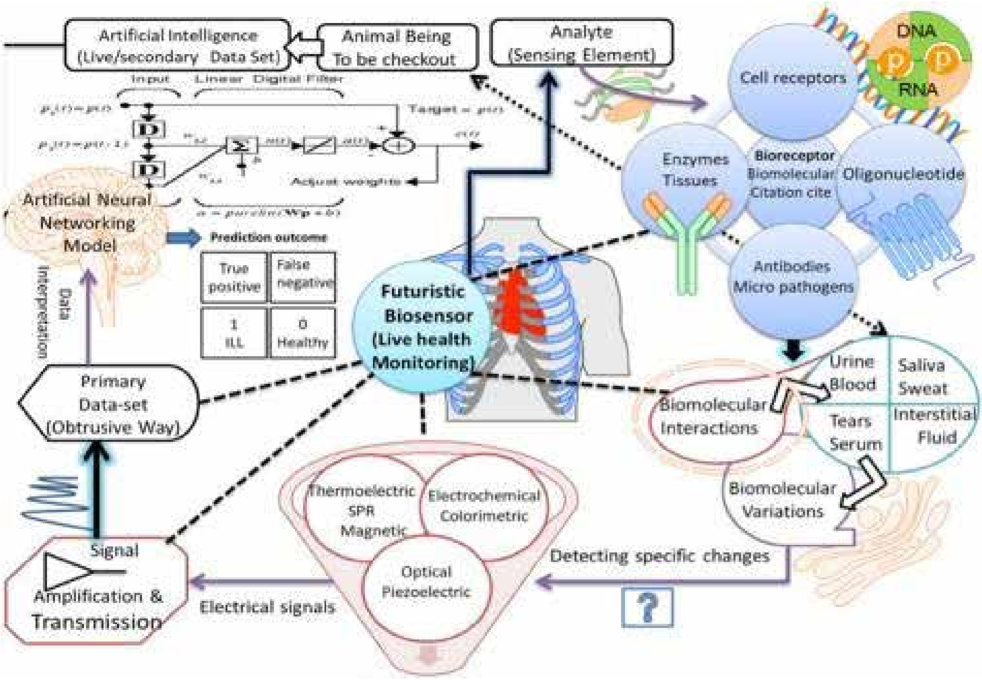


Fig. 1. Artificial Intelligence with Futuristic Biosensor [21]

Biosensors can be used in the diagnosis and treatment of many diseases. It can also be used for MS disease, which is one of them. Multiple sclerosis (MS) is a disease that may affect the brain and spinal cord, causing a wide array of likely signs, including vision, movement of the arm or leg, sensation, or coordination issues. In MS, the accelerometer is the most widely used wearable tracker in researchers have focused, although many other sensor types, including gyroscopes and electrodermal sensors, have also been used. Several nanobiosensors, including those that track MS-related micro-RNAs or protein biomarkers, are used in cerebrospinal fluid. In MS clinical study and practice, the considerations that will decide the implementation of biosensors include reliability, vulnerability to changes in operation, degree of risk to patient life or clinical workflow, cost and privacy ethics, and patient safety. Practical concerns include setting data quality and collection criteria for these instruments, as well as deciding the best sampling rates for multiple data types. Physicians would be most helpful in distributing the evaluated results in real-time. For data sharing, Bluetooth or other wireless communications are very important. With this data transfer method, information can be sent to authorized persons more quickly and reliably. Since multi-sensor systems can provide complementary data for a richer representation of a particular motion or event, they will be useful in gaining wider information in the field of biosensing [22].

Scientists investigated the use of biosensors to detect the COVID-19 virus, which emerged at the end of 2019 and is still ongoing today. The biosensor collects the physical activities of the patient such as heart rhythm, body temperature, oxygen level, respiratory rate during sleep, eating and walking in partnership with the patient and mobile device, collects them in a cloud program and sends them to the relevant health institutions to ensure a rapid and supervised diagnosis and treatment period [23].

3. Conclusion

The aim of this study is to show how the increase of biosensors used with artificial intelligence applications in the field of health will provide convenience for both the authorized healthcare worker and the patient in diagnosis and treatment. The recommendation of the use of biosensors, which is planned for the recently observed COVID-19 disease, to ensure that patients' symptoms are observed at home, can show how effective it can be for health technology to expand the use for other diseases and treatments. Especially recording the information through the results obtained using artificial intelligence and transferring it to the authorities quickly makes it easier. Could this data recording already allow for scientific developments for future studies? As technology and scientific studies continue, we can get answers to these and many other questions thanks to the method transferred in the study you are reading. As a result of this article, the readers have shown the time and knowledge gain of artificial intelligence through biosensors used in the field of health.

Acknowledgements

This research was supported by Bahcesehir University. We thank our colleagues from Bahcesehir University who provided insight and expertise that greatly assisted the research, especially thank you to Prof. Dr. Can Balkaya.

References

[1] Bhalla, N., Jolly, P., Formisano, N., & Estrela, P. (2016). *Introduction to biosensors. Essays in biochemistry, 60(1), 1–8.* doi:10.1042/ebc20150001

[2] Verma, N., Bhardwaj, A., (2015). *Biosensor technology for pesticides—A review,* Springer Science+Business Media New York 20150901

[3] Hopkins, J., Fidanovski, K., Lauto, A., & Mawad, D. (2019). *All-Organic Semiconductors for Electrochemical Biosensors: An Overview of Recent Progress in Material Design. Frontiers in Bioengineering and Biotechnology, 7.* doi:10.3389/fbioe.2019.00237

[4] S. Vasuki, V. Varsha, Mithra, R. Arul Dharshni, S. Abinaya, R. Deva Dharshini, N. Sivarajasekar, March 2019, *Thermal Biosensors and their applications* Department Of Biotechnology, Kumaraguru College of Technology, Coimbatore-641049, India.

[5] Devados, A., Sudhagar, P., Terashima, C., Nakata, K., June 2015, Journal of Photochemistry and Photobiology C Photochemistry Reviews, *Photoelectrochemical biosensors: New insights into promising photoelectrodes and signal amplification strategies*

[6] Pohanka, M. (2018). Overview of Piezoelectric Biosensors, Immunosensors and DNA Sensors and Their Applications. Materials, 11(3), 448. doi:10.3390/ma11030448

[7] Mehrotra, P. (2016). *Biosensors and their applications – A review. Journal of Oral Biology and Craniofacial Research, 6(2), 153–159.* doi:10.1016/j.jobcr.2015.12.002

[8] Kozitsina, A., Svalova, T., Malysheva, N., Okhokhonin, A., Vidrevich, M., & Brainina, K. (2018). *Sensors Based on Bio and Biomimetic Receptors in Medical Diagnostic, Environment, and Food Analysis. Biosensors, 8(2), 35.* doi:10.3390/bios8020035

[9] David D. Luxton, in Artificial Intelligence in Behavioral and Mental Health Care, 2016, An Introduction to Artificial Intelligence in Behavioral and Mental Health Care

[10] Collins, G. S., & Moons, K. G. M. (2019). *Reporting of artificial intelligence prediction models. The Lancet, 393(10181), 1577–1579.* doi:10.1016/s0140-6736(19)30037-6

[11] Haenlein, M., & Kaplan, A. (2019). *A Brief History of Artificial Intelligence: On the Past, Present, and Future of Artificial Intelligence. California Management Review, 000812561986492.* doi:10.1177/0008125619864925

[12] Panch, T., Pearson-Stuttard, J., Greaves, F., & Atun, R. (2019). *Artificial intelligence: opportunities and risks for public health. The Lancet Digital Health, 1(1), e13–e14.* doi:10.1016/s2589-7500(19)30002-0

[13] Pasche, S., Angeloni, S., Ischer, R., Liley, M., Luprano, J., Vorin, G., September 2008, Advences in Science and Technology 57:80-87, Wearable Biosensors for Monitoring Wound Healing

[14] Pandey, P., Datta, M., Malhotra, B., February 2008, Analytical Letters 41(2):159-209, Prospects of Nanomaterials in Biosensors

[15] Tothill IE. Biosensors and Nanomaterials and their Application for Mycotoxin Determination. World Mycotoxin Journal. 2011; 4(4):361-374.

[16] Vashist, S.K., Venkatesh, A.G., Mitsakakis, K. *et al.* Nanotechnology-Based Biosensors and Diagnostics: Technology Push versus Industrial/Healthcare Requirements. *BioNanoSci.* 2**,** 115–126 (2012). https://doi.org/10.1007/s12668-012-0047-4

[17] M. I. Ciolacu, L. Binder and H. Popp, "Enabling IoT in Education 4.0 with BioSensors from Wearables and Artificial Intelligence," *2019 IEEE 25th International Symposium for Design and Technology in Electronic Packaging (SIITME)*, Cluj-Napoca, Romania, 2019, pp. 17-24, doi: 10.1109/SIITME47687.2019.8990763.

[18] M. I. Ciolacu, L. Binder, P. Svasta, I. Tache and D. Stoichescu, "Education 4.0 – Jump to Innovation with IoT in Higher Education," *2019 IEEE 25th International Symposium for Design and Technology in Electronic Packaging (SIITME)*, Cluj-Napoca, Romania, 2019, pp. 135-141, doi: 10.1109/SIITME47687.2019.8990825.

[19] Popenici, S.A.D., Kerr, S. Exploring the impact of artificial intelligence on teaching and learning in higher education. *RPTEL* **12,** 22 (2017). https://doi.org/10.1186/s41039-017-0062-8

[20] Ciolacu, M. I., Binder, L., & Popp, H. (2019). *Enabling IoT in Education 4.0 with BioSensors from Wearables and Artificial Intelligence. 2019 IEEE 25th International Symposium for Design and Technology in Electronic Packaging (SIITME)*. doi:10.1109/siitme47687.2019.8990763

[21] Vashistha, R., Dangi, A. K., Kumar, A., Chhabra, D., & Shukla, P. (2018). Futuristic biosensors for cardiac health care: an artificial intelligence approach. 3 Biotech, 8(8). doi:10.1007/s13205-018-1368-y

[22] Graves, J. S., & Montalban, X. (2020). Biosensors to monitor MS activity. Multiple Sclerosis Journal, 135245851988817. doi:10.1177/1352458519888178

[23] Wong CK, Ho DTY, Tam AR*, et al*, Artificial intelligence mobile health platform for early detection of COVID-19 in quarantine subjects using a wearable biosensor: protocol for a randomised controlled trial, *BMJ Open* 2020;10:e038555. doi: 10.1136/bmjopen-2020-038555

Artificial Intelligence for Proactive Internal Audit in Business

Gülnihal Çiçekdağı[1] and Gökçe Sargut[2]

[1]Big Data Analytics and Management, Faculty of Engineering and Natural Sciences, Bahcesehir University, Graduate Student, Yıldız, Çırağan Cd., 34349 Beşiktaş/İstanbul, gulnihalcicekdagi@gmail.com

[2]Associate Professor of Management, Governors State University, University Park, IL, USA, gsargut@govst.edu

Abstract
Internal audit functions of companies are important in helping companies establish favorable public opinion. Auditors examine almost every aspect of companies and mentor top management teams in accordance with their findings. As the world is changing every day and technological developments affect every business area, the way that the companies perform their tasks and how they serve their customers is also changing. Customers expect faster and more capable products and services day by day. Companies are increasingly pressured to adapt newly-introduced technologies and practices that enhance customer experience. As the speed of both adoption and adaptation are increasingly becoming critical for survival, the internal audit functions of companies should also keep up with the dizzying pace of change to ensure the effective monitoring of the gamut of organizational operations. During the past decade, the terms 'data science', 'machine learning', and 'artificial intelligence (AI)' have become ubiquitous. Various technologies and practices related to these concepts have already begun making their mark in industries as varied as finance and health. Companies are seeking ways to integrate data science and artificial intelligence into their processes to benefit from advantages in speed, quality, efficiency, and effectiveness. Thus, internal audit functions should also consider how AI can change their company, as well as the risks involved in AI, and the most productive uses for AI in internal audit activities. In order to be successful in this important task, an internal audit function should have a good understanding of AI. Moreover, it should explore the correct ways of integrating AI to their activities so as to keep pace with this defining technological era.

Keywords: Internal audit, audit, artificial intelligence, machine learning, data science.

1. Introduction

The concept of 'internal audit' came into view with the establishment of The Institute of Internal Auditors (IIA) in United States in 1941 [1]. It became a vital element for most of the companies from different sectors all over the world. It is a key function for organizations to gain a positive public opinion, because it provides an assurance that the organization has a sufficient system and its operations are accurate, efficient, timely and compliant. To do this, internal audit departments of the companies examine the whole system including risk management, governance and internal controls [2]. In big companies, internal audit function is conducted by separate departments, while in small ones it is conducted by internal auditors. No matter the size of the company, internal auditors should have well-organized audit plans because companies have complex structures. A wide variety of processes and operations requires that huge amounts of information flow through different departments. Auditing is an absolute necessity given such complexity in order to safeguard the organization from potential improper and/or fraudulent activities, which may cause material damage and reputation loss. As such, an internal audit function can only be effective to the extent that auditors are able to stay up-to-date with new applications, visions, and technologies that the companies use. It is within this context that the importance of 'Artificial Intelligence (AI)' needs to be emphasized.

Computers have increasingly become an essential part of our lives since their widespread adoption began in the 1980s. But the immense processing power and the sheer volume of information accumulation it promised took an even more dramatic turn during the last decade, when a new generation of 'Deep Learning' models paved the way for rapid advances in Artificial Intelligence. Today, not only have the terms 'Artificial Intelligence', 'Big Data', and 'Deep Learning' become commonplace, their impact on the ways in which companies operate and serve their customers was almost immediate. Such a rapid change in business practices underscores the importance and challenges of maintaining an informed and capable internal audit function.

In this research paper, the problems encountered by internal auditors, and the usefulness of AI for solving these problems are examined. Why and how internal auditors need to stay up-to-date with the pace of change are also discussed. Finally, some real-life examples of AI usage in internal audit are given for demonstrative purposes.

2. How Artificial Intelligence Can Improve Internal Audit

2.1. Classical Approaches and Problems

In classical internal audit approaches, an annual audit plan is prepared at the beginning of a year, and this plan is followed by the audit department—with the exception of unexpected situations. Deciding which areas of organizational activity and related processes are more critical and possess high risk is very crucial in the audit plan phase. High risk areas should be audited more often and/or with more scrutiny in order to constitute a risk-based audit system. Therefore, in the planning phase, risk assessments are made for all existing processes, to classify them in terms of risk levels. The classical methodologies in risk assessment are mostly based on past experiences and historical data. This is widely regarded as a time-honoured, legitimate approach as auditors have traditionally been trained to familiarize themselves with the inner workings of their company and having a comprehensive top-to-bottom view of the whole organization. But as with all models, such an approach has its own blind spots. In the case of auditing, this often means not being able to identify newly-emerging risks generated as a by-product of the ongoing changes in a company's practices. For example, in a company which has established new processes and operations, the internal audit function probably does not have any experience with these new processes, so an experience-based approach can be misleading.

On the other hand, elements that are subject to internal audit are include huge amounts of transactions, operations, documents, employees, etc. In the classical approaches, samples are taken and examined one by one to detect faulty and fraudulent operations or activities. When such a faulty or fraudulent operation is detected, it is reported to the supervisors of responsible employees so that actions can be taken as necessary. Sampling is a good method, especially if samples are created with effective methods like stratified sampling or cluster sampling. However, examining every single element in all samples is very time and resource consuming for both the internal audit department and the company.

One other issue with classical approaches is that a problem is usually detected long after it occurs, due to its retroactive nature. As operations and transactions are carried out in any given organization, a wide variety of forms and all manner of documents are prepared. As time goes on, all these are gathered and samples are prepared to be audited. Problems have already occurred by the time they are detected, at which point corrections are made, if possible. When corrections are not possible, it can often lead to financial or reputational damages for the company.

As an example of the three problems mentioned above, consider the hypothetical case of a national bank: XBank is a very large organization that has millions of customers, thousands of employees, hundreds of branches and tens of departments. All the units and departments in the bank have varying types of processes, operations, methodologies, documents, etc., all of which are subject to internal audit. For example, a branch is the place where XBank encounters its customers by servicing them with account openings, withdrawals, money transfers, bill payments, etc. On the other hand, the credit department decides which customers qualify for credit, while the accounting department is responsible for recording these credits in balance sheets. All these departments have separate but interrelated responsibilities. The Internal Audit Department of XBank should consider all such complexities during the planning phase. Based on their experience, the department probably has knowledge about which departments or branches pose more risk for the company. In addition, auditors also can establish support for their risk assessments using statistical methods based on past records. However, although they are reasonably capable in accounting for past activity using the established methods, auditors can easily miss newly emerging problems. After planning, internal auditors visit branches and departments to examine their activities within the scope of the samples that were collected. Trying to find problematic issues in samples is very time consuming for auditors because every item in the sample needs to be analysed, even if they are not faulty. Although an internal auditor at a branch can detect faulty transactions, corrections may not be possible after-the-fact. For example, a money transfer made to the wrong account number cannot be reversed. In such situations, the bank would have to reimburse the customer’s loss.

2.2. How to integrate AI applications to Internal Audit

The term Artificial intelligence (AI) refers to the ability of computers to perform like human beings [3]. After the birth of the field in 1956, at a conference at Dartmouth College [4], AI became more and more popular across the world in almost every sector. The reason why it has gained such popularity may just be due to human curiosity or the hope that making computers do everything for us, will make the world a better place to live in. Regardless of the motivation, computers have become such an essential part of our lives that nowadays we can see the terms ‘Artificial Intelligence’, ‘Big Data’, ‘Deep Learning’ everywhere. This also affects the ways that companies operate and serve their customers. As the offerings, activities, and processes of companies continue to change day by day, the internal audit function also has to change to remain effective.

According to a report from 2018 by the global consulting firm Protiviti, data analytics usage in internal audit departments was 76% for the Asia-Pacific region and Europe, and 63% for North America [5]. This shows that a big percentage of internal audit departments had already integrated data analytics into their activities a couple of years ago, which suggests that the numbers are probably higher today. As the very first step of AI is data analytics, we can say that utilizing AI is not rocket science, it is achievable and even there are some real-life examples of it.

We have hinted that the above-mentioned problems that are frequently encountered by internal audit departments, can be overcome by artificial intelligence, but how? From planning phase to reporting phase, all processes of internal audit can be integrated with data science and artificial intelligence to enhance them in terms of speed, quality, and efficiency.

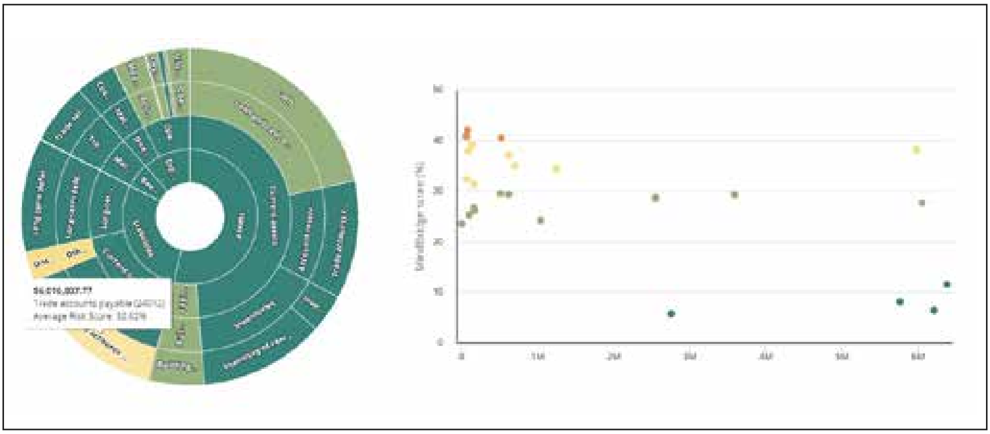


Fig. 1. Example visualizations from MindBridge's 'AI Auditor' [6].

Consider the risk assessment and planning phase, which constitutes the basis for internal audit activities. Past experiences are vital because they give internal auditors spot-on ideas about risky areas. However, the ones which have not been detected yet can be overlooked. Fortunately, thanks to the technological advancements in data science, operations, people, processes etc. that behave similarly can be distinguished from the less important data that exist in huge data universes. Therefore, if data science is incorporated into risk assessments, the problem of failing to notice some risky, fraudulent and faulty elements can be addressed more effectively. Moreover, risk assessments and plans can be prepared by using artificial intelligence. It is impossible for auditors to simultaneously analyse everything and catch red flags. However, artificial intelligence can be used to analyse most processes in a company to detect anomalies and prepare audit plans based these. AI can be a very valuable tool in assisting internal auditors detect fraudulent activities or faulty applications. For example, the software platform 'AI Auditor' by MindBridge assists internal audit departments of companies by processing all transactions to look for anomalies and red flags—as well as assessing the risks related to these. The platform also creates reports and graphics that allow internal auditors to get new insights about the company's activities. An example of such a visualization can be seen in Fig.1 [6].

Risk assessment and planning is the first step of internal audit. After planning, audit activities are conducted in every area included in the plans. At this step, recordings, documents, transactions, processes, etc. are analysed in a deeper sense to detect discrepancies. In classical approaches, internal auditors gather everything together and create samples to examine every item individually. Generally, there are huge amounts of data even in samples, resulting in time-consuming internal audits. Despite the large amount of resources that are often allocated to internal audit tasks, most audits detect only, a small number of errors and very rare fraudulent activities. Therefore, there is a need for questioning the efficiency of internal audit processes. Artificial intelligence can also be useful for examining documents, recordings, operations, and processes in detailed way. For example, 'AuditMap.ai' is an AI-based solution that can read thousands of documents simultaneously to notice key ideas and cluster them for audit purposes [7]. Such AI-based methodologies can not only greatly reduce the time, effort, and resources allocated by the company, but also increase the efficiency of internal audits.

Internal audit has a retroactive nature. In other words, organizational processes often have shortcomings that inadvertently allow mistakes and fraudulent activities. Internal auditors can identify such activities only after they have occurred, which makes the internal audit process retroactive. Although this traditional approach has proven to be useful in finding and correcting discrepancies, internal auditors have been increasingly becoming aware of the fact that instead of detecting a fault, preventing its occurrence is a much more favourable alternative for companies.

Thus, if internal auditors were able to successfully predict the areas that are open to mistakes, they can also prevent them before happening. Artificial intelligence again becomes a key factor in such a paradigm shift. AI should be adopted for internal audit activities because it can be used to investigate past recordings of processes and find patterns of discrepancies, in turn leading to detection of processes which will most likely result in faults. At that point, internal auditors can work on these error-prone processes and explore the required internal controls that need to be adapted to these processes. As such, the auditing department will have data-driven, evidence-based, and actionable suggestions that they can present to the top management. In this way, not just the faults, but also the causes are detected, resulting in a *proactive* auditing process that significantly enhances the performance of the internal audit department.

Additionally, while internal audit departments are seeking ways to integrate AI applications into their activities, the companies themselves are also integrating AI to their every-day operations and processes. Therefore, internal auditors should know in the near future they have to become competent on auditing AI-based processes in their companies. AI is generated by humans, and like any other human-built system, it is prone to biases and mistakes. Therefore, placing blind faith in AI systems can result in big mistakes. This is another reason why employing well-trained auditors—who can also keep their knowledge current—is crucial for all organizations.

According to a survey conducted in December 2017 by 'Tone at the Top', a bimonthly newsletter published by Institute of Internal Auditors (IIA), 65% of the participants responded that internal audit departments of their companies did not integrate AI into their activities as shown in Fig.2 [8]. More strikingly, just 5% of respondents believed that their internal audit departments were either extensively or fully capable of artificial intelligence usage in their activities--which shows that AI did not have a big impact on internal auditing as recently as a couple of years ago. If this survey was conducted today, the results would be slightly better—but a much higher rate of adoption would be necessary for internal auditing practices to remain useful and relevant in today's organizations.

Fig.2. Survey Results from Tone at the Top, December 2017 [8]

Although the survey results are not promising, managers of successful companies are aware of the need for change. To ensure the survival of their companies, they must do what is required. As the internal audit functions are the defence mechanism of the companies, they are the key to survival--but they have to be alert in today's rapidly changing environments. If internal auditors fail to maintain a capable internal control and risk management system, then the company becomes open to the failures. Therefore, they should develop systems to continuously learn and retain appropriate levels of data science, machine learning and artificial intelligence knowledge to serve their companies.

3. Conclusion

The world is rapidly changing in terms of technological improvements, and companies that fail to catch up with this changing era will disappear one by one. Internal audit is a critical function of a company especially in such changing environments, due to its capability in alerting top management in areas that pose a potential for failure. Thus, internal audit departments should have a sufficient knowledge of technological advancements in general and Artificial Intelligence in particular.

AI-integration can also be beneficial for the work of internal auditors. For example, manual processes, inefficient auditing practices, and retroactive structure can be eliminated by artificial intelligence. It goes without saying that keeping up with a dizzying array of technological improvements is easier-said-than-done. However, it is very clear that the companies, as well as their internal audit departments are aware of potential avenues of failure, so they look for ways to adapt to these changes. Moreover, there are real-life applications already in place to serve internal audit functions of the companies. It is worth repeating that to be able fully benefit from these applications, internal auditors should have sufficient knowledge. Therefore, internal audit departments should take necessary steps to improve themselves to better serve their companies.

In conclusion, 'change is good' in a business area to the extent that companies are able find ways to adapt their existing structures and systems. Otherwise, they run the risk of failure and eventual extinction. Today, AI is one of the most popular technologies that the companies are keen on. As internal audit functions have key roles in the companies, they also have to catch up with AI.

References

[1] S. Ramamoorti, Internal auditing: history, evolution, and prospects, 1st edition. The Institute of Internal Auditors, 2003, pp.3-4.

[2] The Institute of Internal Auditors, Internal Auditing: Assurance, Insight, and Objectivity, https://na.theiia.org/about-ia/publicdocuments/pr-value_prop_bro-fnl-lo.pdf. Accessed 11 December 2020.

[3] B. Copeland, "Artificial intelligence". Encyclopaedia Britannica, 11 Aug. 2020, https://www.britannica.com/ technology/artificial-intelligence. Accessed 22 January 2021.

[4] T. Lewis, A Brief Story of Artificial Intelligence, Live Science, 04 December 2014, https://www.livescience.com/49007-history-of-artificial-intelligence.html#:~:text=The%20beginnings%20of%20modern%20AI,%22artificial%20intelligence% 22%20was%20coined. Accessed 22 January 2021.

[5] Consultancy.uk, Data analytics to become a game changer for internal audit, 25 April 2018, https://www.consultancy.uk/news/16863/data-analytics-to-become-a-game-changer-for-internal-audit. Accessed 25 December 2020.

[6] G. Song, How AI and data can power an effective audit plan, https://www.mindbridge.ai/blog/ai-data-power-audit-plan/. Accessed 18 January 2021.

[7] D. Shapiro, Artificial Intelligence for Internal Audit and Risk Management, Towards Data Science, 26 June 2020, https://towardsdatascience.com/artificial-intelligence-for-internal-audit-and-risk-management-94e509129d49. Accessed 05 December 2020.

[8] The Institute of Internal Auditors, Tone at the Top, 86: 4, February 2018, https://global.theiia.org/knowledge/Public%20 Documents / Tone-at-the-Top-February-2018.pdf. Accessed 28 January 2021

A Degradation Model to Predict the Remaining Useful Life of Bearings

Erkan Caner Ozkat[1]

[1]Department of Mechanical Engineering, Faculty of Engineering & Architecture,
Recep Tayyip Erdoğan University, Rize, Turkey,
(ORCID: 0000-0003-0530-5439),
erkancaner.ozkat@erdogan.edu.tr

Abstract

Bearings are widely used in the machinery industry. In particular, they undertake important functions at many critical points. Any failure in these parts causes the machine to stop and leads to unexpected production losses. These unplanned downtimes increase the costs. Therefore, the goal of predictive maintenance is to estimate the remaining useful life of critical machine parts, such as bearings. The main purpose of this study is to determine the machine health indicator from vibration data using machine learning methodology and then use the computed health indicator in the remaining useful life estimation of bearings. The steps required to compute the machine health indicator are listed as feature extraction, feature postprocessing, dimension reduction, and feature fusion. The vibration signal is collected in the time-domain and is transformed to the frequency-domain using the Welch method. The traditional statistical indicators such as mean value, standard deviation, skewness, kurtosis, etc. are utilized to extract features in both domains. Monotonicity analysis is employed to select the most useful features, and they are fused into a single health indicator by the principal component analysis. An exponential degradation model for estimating remaining useful life is developed using the forecasting machine learning method. The proposed methodology is evaluated through a case study involving vibration measurements generated using a ball bearing signal simulator.

Keywords: Predictive maintenance, vibration analysis, machine learning, health indicator, bearing.

1. Introduction

Current technological developments, increasing energy needs, changing and developing competitive conditions make maintenance activities, which constitute one of the main expenditure items of companies, even more important. Furthermore, production losses due to unexpected failures can hinder the implementation of production plans [1-3]. Today, the interruption of production even in minutes causes high costs.

In the event of sudden and unexpected failures, maintenance and repair are time-consuming and spare parts can be difficult to find, increasing costs. Although the maintenance activities after the failure is a kind of maintenance method, considering today's conditions, the understanding of intervening the machine before the failure is more priority and this issue is becoming increasingly important in the machine industry [4-6].

Predictive Maintenance uses predictive tools to determine when maintenance actions are necessary. It is based on continuous monitoring of a machine or process integrity, allowing maintenance to be performed only when it is needed. Moreover, it allows the early detection of failures thanks to predictive tools based on historical data (e.g. machine learning techniques), integrity factors (e.g. visual aspects, wear, coloration different from original, among others), statistical inference methods, and engineering approaches [7,8].

Due to the development of technology in recent years, data obtained from temperature, pressure, lubricant, noise, and vibration sensors are used in many predictive maintenance activities. The decision as to which sensor data to use depends on both the type of problem to be investigated and the operating conditions of the machine. Bearings are widely used in rotating machines. The smooth operation of the machines is closely related to the healthy operation of the bearings. Damage to the bearings is one of the most important reasons that increase the level of vibration. The amount of vibration emitted by the bearings is used to determine their remaining useful life. In addition, it is possible to obtain information about the roughnesses that occur on the surface of the bearings by monitoring the vibrations. Bearings generally operate under harsh environmental conditions and are therefore subject to several types of faults. Advanced feature extraction algorithms need to be developed to effectively detect bearing failures and fully understand vibration behaviour. It is traditionally aimed to extract features from time-domain, frequency-domain, and time-frequency-domain. In the literature, wavelet-energy spectrum, energy spectrum entropy, generalized gaussian density, envelope analysis methods have been widely used to determine the state of bearings in rotating machines [9].

One of the goals of predictive maintenance is to estimate the remaining useful life (RUL) of the machine [10, 11]. The health of a machine deteriorates over time from a healthy state to a faulty state. The RUL is defined as the reaming time from the current health to failure. Depending on the operation, this time can be represented in terms of hours, days, cycles, etc. In general, there are three common models to estimate the RUL: (i) survival model [12, 13]; when data only from the time of failure is known, (ii) degradation model [14, 15]; when failure data is not available, but a safety threshold that indicates failure is known, and (iii) similarity model [16-18], when the complete history (from health state to failure) from similar machines is known.

The main purpose of this study is to determine RUL using a degradation model developed according to the computed health indicator. The time-domain features (root mean square, peak value, signal kurtosis, etc.) and the frequency-domain features (peak frequency, mean frequency, etc.) are extracted from raw vibration signals and are fused into a single health indicator. The model can detect the degradation trend and updates itself when a new observation becomes available.

The remainder of this paper is organized as follows: Section 2 reviews the data structure used in this paper. Section 3 presents each step in the proposed methodology. Section 4 presents the results. Section 5 provides conclusions of the work.

2. Bearing Degradation Dataset

The vibration signals used in this study were obtained from a ball bearing signal simulator with a single point defect on the outer race of the bearing. The ball bearing was designed to operate at high speed and vibration signals were presented to the researchers as an open-source [19]. This vibration signal presented in Fig. 1a contains multiple segments of vibration signals for bearings simulated in different health conditions and each segment stores signals collected for 1 second at a sampling rate of 20 kHz. An outer race defect with a depth ranging from 3 mm to over 3 mm has been developed and is given in Fig. 1b.

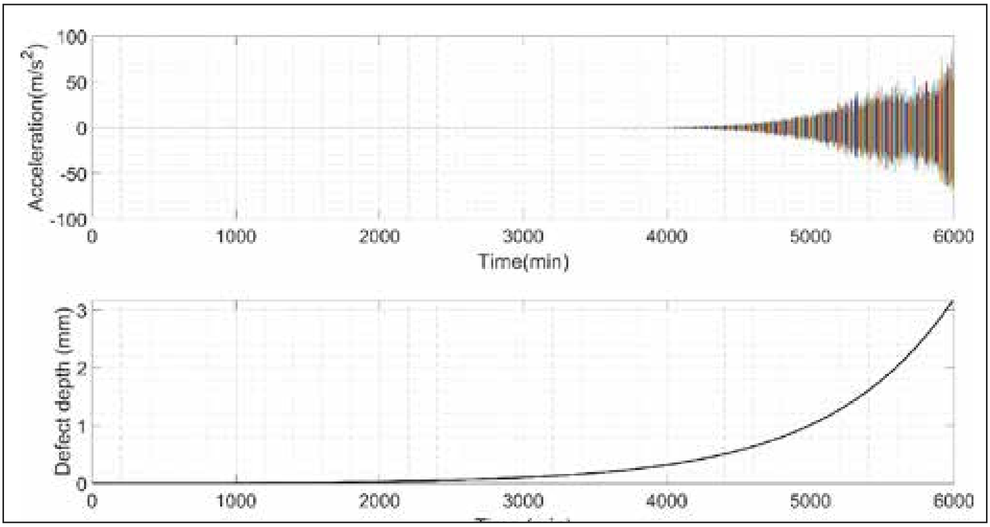


Fig. 1. a) Vibration Signal, b) Defect depth changes

3. Methodology

This study has two objectives: *(i)* to compute the health indicator, and *(ii)* to estimate the RUL. The necessary steps to compute the health indicator are listed as: *(i)* feature extraction, *(ii)* feature postprocessing, and *(iii)* dimension reduction and feature fusion. Having computed the health indictor, a degradation model will be trained to estimate the RUL using the computed health indictor. The next section deals with a detailed description of the proposed methodology.

3.1. Feature Extraction

Bearings often operate under harsh environmental conditions and are therefore subject to several types of faults. Therefore, a median filter is initially applied to the vibration signals collected in the time-domain to remove high-frequency noise components as well as to preserve useful information in the high frequencies. Then, they are transformed into different areas, primarily the frequency-domain, in order to extract error-sensitive features and obtain useful information. In this study, vibration signals are transformed from time-domain to frequency-domain using the Welch method. This method returns the power spectrum density (i.e. *pxx*) and cyclical frequency, (i.e. *f*) for a given vibration signal (i.e. *x(k)*), where x is the sampled vibration signal in time-domain and k is the sample index.

Statistical features are considered to be a simple but powerful tool, which characterize the change of a signal when faults occur. Table 1 presents a collection of statistical features and these are going to be extracted from the time-domain signal and the power spectrum density in frequency-domain for each sampled vibration signal. Furthermore, statistical features in time-domain are significantly divergent under the failure, and frequency-domain is mainly based on the study of the spectral analysis. Therefore, mean peak frequency is extracted from each segment of power spectrum density to distinguish healthy ball bearings from faulty ball bearings.

Table 1. A collection of statistical features

Statistical Feature	**Equation**	
Mean	$x_m = \frac{1}{N_k}\Sigma_{k=1}^{N_k} x(k)$	(1)
Standard Deviation	$x_{std} = \left(\frac{1}{N_k}\Sigma_{k=1}^{N_k}\left(x(k)-x_m\right)\right)^{\frac{1}{2}}$	(2)
Skewness	$x_{skewness} = \frac{\Sigma_{k=1}^{N_k}\left(x(k)-x_m\right)^3}{\left(N_k-1\right)x_{std}^3}$	(3)
Kurtosis	$x_{kurtosis} = \frac{\Sigma_{k=1}^{N_k}\left(x(k)-x_m\right)^4}{\left(N_k-1\right)x_{std}^4}$	(4)
Peak-to-Peak	$x_{p-p} = x_{\max} - x_{\min}$	(5)
Root Mean Square	$x_{RMS} = \left(\frac{1}{N_k}\Sigma_{k=1}^{N_k} x(k)^2\right)^{\frac{1}{2}}$	(6)
Crest Factor	$CF = \frac{x_{\max}}{RMS}$	(7)
Shape Factor	$SF = \frac{x_{RMS}}{\frac{1}{N_k}\Sigma_{k=1}^{N_k}\lvert x(k)\rvert}$	(8)
Impulse Factor	$IF = \frac{x_{\max}}{\frac{1}{N_k}\Sigma_{k=1}^{N_k}\lvert x(k)\rvert}$	(9)
Margin Factor	$MF = \frac{x_{\max}}{\left(\frac{1}{N_k}\Sigma_{k=1}^{N_k}\lvert x(k)\rvert\right)^2}$	(10)
Energy	$E = \Sigma_{k=1}^{N_k} x(k)^2$	(11)

3.2. Feature Postprocessing

It is required to evaluate how well features correlate with the degradation pattern for a reliable RUL estimation. For this purpose, monotonicity is employed in this used to select the most appropriate features for RUL estimation. Monotonicity characterizes the trend of a feature as the system evolves toward failure. As a system gets progressively closer to failure, a suitable condition indicator has a monotonic positive or negative trend, which is defined as follows:

$$monotonicity = mean\left(\left|\frac{positive\big(diff\left(x(k)\right)\big) - negative\big(diff\left(x(k)\right)\big)}{n-1}\right|\right)$$

3.3. Dimension Reduction and Feature Fusion

This section focuses on fusing the useful features into a single health indicator. All useful features are in time series and they operate under different conditions. Before performing dimension reduction, it is a good practice to normalize the useful features to the same scale. Thus, z-score normalization is employed in this study. In the literature, principal component analysis (PCA) which is a statistical dimensionality reduction procedure has been widely used to transform a set of correlated input variables possibly into a set of linearly uncorrelated principal components. In addition, the first principal component is used as the health indicator [20].

3.4. Fit Degradation Model

The main purpose of this study is to estimate the RUL of a bearing. The term RUL refers to the expected lifetime of a system or a machine before it needs maintenance. A typical model that estimates the RUL is based on the time evolution of the health indicator and predicts the remaining time until the health indicator crosses some threshold value indicative of a fault condition. In this study, a degradation model is proposed to estimate the RUL for bearing running at high speed and constant load. The model considers the health indicator, in other words, the computed first principal component, as a time series; and hence, the first 200 values are employed to develop an initial time series model, then once 10 new values are available, use the last 100 values to update the time series model. The updated time series model is used to compute a 10 step ahead forecast.

4. Case Study & Results

In this study, a method based on the remaining useful life estimation using the vibration signal generated in the bearing was proposed. The vibration signals were obtained from a ball bearing signal simulator with a single point defect on the outer race of the bearing. The vibration signals were collected at 10 min intervals and a total number of 600 data sets were recorded for analysis. The vibration data were sampled at a sampling rate of 20 kHz for 1 s.

It has been mentioned previously that due to measurement noise, variation of operational conditions, and the stochastic nature of the degradation processes, the raw vibration signals cannot be directly used as the inputs of the prediction models. Therefore, a median filter was utilized to remove the high-frequency noise components. Healthy and faulty bearing signal and their filtered values are presented in Fig. 2a, where the peak amplitude gradually increases towards failure. Furthermore, each vibration segment is transformed into the frequency-domain using the Welch method. The power spectrum density (i.e. *pxx*) is considered a powerful tool for vibration signal analysis in frequency-domain, and its changes over time are shown in Fig 2b.

The change in acceleration in the time-domain manifests itself in the power spectral density in the frequency-domain. Failure occurs in the bearing as time passes. The colorbar indicates the power spectral density in Fig. 2b. It is observed that the power spectral density value between 1.5 kHz and 4 kHz gradually increases as the machine condition degrades. The peak value 0.53 dB/Hz was obtained in 2.8 kHz at 6000 min.

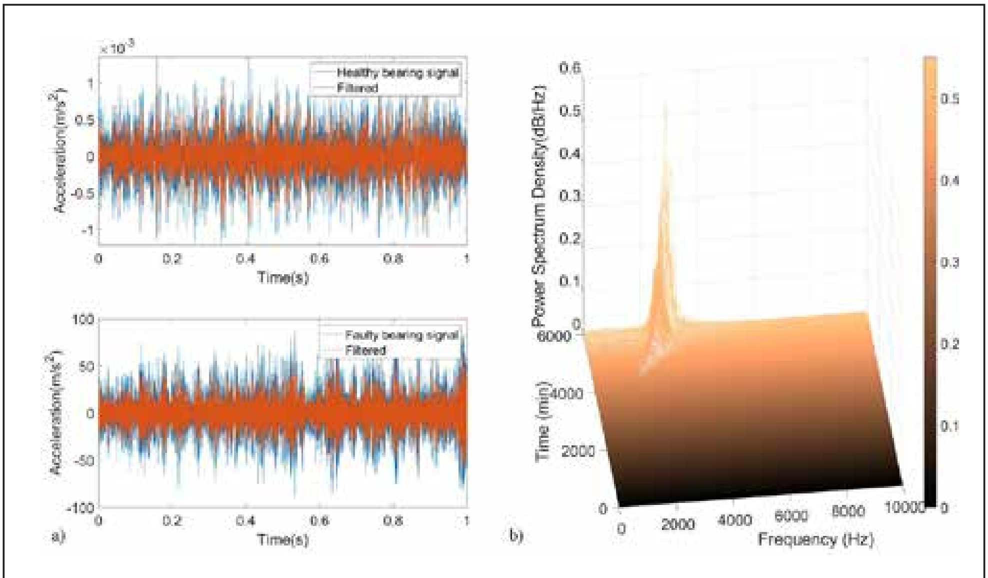


Fig. 2. a) The gathered vibration signals, b) Power spectrum density changes over time

The statistical features of the collected vibration signal in time-domain as well as the statistical features of the power spectrum and mean peak frequency will be potential indicators of the bearing degradation. The change of some properties in both the time domain and the frequency domain with respect to time is depicted in Fig. 3. It can be easily detected from all subplots that there is a threshold value, 3000 min to distinguish the trend change.

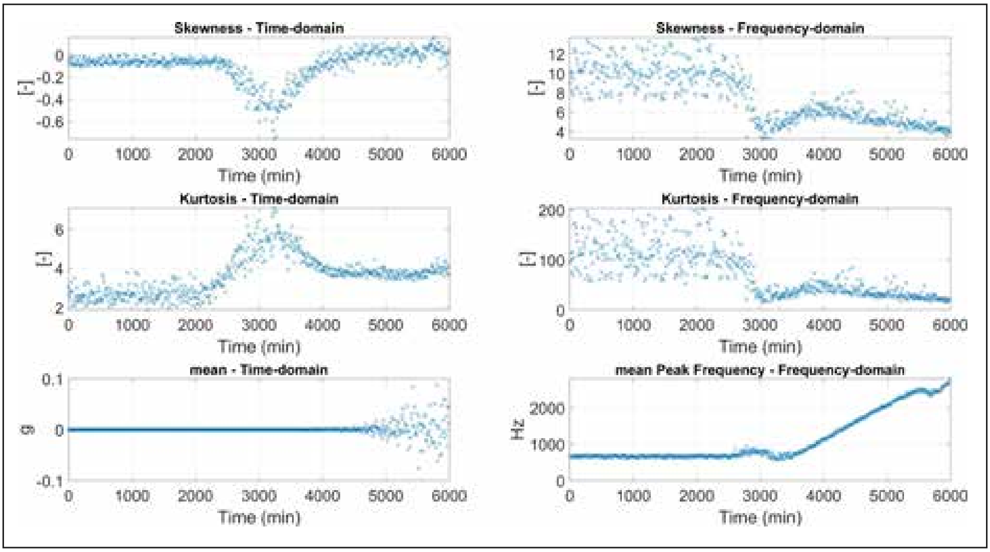


Fig. 3. The change of some extracted features over time

Skewness is a measure of the asymmetry of the data around its mean. The skewness value in the time-domain is almost constant at zero before 3000 min, but suddenly it reaches negative values around 3000 min, then it returns to zero again. Similarly, kurtosis highlights sharp peak signals in the data. The smallest kurtosis values in the time-domain are found below 3000 min, and after this time kurtosis values are saturated around four. On the other hand, a scattered distribution before 3000 min and a more compact distribution after 3000 min were observed for both kurtosis and skewness value in the frequency-domain. This leads to the conclusion that the fluctuation area occurring at 3000 min can act as a boundary between faulty and error-free runs. Furthermore, the computed mean peak frequency is flat before 3000 min and it gradually increases after the threshold value. In some cases, the occurrence of background noise and other sources of vibration signals may prevent bearing faults from being detected through the observation of changes. Therefore, it is required to obtain features that are more correlated with the degradation pattern. In this study, monotonicity was utilized to evaluate the computed features that were given in Table 1. A monotonicity value close to one means that the feature is monotonic and suitable for RUL prediction, whereas a monotonicity value close to zero indicates that the feature is non-monotonic and not appropriate for RUL prediction [9]. The result of monotonicity is given in Fig. 4. A value of 0.1 was set as the threshold to select the appropriate features. As a result, the peak-to-peak, standard deviation (std), square root mean (RMS), energy features computed from the vibration signals in the time-domain, and the mean value derived from the power spectral density (PXXMean) in the frequency-domain were selected as appropriate features.

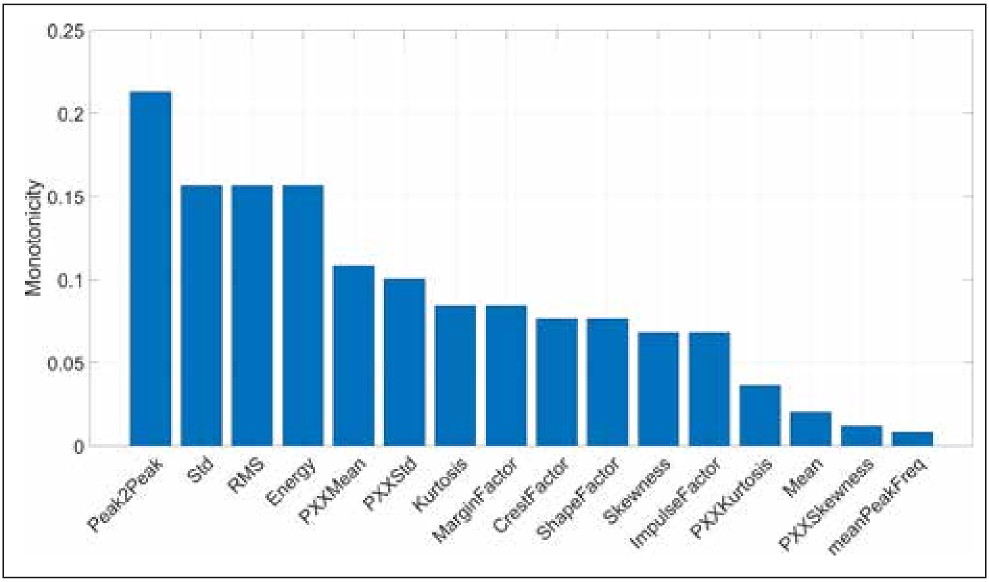


Fig. 4. Evaluation of the computed features using monotonicity

Principal Component Analysis (PCA) is utilized for dimension reduction and feature fusion in this study. As a result of the PCA, the first principal component is employed as the health indicator for RUL estimation, which is shown in Fig. 5. It indicates that the first principal component is exponentially increasing as the machine approaches failure. In the next step, the first principal component is considered as a time series and a degradation model is developed using forecasting methodology. The degradation model estimates the RUL by predicting when a monitored signal will cross a predefined threshold. The selection of threshold is usually based on the historical records of the machine or some domain-specific knowledge. In this study, 0.1% of the peak value of PCA is taken as the threshold value. The degradation model developed is a dynamic model that updates itself as new data becomes available. An alarm is set to give a warning when it reaches 95% of the threshold value. As the results show that the alarm is triggered at 4790 min which states that maintenance is required for the bearing. Furthermore, it is estimated that the RUL is 70 min to reach the failure threshold.

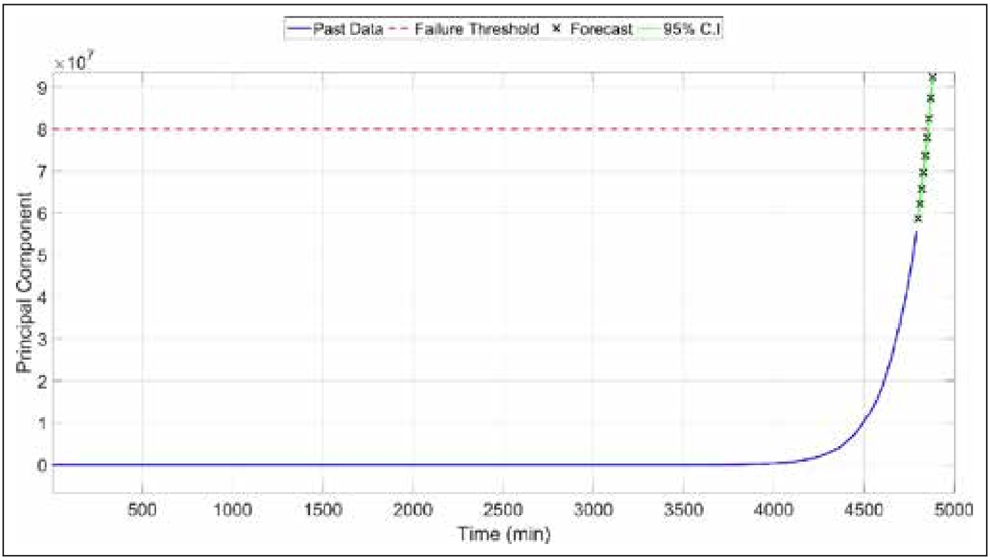


Fig. 5. Estimated values of the first principal component using the degradation model

5. Conclusions

Estimating the RUL is considered a challenge for real-world problems such as aircraft engines, high-speed shaft bearing, etc. This paper developed a new strategy to determine the machine health indicator from vibration data using machine learning methodology and then employed the computed health indicator in the remaining useful life estimation of bearings. The vibration signal was collected in the time-domain and was transformed to the frequency-domain using the Welch method. The traditional statistical indicators such as mean value, standard deviation, skewness, kurtosis, etc. were utilized to extract features in both domains. Monotonicity analysis was conducted to compare the efficiency of the computed features, and the results showed that peak-to-peak, standard deviation, root mean square, energy and mean power spectral density value were able to detect early faults. The selected features were fused into a single health indicator using PCA and an exponential degradation model was developed using the computed health indicator. The achieved results indicated that 70 min was remaining to reach the failure threshold. In fact, at the time when the alarm was triggered, the defect depth was 0.8 mm, but there is another 1210 minutes before the 3mm of damage occurred.

References

[1] Nejad, A. R., Odgaard, P. F., Gao, Z., & Moan, T. (2014). A prognostic method for fault detection in wind turbine drivetrains. Engineering Failure Analysis, 42, 324-336.

[2] Teng, W., Han, C., Hu, Y., Cheng, X., Song, L., & Liu, Y. (2020). A Robust Model-Based Approach for Bearing Remaining Useful Life Prognosis in Wind Turbines. IEEE Access, 8, 47133-47143.

[3] Kizrak, M. A., & Bolat, B. (2019). Uçak Motoru Sağlığı için Uzun-Kısa Süreli Bellek Yöntemi ile Öngörücü Bakım. Bilişim Teknolojileri Dergisi, 12(2), 103-109.

[4] Ahmad, R., & Kamaruddin, S. (2012). An overview of time-based and condition-based maintenance in industrial application. Computers & industrial engineering, 63(1), 135-149.

[5] Wei, J., Dong, G., & Chen, Z. (2017). Remaining useful life prediction and state of health diagnosis for lithium-ion batteries using particle filter and support vector regression. IEEE Transactions on Industrial Electronics, 65(7), 5634-5643.

[6] Elforjani, M., Shanbr, S., & Bechhoefer, E. (2018). Detection of faulty high speed wind turbine bearing using signal intensity estimator technique. Wind Energy, 21(1), 53-69.

[7] Carvalho, T. P., Soares, F. A., Vita, R., Francisco, R. D. P., Basto, J. P., & Alcalá, S. G. (2019). A systematic literature review of machine learning methods applied to predictive maintenance. Computers & Industrial Engineering, 137, 106024.

[8] Özkat, E . (2021). Makine Öğrenmesi Metodolojisi Kullanılarak Yüksek Hızlı Rulmanlarda Sağlık Göstergesinin Belirlenmesi . Avrupa Bilim ve Teknoloji Dergisi , Ejosat 2021 Ocak , 176-183 . DOI: 10.31590/ejosat.843465

[9] Elasha, F., Shanbr, S., Li, X., & Mba, D. (2019). Prognosis of a wind turbine gearbox bearing using supervised machine learning. Sensors, 19(14), 3092.

[10] Selcuk, S. (2017). Predictive maintenance, its implementation and latest trends. Proceedings of the Institution of Mechanical Engineers, Part B: Journal of Engineering Manufacture, 231(9), 1670-1679.

[11] Zheng, Y. (2019). Predicting remaining useful life based on Hilbert–Huang entropy with degradation model. Journal of Electrical and Computer Engineering, 2019.

[12] Zupan, B., Demšar, J., Kattan, M. W., Beck, J. R., & Bratko, I. (2000). Machine learning for survival analysis: a case study on recurrence of prostate cancer. Artificial intelligence in medicine, 20(1), 59-75.

[13] Sikorska, J. Z., Hodkiewicz, M., & Ma, L. (2011). Prognostic modelling options for remaining useful life estimation by industry. Mechanical systems and signal processing, 25(5), 1803-1836.

[14] Li, Z., Goebel, K., & Wu, D. (2019). Degradation modeling and remaining useful life prediction of aircraft engines using ensemble learning. Journal of Engineering for Gas Turbines and Power, 141(4).

[15] Ramasso, E., & Gouriveau, R. (2014). Remaining useful life estimation by classification of predictions based on a neuro-fuzzy system and theory of belief functions. IEEE Transactions on Reliability, 63(2), 555-566.

[16] Bektas, O. (2018). Çok Boyutlu Turbofan Motoru Durum İzleme Verilerinde Prognostik Uygulamalarının Değerlendirilmesi. VII. Ulusal Havacılık ve Uzay Konferansı, (pp. 1-10). Samsun

[17] Bektas, O., Jones, J. A., Sankararaman, S., Roychoudhury, I., & Goebel, K. (2019). A neural network framework for similarity-based prognostics. MethodsX, 6, 383-390.

[18] Bektas, O., Jones, J. A., Sankararaman, S., Roychoudhury, I., & Goebel, K. (2019). A neural network filtering approach for similarity-based remaining useful life estimation. The International Journal of Advanced Manufacturing Technology, 101(1), 87-103.

[19] Mathworks. (2021). Condition Monitoring and Prognostics Using Vibration Signals. Retrieved May 24, 2021, from https://uk.mathworks.com/help/predmaint/ug/wind-turbine-high-speed-bearing-prognosis.html#WindTurbineHighSpeedBearingPrognosisExample-4

[20] Holland, S. M. (2008). Principal components analysis (PCA). Department of Geology, University of Georgia, Athens, GA, 30602-2501.

Creating a Unified Data Collection System for Infrastructure Projects

[1]Dr. Yunus Emre Ayözen

General Manager of Strategy Development Department in Republic of Turkey Ministry of Transport and Infrastructure

E-mail: yunusemre.ayozen@uab.gov.tr

[1]Bilge Büdeyri, M.Sc.

Head of Strategy Department at Strategy Development Department in Republic of Turkey Ministry of Transport and Infrastructure

E-mail: bilge.budeyri@uab.gov.tr

[2] A. Sinan Alemdar, M.Sc.

Program Manager at EMAY International Engineering and Consultancy

E-mail: salemdar@emay.com

[3]Ugur Kaya M.Sc.

Design Coordinator at EMAY International Engineering and Consultancy

E-mail: ugur.kaya@emay.com

[4]Assoc. Prof. Dr. Onur Behzat Tokdemir

Dept. of Civil Engineering, Middle East Technical Univ., (corresponding author).

E-mail: tokdemir@metu.edu.tr

Abstract

Infrastructure projects are the most critical part of development in any country. As Turkey is developing very fast, many projects are in the works, from speed rails to highways. The Ministry of Transport and Infrastructures manages all these mega-projects under five different infrastructure divisions: Aviation, Communication, Maritime, and Railway to Roads. Even though a unified GIS (Geographic Information System) provides current asset information for these departments, managing ongoing projects is done with separate systems with separate databases. This study aims to explain the underlying principles of a unified database system for all these agencies to manage ongoing mega-projects. The new system will provide unified access and support necessary executive dashboards, KPIs (Key Performance Index), and field reports. The system is called UAB-PYS during development. UAB-PYS will enable better control of all infrastructure projects and create a big data environment. While gathering new data, all the existing datasets will be consolidated. The historical information will also be used in future AI (Artificial Intelligence) based management systems.

Keywords: Infrastructure, PMS, Big Data, AI.

1. Introduction

Transportation and communication became one the most critical activities for any nation. The growing human population is urging for more systematic and accessible transportation webs throughout the globe. Every day, humankind steps out beyond exploring other cultures and lifestyles—furthermore, the digital and physical movement of the public fuels economic growth. The technology of today makes it possible to design with information to clarify the truth of requirements and future success. Especially with the pandemics, improvements in telecommunication infrastructure became more urgent than ever as people were bound to stay home.
The Ministry of Transportation and Infrastructure of Turkey operates and leads many sectors, such as aviation, highways, and telecommunication, founded in 1939. With the vision of being an institution that provides safe transportation and fast access, The Ministry reached its final organizational structure in 2018. Turkey has large cities with a growing population and eventually connects more than ever due to growing technology as fast as our population. The Ministry of Transportation and Infrastructure of Turkey carefully analyses the processes and follows the world each day, finds better solutions for the citizens, works upon more efficient transportation and communication systems, and reaches out to the voices and cares for public opinions.

Each sectoral investment and policy about investment has been especially implementing for the last 20 years. Turkey's infrastructure has been completely renewed in transportation by investing 932 billion Turkish liras in this period. In 2021, 30.7 % of the total national budget is shared to develop the transportation and communication sector. The distance of the divided road in the highway sector has been reached 28,200 kilometers, and the length of the highway has been increased from 1710 kilometers to 3 thousand 523 kilometers.

Also, one of the most significant changes could be seen in the aviation sector in Turkey. Aviation number is reached to 56, and Istanbul Airport, which opened in 2018, will be able to accommodate 200 million passengers a year by using a single terminal. These numbers make it the world's largest terminal. Also, many railway projects will be implemented in line with the 2023 targets of the country. Turkey will have a total railway length of 17,525 km after completion of a 5,509-km long new line, and the length of signaled lines will reach 6,382 km. Turkey has been making investments in its existing ports. Large projects increase the capacity of the country and meet the freight transportation needs. In this respect, one of the new ports in İzmir, called Candarli port, will be the biggest port in Europe and the 10th biggest in the world. The port will have a capacity of 12 million TEU capacity for freight transport.

Another important investment decision about information and communication technologies has been taking by the Republic of Turkey Ministry of Transport and Infrastructure. According to the address-based population, the young population (15-24 age) made up 15.6% of the total population of Turkey. This shows potential as a technology-driven young population continues to buy consumer electronics, software, and products. The number of broadband subscribers reached 85 million, and the number of mobile phone subscribers rose to 84 million. Many of the institutional Works have been transferred to electronic platforms, and today electronic government platform users approach 54 million.

Based on this idea, the Ministry of transportation and infrastructure is willing to provide safer, faster, and efficient transportation by using engineering and technology to allow people to live better and contribute to the country's development in transportation, maritime, communication, space, and information technologies. The Ministry operates epochal mega projects upon infrastructure sectors, such as the İstanbul Grand Airport project hosting 200 million passengers and flights to more than 300 different destinations in 2028.

2. The New Infrastructure Project Management System

There are many databases used in the Ministry, most of which created as the needs arose in different years. However, with many mega projects in the pipeline, there is a need for a new Project Management System. This system will be fed data from different databases, as seen in Figure 1. A unified solution is needed because as the size and complexity of the projects increased; there is an urgency for faster and easy-to-use systems. The design of the new system called UAB-PYS started in January 2021 (Figure 2).

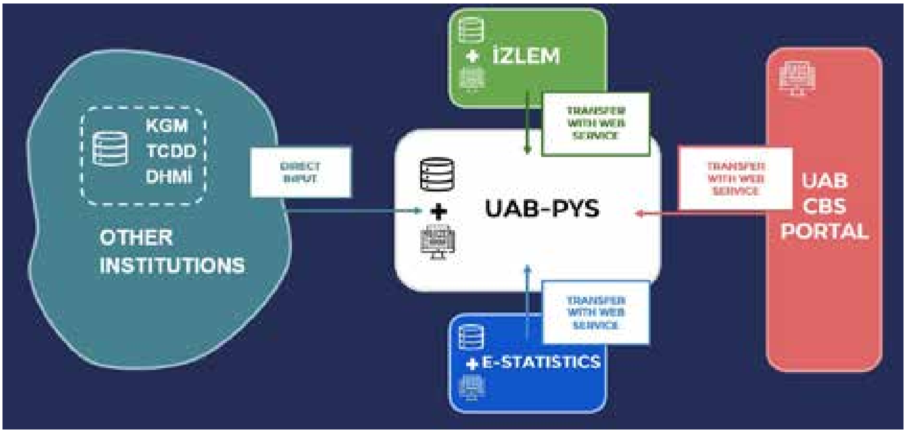


Fig. 1. New Structure of New Project Management System (UAB-PYS)

The new consolidated management system covers all necessary themes of ministry for the program management. Previously developed IZLEM is used to manage project funds monthly. The new portal, UAB-PYS will get connected with already running system. The new interconnection between both databases is helping to flow cost information from IZLEM to UAB-PYS. GIS and statistics data are also connected to system in the same manner. The other departments, which do not have any database management platform are also enriched with the new system without additional effort.

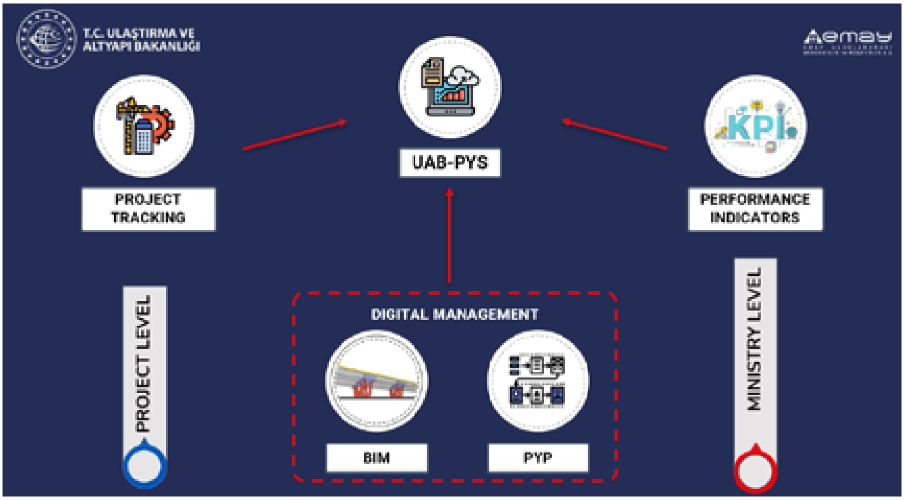


Fig. 2. Unified System (UAB-PYS)

The executive dashboard is designed to present required information on different levels of detail. As the ministry requirements the information is sorted out from detailed levels. On the other hand, project manager of sub sectors can get access to more detailed information, that aids decision making in action. The variety of different levels of detail is managed by adaptation of unified code to all projects in all sub sectors. The new unified code breaks down the works until it reaches only to the elements, which has significant effects on cost and time (Figure 3).

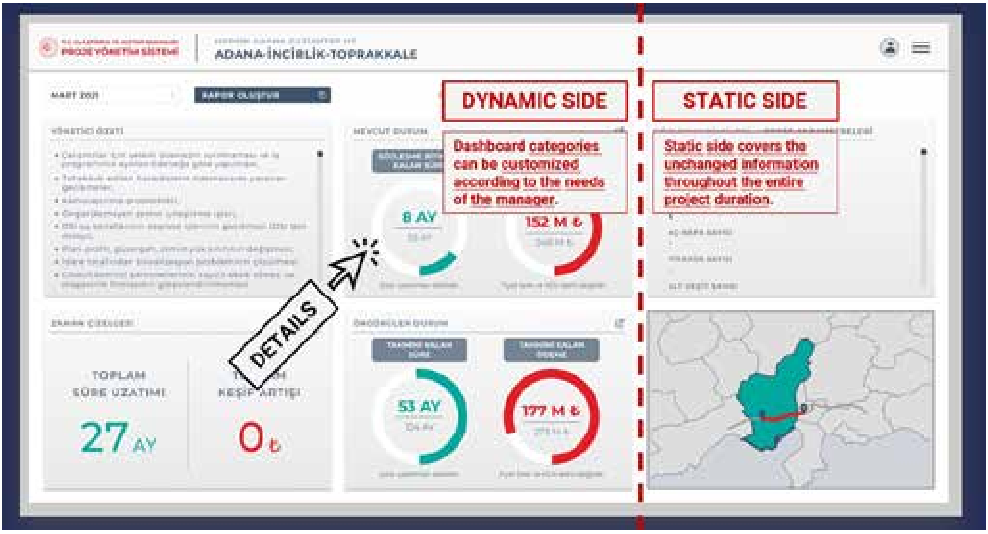


Fig. 3. Executive Dashbord

In spite of all efforts in technological changes, the new system has also be supported well enough to manage all information flow all around to Turkey. The social habits were one of barriers against on adaptation to new changes on work breakdown structure. Therefore, supporting documents were transmitted, that transcribe what to do and how to reach technical requirements in and between organizations with new cultural aspect.

The new documents supported byUAB-PYS are published in May 2021. The first one is for Project Management Procedure (Figure 4). This procedure underlines all the rules to use the system and streamlines the rules to report and control all of the projects (https://sgb.uab.gov.tr/dijital-yonetim). This procedure has covered:

- Scope Management,
- Time Management,
- Cost Management,
- Quality Management,
- Information Management,
- Document Management,
- Risk Management,
- Procurement Management,
- Stakeholder Management, and
- Health, Safety and Environment Management.

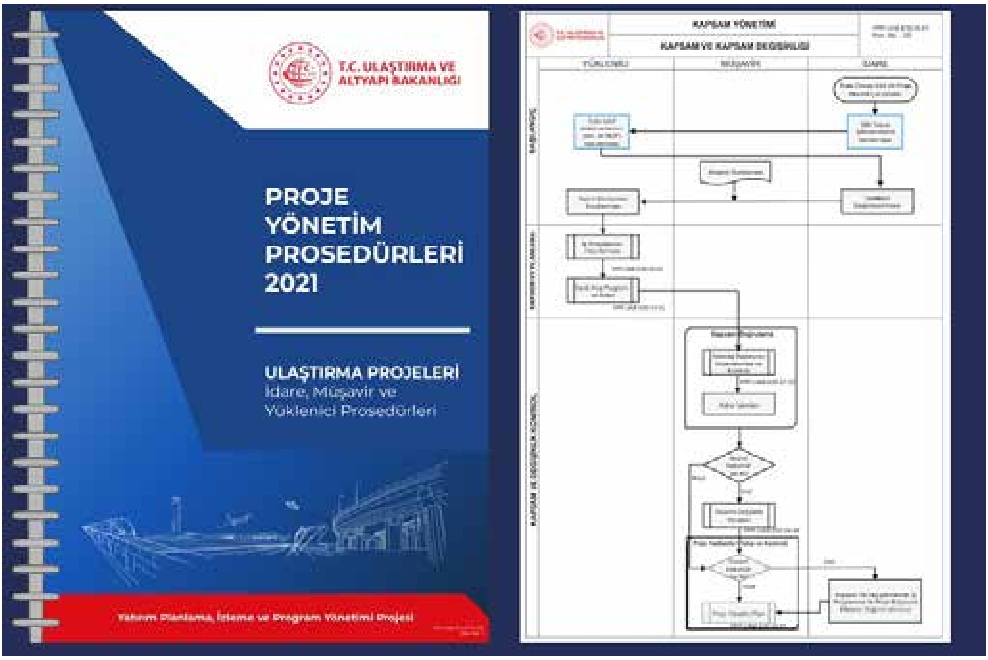


Fig. 4. Project Management Procedure

As part of the UAB-PYS, there are also BIM (Building Information Modeling) Technical Specifications published (Figure 5). This document covered (https://sgb.uab.gov.tr/dijital-yonetim):

- 3D Visualization,
- 2D Construction Documents,
- 3D Coordination,
- Quantity Take-off and Budget Preparation,
- 4D Planning,
- Engineering Analysis,
- Facility Management,
- Modeling Standards,
- Reference Requirements, and,
- Information Management Requirements.

Fig. 5. BIM Technical Specifications and Bid Documents

As these procedures and technical specifications are introduced, training sessions and follow-up meetings were performed. However, the system is still too young to be fully executed in all projects. It will be completed at the end of the year 2021. When it is finished, it will provide management data for both project-level and ministry levels (Figure 6).

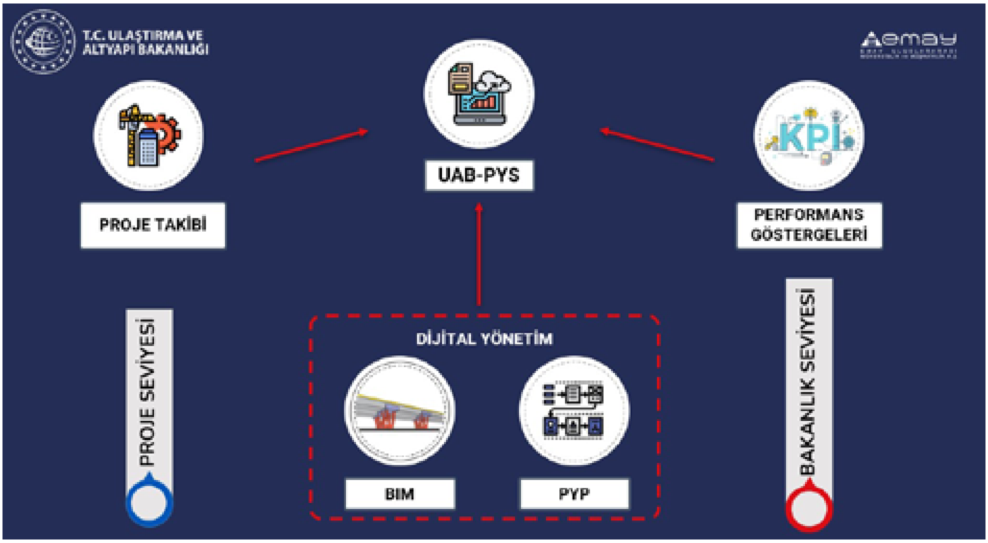


Fig. 6. UAB-PYS System Architecture

3. Conclusion

As the Republic of Turkey's infrastructure is growing fast, there is an urgent need for Project Management Systems. With the use of UAB-PYS, many problems can be solved, cost and time overruns can be predicted early, and millions of dollars can be saved. Therefore the success of this system is crucial for the management of upcoming mega infrastructure projects. The success of the system also requires a lot of training and follow-up, as it will take users to get accustomed to the system. Once it is fully up and running, UAB-PYS will make an amazing contribution to the management of infrastructure projects.

Acknowledgments

This study is supported by Emay Engineering and Consulting. The authors would like to thank for the given data, expert opinion, and the fund given to researchers.

References

- Project Management Procedure, Accessed on 30-5-2021, https://sgb.uab.gov.tr/dijital-yonetim
- BIM (Building Information Modeling) Technical Specifications, Accessed on 30-5-2021, https://sgb.uab.gov.tr/dijital-yonetim

ICAIBDEA International Conference on AI and Big Data in Engineering Applications 2021

IAU Florya Campus Istanbul, Turkey June 14-15, 2021

ICAIBDEA is an acronym for International Conference on Artificial Intelligence (AI) and Big Data in Engineering Applications. Artificial intelligence, collection and control of big data are widely used for practical engineering applications for complex problems and decision support systems. Proceedings include case studies, innovative techniques in engineering applications strongly welcome to our conference. The principal mission of ICAIBDEA is bridging the gap between universities and sectors, encouraging interdisciplinary research applications and bringing multidiscipline researchers together.

CONFERENCE THEME

Topics covered include engineering applications in industrial product design, industrial engineering, engineering management/disaster management, energy systems, smart materials, smart systems, smart cities, transportation, cyber security, agriculture, mechatronics, robotics, control systems/monitoring/GIS, civil engineering/structures, infrastructures, information systems/expert systems, bioengineering/bioinformatics, computer aided design and applications, database/data mining, and other AI and Big Data analytics engineering applications.

CHAIRS AND COMMITTEES

Conference Committee Chair and Editor: Prof.Dr. Can BALKAYA

International Program & Evaluation Committee:

Prof.Dr Can BALKAYA (Turkey)	İstanbul Aydın University
Prof.Dr. Ali OKATAN (Turkey)	İstanbul Aydın University
Prof.Dr. Mehmet GÜLGÜN (Turkey)	Sabancı University
Prof.Dr. Erol TUTUMLUER (USA)	University of Illinois at Urbana-Champaign, IL
Prof.Dr. İlhami ÇOLAK (Turkey)	Nişantaşı University
Asst.Prof.Dr. Serkan AYVAZ (Turkey)	Bahçeşehir University
Prof.Dr. Oğuzhan HASANÇEBİ (Turkey)	Middle East Technical University
Assoc.Prof.Dr. Onur Behzat TOKDEMİR (Turkey)	Middle East Technical University
Assoc.Prof.Dr. Gökçe SARGUT (USA)	Governors State University, University Park, IL
Dr. Ahmet ÇITIPITIOĞLU, PE (Turkey)	TAV Construction
Dr. Erol KALKAN, PE (USA)	QuakeLogic Inc., Roseville, CA, USA
Prof.Dr. Oktay ÖZCAN (Turkey)	İstanbul Aydın University
Prof.Dr. Ömer Faruk ERKENDİRCİ (Turkey)	İstanbul Aydın University
Prof.Dr. Zeynep Dilek HEPERKAN (Turkey)	İstanbul Aydın University

ISTANBUL AYDIN UNIVERSITY PUBLICATIONS

ISBN 978-6257783286