



**International Conference on Interdisciplinary Applications
of Artificial Intelligence 2021**

PROCEEDINGS BOOKS

Volume 3
EXTENDED ABSTRACT BOOK

e-ISBN: 978-625-00-9716-8

EDITORS

Dr. Yunus ÖZEN
Dr. Alperen BAL
Dr. Ferhat SAYIM
Dr. Murat GÖK

Prepared By
Yasemin KORKUT

PREFACE

International Conference on Interdisciplinary Applications of Artificial Intelligence 2021 (ICIDAAI'21) was organized online from 21-23 May 2021. This was the first virtual conference which was organized in collaboration with Yalova University, Istanbul University, Kocaeli University, Sakarya University, Bursa Technical University, UET Lahore University, Matej Bel University, University of Tlemcen, Universite 8 Mai 1945 Guelma, International Vision University, Bulgarian Academy of Sciences, Lahore Leads University and IQRA National University.

There were 106 presentations for the virtual conference. A secured platform was used for virtual interactions of the participants. After the virtual conference, there was a call for full papers to be considered for publication in the conference proceedings. Manuscripts were received and they were processed and peer reviewed. These manuscripts cover a range of Artificial Intelligence topics from social sciences to physical sciences.

We hope that these chapters will add to literature, and they will be useful references. To conclude, ICIDAAI'21 was a successful event, and we would like to thank all those who have contributed. We would also like to thank the Organizing and International Advisory committee members, the participants, and the reviewers.

Editors

COMMITTEES

Honorary Board

- Dr. Suat CEBECİ (Yalova University)
- Dr. Mahmut AK (Istanbul University)
- Dr. Fatih SAVAŞAN (Sakarya University)
- Dr. Sadettin HÜLAGÜ (Kocaeli University)
- Dr. Dragomir DOBRUDZHALIEV (Bulgarian Academy of Science)
- Dr. Ellagoune Salah (University 8 Mai 1945 Guelma)
- Dr. Türkay DERELİ (Iskenderun Technical University)
- Dr. Zekai ŞEN (Medipol University)

Advisory Board

- Dr. Alpaslan FIĞLALI (Kocaeli University)
- Dr. Cemalettin KUBAT (Sakarya University)
- Dr. Harun TAŞKIN (Sakarya University)
- Dr. İsmail ADAK (Yalova University)
- Dr. İsmail HAKKI CEDİMOĞLU (Sakarya University)
- Dr. M. Melih İNAL (Kocaeli University)
- Dr. Metin ÇANCI (Yalova University)
- Dr. Mustafa ÖZTAŞ (Yalova University)
- Dr. Nilgün FIĞLALI (Kocaeli University)
- Dr. Orhan TORKUL (Sakarya University)
- Dr. Sibel AKOVA (Yalova Üniversitesi)
- Dr. Rafet BOZDOĞAN (Yalova University)
- Dr. Sevinç GÜLSEÇEN (Istanbul University)
- Dr. Tauseef Aized KHAN (UET Lahore, Pakistan)
- Dr. Yacine LAFIFI (University 8 May 1945 Guelma, Algeria)
- Dr. Zerrin ALADAG (Kocaeli University)

Organizing Committee

- Dr. Alperen BAL (Yalova University)
- Dr. Çiğdem SELÇUKCAN EROL (Istanbul University)
- Dr. Ferhat SAYIM (Yalova University)
- Dr. Halil İbrahim CEBECİ (Sakarya University)
- Dr. İhsan Hakan SELVİ (Sakarya University)
- Dr. Murat GEZER (Istanbul University)
- Dr. Muhammad Saeed Khan (Lahore Leads University)
- Dr. Murat GÖK (Yalova University)
- Dr. Safiye TURGAY (Sakarya University)
- Dr. Yunus ÖZEN (Yalova University)
- Dr. Zeki ÖZEN (Istanbul University)
- Dr. Selçuk Öğütçü (Kocaeli University)
- Yasemin KORKUT (Yalova University)
- Özgür ÖZKAYA (Yalova University)
- Merve ŞİŞÇİ (Sakarya University)
- Bilal TORKUL (Yalova University)

Scientific Committee

- Dr. Alptekin Erkollar (ETCOP U. Digitalization Department Austria)
- Dr. Birgit Oberer (Austrian Government, Ministry of Education)
- Dr. Rashmi Gujrati (KC Group of Institutions, India)

- Dr. Selim Zaim (Istanbul Şehir Üniversitesi)
- Dr. Yaşar Sarı (Eskişehir Osmangazi Üniversitesi)
- Dr. Abdulgaffar AĞAOĞLU (Yalova University)
- Dr. Abdulkadir HIZIROĞLU (İzmir Bakırçay University)
- Dr. Abdülkadir TEPECİK (Yalova University)
- Dr. Abdullah TAŞKESEN (Bingöl Üniversitesi)
- Dr. Adem GÖLEÇ (Erciyes University)
- Dr. Adem TUNCER (Yalova University)
- Dr. Adil BAYKAŞOĞLU (Dokuz Eylül University)
- Dr. Ahmet Kürşad TÜRKER (Kırıkkale University)
- Dr. Ahmet ÖZMEN (Sakarya University)
- Dr. Ahmet Yağmur ERSOY (Sakarya University)
- Dr. Ajith ABRAHAM (Machine Intelligence Research Labs)
- Dr. Alfred VELLA (Open University)
- Dr. Alperen BAL (Yalova University)
- Dr. Andrew KUSIAK (University of Iowa)
- Dr. Ayhan DEMİRİZ (Gebze Technical University)
- Dr. Ayhan İSTANBULLU (Balıkesir University)
- Dr. Ayşe Tansel ÇETİN (Yalova University)
- Dr. Barış KANTOĞLU (Düzce Üniv.)
- Dr. Bendjebar SAFIA (University 8 Mai 1945 Guelma)
- Dr. Birol ELEVİLİ (19 Mayıs Üniversitesi)
- Dr. Burcu OKKALIOĞLU (Yalova University)
- Dr. Buket KARATOP (Istanbul Cerrahpaşa University)
- Dr. Celal ÇEKEN (Sakarya University)
- Dr. Cem Sefa SÜTÇÜ (Marmara University)
- Dr. Cemalettin KUBAT (Sakarya University)
- Dr. Cemil ÖZ (Sakarya University)
- Dr. Cercis İKİEL (Sakarya University)
- Dr. Cevdet KIZIL (İstanbul Medeniyet University)
- Dr. Çiğdem SELÇUKCAN EROL (Istanbul University)
- Dr. Derya DELİKTAŞ (Dumlupınar Üniversitesi)
- Dr. Derviş KARABOGA (Erciyes University)
- Dr. Devrim AKGÜN (Sakarya University)
- Dr. Duc Truong PHAM (University of Birmingham)
- Dr. Ebubekir KOÇ (Istanbul Fatih Sultan Mehmet University)
- Dr. Ecir Uğur KÜÇÜKSİLLE (Isparta Süleyman Demirel University)
- Dr. Elif CESUR (Medeniyet Üniversitesi)
- Dr. Elif KARTAL (Istanbul University)
- Dr. Emrah CENGİZ (Istanbul University)
- Dr. Emrah DOĞAN (Sakarya University)
- Dr. Emre AKADAL (Istanbul University)
- Dr. Emre DANDIL (Bilecik Şeyh Edebali University)
- Dr. Ercan ÖZTEMEL (Marmara University)
- Dr. Erman COŞKUN (İzmir Bakırçay University)
- Dr. Eyüp ÇALIK (Yalova University)
- Dr. F. Önay KOÇOĞLU (Istanbul University)
- Dr. Fatih ÇALIŞKAN (Düzce University)
- Dr. Fatih ÇALLI (Sakarya University)
- Dr. Fatih HOCAOĞLU (Medeniyet Üniversitesi)
- Dr. Farzand Ali JAN (IQRA National University)
- Dr. Fehim FINDIK (Sakarya University of Applied Sciences)
- Dr. Ferdi SÖNMEZ İstanbul Arel University
- Dr. Ferhat SAYIM (Yalova University)
- Dr. Fidan BAYRAKTUTAN (Istanbul University)
- Dr. Fuat KOSANOĞLU (Yalova University)
- Dr. Georgi ANGELOV (Bulgarian Academy of Science)

- Dr. GülGönül BOZOĞLU (Batı Yalova University)
- Dr. Gökhan SİLAHTAROĞLU (Medipol University)
- Dr. Gültekin ÇAĞIL (Sakarya University)
- Dr. Güneş GENÇYILMAZ (Istanbul Aydın University)
- Dr. Gürkan ÖZTÜRK (Eskişehir Teknik Üniversitesi)
- Dr. Güneş HARMAN (Yalova University)
- Dr. Haldun AKPINAR (Marmara University)
- Dr. Halil İbrahim CEBECİ (Sakarya University)
- Dr. Halimi KHALED (University 8 Mai 1945 Guelma)
- Dr. Hamid ARABNIA (University of Georgia)
- Dr. Harun Reşit YAZGAN (Sakarya University)
- Dr. Hasan Erdiñç KOÇER Konya Selçuk University
- Dr. Hikmet ERBIYIK (Yalova University)
- Dr. Hüseyin GÜRÜLER (Muğla Sıtkı Koçman University)
- Dr. İbrahim SABUNCU (Yalova University)
- Dr. İkram DAŞTAN (Yalova University)
- Dr. İhsan EROZAN (Dumlupınar Üniv.)
- Dr. İhsan Hakan SELVİ (Sakarya University)
- Dr. İhsan TOKTAŞ (Yıldırım Beyazıt Üniversitesi)
- Dr. İsmail EKİNCİOĞLU (Dumlupınar Üniv.)
- Dr. İrfan KÖSESOY (Yalova University)
- Dr. Kürşat TÜRKER (Kırıkkale University)
- Dr. Levent ÇALLI (Sakarya University)
- Dr. M. Furkan KIRAÇ (Özyeğin University)
- Dr. Maged M. DESSOUKY (University of Southern Carolina)
- Dr. Malgorzata PANKOWSKA (University of Economics in Katowice, Poland)
- Dr. Mehenaoui ZOHRA (University 8 Mai 1945 Guelma)
- Dr. Mehmet Emin AYDIN (University of the West of England, Bristol)
- Dr. Mehmet Emin TAĞLUK (İnönü Üniversitesi)
- Dr. Mehmet TOPAKÇI (Akdeniz Üniversitesi)
- Dr. Mete GÜNDOĞAN (Ankara Yıldırım Beyazıt University)
- Dr. Metin SAYGILI (Sakarya University of Applied Sciences)
- Dr. Mine AKSOY KAVALCI (Yalova University)
- Dr. Muhammed Kürşad UÇAR (Sakarya University)
- Dr. Muharrem DÜĞENCİ (Karabük University)
- Dr. Murat AZALTUN (Yalova University)
- Dr. Murat GEZER (Istanbul University)
- Dr. Murat GÖK (Yalova University)
- Dr. Murat OKKALIOĞLU (Yalova University)
- Dr. Mustafa YILMAZ (Sakarya University)
- Dr. Müfit ÇETİN (Yalova University)
- Dr. Natalija LEPKOVA (Vilnius Gediminas Technical University, Lithuania)
- Dr. Nejat YUMUŞAK (Sakarya University)
- Dr. Nizami AKTÜRK (Gazi Üniversitesi)
- Dr. Numan ÇELEBİ (Sakarya University)
- Dr. Osman Hilmi KOÇAL (Yalova University)
- Dr. Osman Nuri ŞAHİN (Alanya Alaaddin Keykubat University)
- Dr. Ömer Kadir MORGÜL (Sakarya Üniversitesi)
- Dr. Özden ÜSTÜN (Dumlupınar Üniversitesi)
- Dr. Özer UYGUN (Sakarya University)
- Dr. Raşit CESUR (Sakarya University)
- Dr. Robertas DAMASEVICIUS (Kaunas University of Technology, Lithuania)
- Dr. Selami ÖZCAN (Yalova University)
- Dr. Semra BORAN (Sakarya University)
- Dr. Serdar SOLAK (Kocaeli University)
- Dr. Serkan AYVAZ (Bahçeşehir Üniv.)
- Dr. Serra ÇELİK (Istanbul University)

- Dr. Sevinç GÜLSEÇEN (Istanbul University)
- Dr. Sushil K. SHARMA (Ball State University, USA)
- Dr. Süleyman ERSÖZ (Kırıkkale University)
- Dr. Şeref SAĞIROĞLU (Gazi University)
- Dr. Tijen ÖVER ÖZÇELİK (Sakarya University)
- Dr. Tufan ADIGÜZEL (Bahçeşehir University)
- Dr. Tuğba TUNACAN (Abant İzzet Baysal)
- Dr. Tuğrul TAŞCI (Sakarya University)
- Dr. Türkay DERELİ İskenderun Technical University
- Dr. Ufuk KULA (American University of Middle East)
- Dr. Umut ALTINIŞIK (Kocaeli University)
- Dr. Ümit KOCABIÇAK (Sakarya University)
- Dr. Volkan POLAT (Yalova University)
- Dr. Yağız ÜRESİN (Istanbul University)
- Dr. Yasin PAUNSKI (Bulgarian Academy of Science)
- Dr. Yunus ÖZEN (Yalova University)
- Dr. Yücel BATU Salman (Bahçeşehir Üniv.)
- Ahmet Emrah ATEŞ (Yalova University)
- Başak ÇETİNGÜÇ (Yalova University)
- Fatih ASLAN (Yalova University)
- Hüseyin SAVRAN (Yalova University)
- Kıymet ENSARİOĞLU (Yalova University)
- Mahir ATMIŞ (Yalova University)
- Muhammed TEKİN (Yalova University)
- Özgür ÖZKAYA (Yalova University)
- Uğur YILDIZ (Kocaeli University)
- Yunus Emre AVCI (Yalova University)

CONTENTS

Entrepreneurship in Artificial Intelligence Products with Crowdfunding Tools: Comparison of Turkey and the USA	1
Investigation of the Effects of Different Learning Rules on the Success of Artificial Neural Network Forecasting Models in the Estimation of the BIST-30 Index Value by ANOVA Analysis	3
Training Simple Recurrent Deep Artificial Neural Network for Forecasting by Using Particle Swarm Optimization	6
Size Optimization of Trusses with Automatic Member Grouping: Performance Issues	8
On the Road to Development of a Hyperheuristic: Experiments on a Size Optimization Problem	11
Computer-Automated Design of RC Residential Buildings: First Thoughts and Initial Tests	13
A New Recurrent Deep Artificial Neural Network Based on Multiplicative Neuron Model for Forecasting	16
Exploring Transformers in Natural Language Generation: GPT, BERT, and XLNet	18
Deeper in Data Science: Geometric Deep Learning	21
Examination of European Union Horizon 2020 Information and Communication Technologies Projects with Social Network Analysis Method	23
Prediction of CO ₂ Emission's Time Series by Using Support Vector, Artificial Neural Networks and Classical Time Series Analysis	26
AI Applications in a Home Improvement Retailer, Koçtaş	28
An Architectural Data Study Through a Natural Language Processing Based Networked Content Tool	31
A New Approach To Potential Distribution For Grounding With Using Artificial Intelligence Methods	34
Study of Consumer Perception of Recommender Systems in E-Commerce Applications	37

Entrepreneurship in Artificial Intelligence Products with Crowdfunding Tools: Comparison of Turkey and the USA

Gökçe Öztürk
Foreign Trade Department
Istanbul Yeni Yüzyıl University
Istanbul Turkey
gokceozturk90@gmail.com

Burak Öçlü
Business Administration Department
Istanbul Kültür University
Istanbul, Turkey
burak.oclu@gmail.com

Abstract—Land, labor, enterprise, and capital are mentioned that four basic factors of production. With the emerging internet technologies, people are able to explain their ideas more easily and efficiently at a global level. Developments have led to the modification of the factors of production to the situation. When entrepreneurs realized that they could use technology to meet their capital needs, they continued their search for funds through individualization (1). In this context, crowdfunding platforms, which do not have any physical facilities, are born digitally and operate digitally, have started to be established. Capital, which is also an important problem for developers of artificial intelligence that can be used in almost every aspect of our daily lives, can be solved in this way. Aware of this, artificial intelligence developers have begun to express themselves as more members of crowdfunding sites (2). The main sources of funding for mass funding sites include individual donations. Thus, people who are involved in daily life as consumers have also gained the role of funders. This situation also provides spiritual satisfaction for people both in terms of internal satisfaction and the possibility that they may leave a lasting idea or product in the world. **Keywords**—crowdfunding, capital, entrepreneurship

I. METHOD

Our study examines crowdfunding websites that are operating in Turkey or the USA and listed with the keywords "crowdfunding", "funding", "find funds" in the Google search engine. Although crowdfunding in the USA is a deep-rooted concept, it is still a new concept in Turkey (3). Within the websites that meet the criteria, searches were carried out with the keyword "Artificial Intelligence" (AI). Among the results, the fund targets of the campaigns, the tools they use in promotion, and the products/services sold for the fund were compared. Our study aims to guide entrepreneurs who will use mass funding in the field of artificial intelligence.

The term "kitleles fonlama" is used to reach the crowdfunding websites which are operated in Turkey. For the content analysis, top two website was chosen in search results according to Alexa Page Rank. On the first page, there are 12 websites with paid and organic results. In the results, only 2 of them are actively performed as crowdfunding websites. They are fongogo.com and arikovani.com. In-page search, we used the term "yapay zeka" keyword to represent AI results. Arikovani.com has five projects, and fongogo.com has three projects according to AI.

To identify crowdfunding websites based in the USA, the proxy has changed for USA IP location. On the first page, there are 12 websites with paid and organic results. In the results, 5 of them are actively performed as crowdfunding websites. The top 2 websites are indiegogo.com and kickstarter.com. In-page search, "artificial intelligence" was used as a keyword. Indiegogo.com has seven results. Kickstarter.com has 256 results.

II. RESULTS AND DISCUSSION

The webpage has various visual elements. The publishers can able to use pictures, graphics, gif elements, videos to present their idea. Crowdfunding websites are one of the platforms to use rich media elements. Entrepreneurs who register crowdsourcing platforms wish to find funds in a limited time period. For that reason, they also have limited time to represent themselves. They try to tell their ideas to potential supporters. Both Turkey and the USA originated crowdsourcing platforms that have the feature for rich media options. On the other hand, artificial intelligence is an emerging concept along with entrepreneurs. USA Crowdsourcing platforms are frequently supporting artificial intelligence projects. Turkey crowdfunding websites have a limited number of results according to artificial intelligence. One of the important elements of the crowdfunding system is assurance. The lack of legal regulation regarding the crowdfunding concept in Turkey makes it difficult to establish assurance. Before a product or service goes into production, it is displayed on crowdfunding sites. Entrepreneurs can create prototypes. However, Turkish visitors are not willing to support the early stage of the project. Therefore, awareness should be created in terms of consumer (4). There are three main actors in crowdfunding websites. There are entrepreneurs, visitors, and website platforms. Each one wants to maximize their profit. Entrepreneurs want to get more support. Visitors want to increase their benefit of the product. Website platforms want to reach more visitors and entrepreneurs. Crowdfunding websites should reach more audiences to increase AI-related projects. To create awareness among AI developers, crowdfunding websites can enlarge their promotion activities to local developer communities. To integrate university students into the system, crowdfunding websites can create special activities for young developers. Kickstarter.com has a guide for young developers, which is called "Creator

Handbook." The handbook enables young entrepreneurs to create well-developed project pages.

In future studies, the awareness of the crowdfunding concept can be examined on the basis of various countries. Thus, the activities of crowdfunding sites established regardless of location in various countries can be differentiated. Each culture's unique entrepreneurship understanding can be addressed. Thus, more inclusive strategies can be put forward for artificial intelligence developers.

REFERENCES

- [1] Pierrakis Y, Collins L. Crowdfunding: A new innovative model of providing funding to projects and businesses. Available at SSRN 2395226. 2013 May 5.
- [2] Gierczak MM, Bretschneider U, Haas P, Blohm I, JM L. Crowdfunding-The New Era of Fundraising. Crowdfunding in Europe-State of The Art in Theory And Practice. Springer Science+ Business Media, Wiesbaden. 2015.
- [3] Funk AS. Institutions Influencing the evolution of crowdfunding in China: A Review of the World Bank Report on Crowdfunding's Potential for the Developing World. Crowdfunding in Europe. 2016:125-31.
- [4] GABISON G. Understanding crowdfunding and its regulations. 2019.

Investigation of the Effects of Different Learning Rules on the Success of Artificial Neural Network Forecasting Models in the Estimation of the BIST-30 Index Value by ANOVA Analysis

Muhammed Ordu
Department of Industrial Engineering
Osmaniye Korkut Ata University
Osmaniye, Turkey
muhammedordu@osmaniye.edu.tr

Abstract—Stock markets are an investment tool that shows instant changes in share values and keeps high returns and risks together. Therefore, it has been the subject of many forecasting studies. In this study, the effects of learning rules on the forecasting accuracy values of artificial neural network models developed to predict the BIST-30 index values are examined. As a result, the learning rules clearly cause to produce the significant differences on the model success. Through the Post Hoc Test conducted, the learning rule (i.e., Step) was determined as more successful than the others.

Keywords—ANOVA analysis, artificial neural network, BIST-30 index, forecasting, learning rule

I. INTRODUCTION

Individuals can make their investments by directing the savings to a number of investment tools from the past to the present. It is the aim to secure the future by saving or to reach a planned amount. For this reason, various investment tools have existed throughout the history. The followings are given as an example: Gold, deposits, foreign exchange, repo, bonds and stocks. Among these, stocks are the type of investment tools which the risk and return are very high. In addition, stocks have always attracted the attention of investors. In the circumstances, it has always been a matter of curiosity to predict stock market prices and what the future price of stocks will be. The desire to gain more profit by keeping the return high and the risk low has forced the researchers to conduct scientific research on stock markets.

Artificial neural networks (ANN) are popularly used in forecasting the stock market prices. Learning rules in the methodology of artificial neural networks are one of the most important factors affecting the success of the models. In this study, it is aimed to investigate the effects of learning rules on the success of ANN models.

The remaining of the paper is organized as follows: Section 2 summarizes the literature of ANN application on stock markets, Section 3 explains the material and methods. Section 4 discusses the results and Section 5 concludes the study, respectively.

II. LITERATURE REVIEW

In the literature, stock markets have been the focus of many research areas for few decades, especially in the field of forecasting. Stock market index values were estimated by using both univariate time series and multivariate time series. For example, Guresen, Kayakutlu and Daim [1] Laboissiere, Fernandes and Lage [2], Qui, Song and Akagi [3] and Moghaddam, Moghaddam and Esfandyari [4].

Some forecasting models are compared with ANN method to produce better estimates of stock prices. For example, ANN and Adaptive Exponential Smoothing Method by Faria, Albuquerque, Gonzalez, Cavalcante and Albuquerque [5], ANN and Support Vector Machine (SVM) by Kara, Boyacioglu and Baykan [6] and, Regression, ANN and SVM by Sheta, Ahmed and Faris [7].

These forecasting methods have been integrated with different methods to generate more reliable and accurate estimates. For example, genetic algorithms and ANN integrated by Kim and Han [8] on the other hand, Hadavandi, Shavandi and Ghanbari [9] propose an approach combining the ANN and genetic fuzzy systems.

A number of researchers have reviewed the literature regarding to the artificial neural network forecasting models developed within a certain period. For example, Dase and Pawar [10] investigate the literature for the application of ANN to predict the stock markets for the period between 1990 and 2010. Li and Ma [11] and Vui, Soon, On, Alfred and Anthony [12] research how studies applied the ANN to financial stock markets.

In this study, it is aimed to fill the gap of the literature in this area by investigating the effects of learning rules on the forecast accuracy of forecasting models developed with ANN.

III. MATERIAL AND METHOD

This study proposes a two-stage approach: In the first stage, an artificial neural network model is developed to estimate the BIST-30 Index values. In the second stage, an

ANOVA analysis is conducted for the effects of different learning rules on the success of the forecasting model in stock market index prediction. The structure of the developed ANN model is as follows: The number of hidden layers is one and the number of hidden neurons is five. The activation function is the Hyperbolic Tangent, and the following six different learning rules are taken into account: Step, Momentum, Conjugate Gradient, Levenberg-Marquardt, Quickprop and Delta Bar Delta. Our model includes 10 inputs and 1 output. The followings are our inputs: Exchange rate of USA dollar, industrial production index value, consumer price index changes, gold price, deposit interest rate, rate of export change, rate of import change, unemployment rate, GDP value, the BIST-30 index value at the previous period. Output is the estimates of BIST-30 index value. The Adjusted R-Squared is considered as the forecast accuracy.

In the second stage, the effects of six learning rules above on the forecast accuracy measured by our ANN models are examined. After that, one-way ANOVA analysis is performed and then, whether there exist the significant differences between learning rules or not is determined by carrying out the Post Hoc Tests (i.e., Tukey).

IV. RESULTS AND DISCUSSION

In the study, six different artificial neural network models have been developed depending on the number of learning rules. Each of these models was run with 50 replications. Thus, a dataset consisting of a total of 300 data was obtained. This data set contains the Adjusted R-Squared value and learning rule for each experiment. One-way ANOVA analysis was conducted to examine the effects of learning rules on the Adjusted R-Squared values. Table I shows the results of the normality test and according to the significance values obtained as a result of the normality test, the dataset belonging to each learning rule conforms to the normal distribution.

TABLE I. TEST OF NORMALITY

Dependent Variable	Learning Rules	Kolmogorov-Smirnov		
		Statistic	df	Sig.
R-Squared	Step	0.118	50	0.079
	Momentum	0.100	50	0.200
	Conjugate Gradient	0.099	50	0.200
	Levenberg-Marquardt	0.106	50	0.200
	Quickprop	0.087	50	0.200
	Delta Bar Delta	0.093	50	0.200

Table II presents the results of ANOVA analysis. Accordingly, significance value is clearly less than 0.05 and there exist the significant differences in terms of the Adjusted R-Squared values amongst the learning rules.

TABLE II. ANOVA ANALYSIS

ANOVA Analysis	Sum of Squares	df	Mean Square	F	Sig.
Between Groups	4.372	5	0.874	283.540	0.000
Within Groups	0.907	294	0.003		
Total	5.279	299			

The Post Hoc Test is carried out to better understand which the learning rules are different from each other and for this, Tukey test is conducted. Table III gives the results of the Tukey test. According to the results, significance values are less than 0.05 and means are different from each other. Step is more successful than others, Momentum gives more accurate estimates than only Quickprop which produces the better estimates than all learning rules except Step. Conjugate Gradient have the better of Momentum and Delta Bar Delta. On the other hand, Levenberg-Marquardt is superior than only Momentum, Conjugate Gradient and Delta Bar Delta.

TABLE III: THE TUKEY TEST RESULTS

Learning Rules (1)	Learning Rules (2)	Mean Differences (1-2)	Standard Error
Step (S)	M	0,255337300*	0,011106949
	CG	0,204929320*	0,011106949
	L-M	0,112319620*	0,011106949
	Q	0,035980540*	0,011106949
	DBD	0,341933280*	0,011106949
Momentum (M)	S	-0,255337300*	0,011106949
	CG	-0,050407980*	0,011106949
	L-M	-0,143017680*	0,011106949
	Q	-0,219356760*	0,011106949
	DBD	0,086595980*	0,011106949
Conjugate Gradient (CG)	S	-0,204929320*	0,011106949
	M	0,050407980*	0,011106949
	L-M	-0,092609700*	0,011106949
	Q	-0,168948780*	0,011106949
	DBD	0,137003960*	0,011106949
Levenberg-Marquardt (L-M)	S	-0,112319620*	0,011106949
	M	0,143017680*	0,011106949
	CG	0,092609700*	0,011106949
	Q	-0,076339080*	0,011106949
	DBD	0,229613660*	0,011106949
Quickprop (Q)	S	-0,035980540*	0,011106949
	M	0,219356760*	0,011106949
	CG	0,168948780*	0,011106949
	L-M	0,076339080*	0,011106949
	DBD	0,305952740*	0,011106949

V. CONCLUSION

This study has addressed the importance of variations of the different parameters in the artificial neural network forecasting models. Learning rules is one the parameters in ANN modelling and closely related to the more robust and accurate estimates. In this study, six different learning rules are investigated in terms of their effects on forecast accuracy (i.e., Adjusted R-Squared). All of them excluding Step and Delta Bar Delta is able to dominate each other in the study. In this study, Step is the best learning rule whereas Delta Bar Delta is the worst. This study reveals that different parameters can be successful in producing better estimates.

REFERENCES

- [1] E. Guresen, G. Kayakutlu, and T. U. Daim, "Using artificial neural network models in stock market index prediction," *Expert Systems with Applications*, vol. 38, pp. 10389-10397, February, 2011.
- [2] L. A. Laboissiere, R. A. S. Fernandes, and G. G. Lage, "Maximum and minimum stock price forecasting of Brazilian power distribution companies based on artificial neural networks," *Applied Soft Computing*, vol. 35, pp. 66-74, October, 2015.
- [3] M. Qiu, Y. Song, and F. Akagi, "Application of artificial neural network for the prediction of stock market returns: The case of the Japanese stock market," *Chaos, Solitons and Fractals*, vol. 85, pp. 1-7, April, 2016.

- [4] A. H. Moghaddam, M. H. Moghaddam, and M. Esfandiyari, "Stock market index prediction using artificial neural network," *Journal of Economics, Finance and Administrative Science*, vol. 21, pp. 89-93, December, 2016.
- [5] E. L. Faria, M. P. Albuquerque, J. L. Gonzalez, J. T. P. Cavalcante, and M.P. Albuquerque, "Predicting the Brazilian stock market through neural networks and adaptive exponential smoothing methods", vol. 36, pp.12506-12509, December, 2009.
- [6] Y. Kara, M. A. Boyacioglu, and Ö. K. Baykan, "Predicting direction of stock price index movement using artificial neural networks and support vector machines: The sample of the Istanbul Stock Exchange," *Expert Systems with Applications*, vol. 38, pp. 5311-5319, May, 2011.
- [7] A. F. Sheta, S. E. M. Ahmed, and H. Faris, "A Comparison between Regression, Artificial Neural Networks and Support Vector Machines for Predicting Stock Market Index," *International Journal of Advanced Research in Artificial Intelligence*, vol. 4, pp. 55-63, July, 2015.
- [8] K. Kim, and I. Han, "Genetic algorithms approach to feature discretization in artificial neural networks for the prediction of stock price index," *Expert Systems with Applications*, vol. 19, pp. 125-132, August, 2000.
- [9] E. Hadavandi, H. Shavandi, and A. Ghanbari, "Integration of genetic fuzzy systems and artificial neural networks for stock price forecasting," *Knowledge-Based Systems*, vol. 23, pp. 800-808, December, 2010.
- [10] R. k. Dase, and D. D. Pawar, "Application of Artificial Neural Network for stock market predictions: A review of literature," *International Journal of Machine Intelligence*, vol. 2, pp. 14-17, December, 2010.
- [11] Y. Li, and W. Ma, "Applications of Artificial Neural Networks in Financial Economics: A Survey," *2010 International Symposium on Computational Intelligence and Design, China*, pp. 211-214, October, 2010.
- [12] C. S., Vui, G. K. Soon, C. K. On, R. Alfred, and P. Anthony, "A Review of Stock Market Prediction with Artificial Neural Network (ANN)," *2013 IEEE International Conference on Control System, Computing and Engineering, Malaysia*, pp. 477-482, December, 2013.

Training Simple Recurrent Deep Artificial Neural Network for Forecasting by Using Particle Swarm Optimization

Eren Bas
Department of Statistics
Giresun University
Giresun, Turkey
eren.bas@giresun.edu.tr

Erol Eğrioğlu
Department of Statistics
Giresun University
Giresun, Turkey
erol.egrioglu@giresun.edu.tr

Emine Kolemen
Department of Statistics
Giresun University
Giresun, Turkey
emine.kolemen@giresun.edu.tr

Abstract— Deep artificial neural networks have been popular for time series forecasting literature, recent years. The recurrent neural networks present more suitable architectures for forecasting problem than other deep neural network types. The simplest deep recurrent neural network type is simple recurrent neural networks according to the number of employed parameters. These neural networks can be preferred to solve the forecasting problem because of their simple structure if they are well trained. Unfortunately, the training of simple recurrent neural networks is problematic because of exploding or vanishing gradient problem. In this study, a new training algorithm is proposed based on particle swarm optimization. The algorithm does not use gradients so it has not vanished or exploding gradient problem. The performance of the new training algorithm is compared with long short term memory which is trained by adam algorithm by using some stock exchange time series. At the end of the analysis processes, the proposed method produces more accurate forecast results than long short term memory.

Keywords—deep learning, recurrent neural networks, forecasting, particle swarm optimization

I. INTRODUCTION

Early forecasting methods were based on probability theory and they were generally statistical methods. Recent years, machine learning methods and their hybridization with statistical methods have become popular, day by day. Machine learning methods do not need any probabilistic or statistical assumptions. Machine learning methods use nonlinear structures and soft models. Artificial neural networks are an important class of machine learning methods. Artificial neural networks can be classified into two groups shallow and deep artificial neural networks. Deep artificial neural networks use more parameters than shallow artificial neural networks. Deep artificial neural networks use parameter sharing approach so they process more data with less parameter than shallow neural networks. Deep artificial neural networks produced successful forecast results in the forecast competitions. Especially, the methods based on long short term recurrent neural networks have top ranks in the competitions. Long short term memory artificial neural network proposed by Hochreiter et al. (1997) to solve vanishing and exploding gradient problem of simple recurrent artificial neural networks. Although the problems of simple recurrent

artificial neural networks, the number of parameters in long short term memory dramatically increase because of various gates. Instead of using various gates in the recurrent neural networks, gradient-free algorithms can be preferred to train simple recurrent neural networks. So forecasting problems can be solved by using fewer number parameters than long short term memory.

The recurrent neural networks employed for forecasting problem in some researches in the literature. Moalla et al. (2017) and Yao et al. (2018) proposed a particle swarm optimization based long short term memory forecasting method. Chung and Shin (2018) proposed long short term memory based on a genetic algorithm approach to forecasting the Korean Stock Price Index. Chen et al. (2018), Stajkowski et al. (2020) used long short term memory with a genetic algorithm to predict network traffic. Shao et al. (2019) proposed a nickel-metal price prediction model based on a particle swarm optimization algorithm with long-short-term memory. Lu et al. (2020) proposed a hybrid method based on the convolutional neural network and long short term memory for forecasting of stock price data. Gundu and Simon (2020) proposed the long-term memory neural network model based on an advanced particle swarm optimization for forecasting the closing price of the Indian Energy Exchange.

In this study, a new gradient-free algorithm based on a modified particle swarm optimization method is proposed for the training of the simple deep recurrent neural network to forecast single-variable time series. The proposed method method uses a fewer number of parameters than long short term memory. Moreover, the proposed method does not have vanishing or exploding gradient problem because particle swarm optimization does not need gradients of the objective function. In the second section of the paper, the proposed method is introduced. In the third section, application results are given. The last section is about conclusions and discussions.

II. A NEW TRAINING ALGORITHM

The new training algorithm is given step by step as follows:

Step 1. The parameters of the particle swarm optimization algorithm are determined.

Step 2. The initial positions and velocities are randomly generated. The positions of the particle swarm optimization are weights and biases of simple deep recurrent neural networks. The outputs of simple deep recurrent neural network with one hidden layer are calculated with the following formulas.

$$h_t = f(x_t S + h_{t-1} W + b_1) \quad (1)$$

$$\hat{x}_t = f(h_t V + b_y) \quad (2)$$

The total number of weights and biases is $(p + h + 2)h + 1$ because the dimensions of weights and biases are $S: p \times h$, $W: h \times h$, $b_1: 1 \times h$, $V: h \times 1$ and $b_y: 1 \times 1$.

Step 3. For each particle, the fitness function values are calculated as mean square error for training data.

Step 4. Pbest (the memory for each particle) and gbest (the memory for the swarm) are constitute Step 5. The new velocities and positions are calculated.

Step 6. Pbest and gbest are updated.

Step 7. The early stopping rule is checked. If the rule is not satisfied go to Step 5.

Step 8. The algorithm is ended.

III. APPLICATION

In the application section, Dow Jones data set is used. The data set was downloaded from Yahoo Finance Website (<https://finance.yahoo.com>). The series is called "Dow Jones Industrial Average (DJI), DJI-DJI Real-Time Price. Currency in USD". Data set is constituted from 5 daily opening prices for years 2014-2018. Five time series are solved by using Long Short Term Memory (LSTM) deep neural network, simple recurrent neural network (SRNN) artificial neural network method. The number of inputs is changed from 1 to 5 by 1 increment in all artificial neural networks applications. In application of LSTM and SRNN number of hidden layers is changed from 1 to 5 by 1 increment. Each methods is applied 30 times by using random initial weights. In the application, the time series is divided into three parts as training, validation and testing data. Training data was used to train artificial neural network and validation data were used to select best configuration or parameter tuning in the architecture of the artificial neural network. The test set was used to compare performance of the different artificial neural network methods. The best conclusion are given in Table 1 for methods.

The statistics for mean and standard deviation values of root of mean square errors (RMSE) values calculated for the test set are given in Table 1. The SRNN is the best method according to mean, standard deviation for 2017 and 2018 years. According to all results, it can be said that SRNN method can produce more robust and accurate results than the LSTM.

TABLE I. STATISTICS FOR RMSE VALUES FOR THE TEST DATA

Year	Methods	Mean	Standard Deviation
2017	LSTM	107.3676	2.5088
	SRNN-PSO	104.1725	0.7734
2018	LSTM	454.4992	9.4840
	SRNN-PSO	438.3752	0.9052

REFERENCES

- [1] Hochreiter S, Silva PCL, Schmidhuber J, "Long short-term memory," Neural Computations vol. 9, pp.1735-1780, 1997.
- [2] Moalla, H, Elloumi, W, Alimi, A. M, "Hybrid LSTM trained by PSO for online handwriter identification," In International Conference on Neural Information Processing pp. 41-50, November 2017.
- [3] Yao, Y, Han, L, Wang, J, "Long short-term memory ship motion prediction based on particle swarm optimization," In 2018 IEEE CSAA Guidance, Navigation and Control Conference (CGNCC) pp. 1-5 IEEE, August 2018.
- [4] Chung, H, Shin, K. S, "Genetic algorithm-optimized long short-term memory network for stock market prediction," Sustainability, vol. 10, pp. 3765, October 2018.
- [5] Chen, J, Xing, H, Yang, H, Xu, L, "Network Traffic Prediction Based on LSTM Networks with Genetic Algorithm," In International Conference On Signal And Information Processing, Networking And Computers pp. 411-419, Springer, Singapore, November 2018.
- [6] Shao, B, Li, M, Zhao, Y, Bian, G, "Nickel Price forecast based on the LSTM neural network optimized by the improved PSO Algorithm," Mathematical Problems in Engineering, September 2019.
- [7] Stajkowski, S, Kumar, D, Samui, P, Bonakdari, H, Gharabaghi, B, "Genetic-algorithm-optimized sequential model for water temperature prediction," Sustainability, vol.12, pp. 5374, July 2020.
- [8] Lu, W, Rui, H, Liang, C, Jiang, L, Zhao, S, Li, K, "A method based on GA-CNN-LSTM for daily tourist flow prediction at Scenic Spots," Entropy, vol.22, pp.261, February 2020.
- [9] Gundu, V, Simon, S. P, "PSO-LSTM for short term forecast of heterogeneous time series electricity price signals," Journal of Ambient Intelligence and Humanized Computing, pp. 1-11, 2020.

Size Optimization of Trusses with Automatic Member Grouping: Performance Issues

Yusuf Cengiz Toklu (*Prof.*)
Department of Civil Engineering
Beykent University
İstanbul, Turkey
cengiztoklu@gmail.com

Hakan Özbaşaran (*Assoc. Prof.*)
Department of Civil Engineering
Eskişehir Osmangazi University
Eskişehir, Turkey
ozbasaran@ogu.edu.tr

EXTENDED ABSTRACT

The size optimization process searches for the member sections of the structure that lead to best feasible design. The size optimization is preferred to simultaneous size, shape and topology optimization mostly to protect the architectural form. It is well known by the practitioners that the number of the section types used in a structural design significantly affects the applicability; the construction time of a design increases by the section type count since there are many factors involved from supply to assembly mistakes. Therefore, the structural members are grouped and same sections are used for the members in a group. Considering this additional constraint (member groups), the size optimization procedure becomes a more difficult problem. It is not easy (even with structural analysis guidance) to give an answer to the questions: “Which member belongs to which group?” and “What is the least number of member groups to obtain a reasonable near-optimal design?” There are two common practices to determine the member-grouping configuration. The former is linking the members to groups using the engineering insights before the optimization attempt. Of course, the researchers in the field know that pre-grouping the members is not “overcoming” the problem but “evading”. Satisfactory results may be obtained for simple structures with this method; however assessment of the member-grouping information gets harder with the structure complexity and member count. The latter is building the design vector so as to include the member grouping information in addition to section sizes. This option surely is the only exact way to determine the optimal member-grouping configuration for now; however it requires far more computational effort compared to the size optimization of structures with pre-grouped members. The authors think that seeking new design vector formats and design vector interpretation schemes is a good point to start to improve the performance of size optimization procedures with automatic member grouping. There are a vast number of studies on size optimization of trusses that involve testing the performance of a specific metaheuristic (maybe an improved version) through the popular pre-grouped benchmark problems. On the other hand, there are only a limited number of works on automatic member grouping. The de facto encoding [1] for size optimization

of trusses with automatic member grouping generates a two part design vector. The first part is for the section

types used in the structure. Each value of this part contains a cross-section size, which may be the cross-section area for optimization with continuous variables (Fig. 1a) or the section label for optimization with discrete variables (Fig. 1b). The second part pairs each member with a section from the first part. Thus, the maximum number of allowed sections constraint is automatically satisfied without any other additional operation. This study evaluates the performance of the “plain” decoding for the de facto encoding for size optimization of trusses with automatic member grouping.

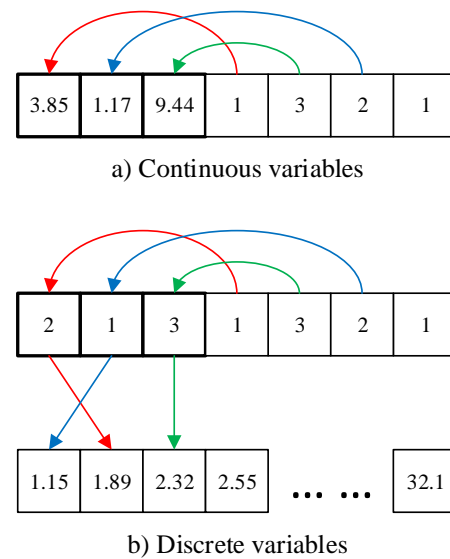


Fig. 1. Plain decoding

The “plain” decoding directly decodes the introduced design vector as given in Fig. 1. For example, the member cross-section areas of the candidate given in Fig. 1a are set to 3.85, 9.44, 1.17, and 3.85 square units, respectively; on the other hand, the member cross-section areas of the structure represented by the design vector given in Fig. 1b are set to 1.89, 2.32, 1.15, and 1.89 square units, respectively. In this study, the Jaya Algorithm (JA) introduced by Rao in 2016 [2] is selected as the optimization algorithm. The JA is based on the concept that a candidate solution may be improved by carrying it towards the best solution and away from the worst solution. The detailed explanation and a numerical example can be found in the original manuscript. The

performance of the “plain” decoding is evaluated through the 10-bar planar cantilever (Fig. 2), which is one of the popular truss-sizing optimization benchmark problems. Since the original problem has no element groups, the cross-section area of each member is a design variable. This problem has two versions as continuous and discrete. In the continuous version, the design variables can be selected between 0.1 in^2 and 35.0 in^2 , respectively. On the other hand, the cross-section areas should be selected from the set: {1.62, 1.80, 1.99, 2.13, 2.38, 2.62, 2.63, 2.88, 2.93, 3.09, 3.13, 3.38, 3.47, 3.55, 3.63, 3.84, 3.87, 3.88, 4.18, 4.22, 4.49, 4.59, 4.80, 4.97, 5.12, 5.74, 7.22, 7.97, 11.50, 13.50, 13.90, 14.20, 15.50, 16.00, 16.90, 18.80, 19.90, 22.00, 22.90, 26.50, 30.00, 33.50} in^2 in the discrete version. This structure is optimized under element stress and node displacement constraints. The allowed stress for the members is $\pm 25 \text{ ksi}$. The allowed nodal displacements are ± 2.0 inches in both vertical and horizontal directions. Finally, the structural material has a unit weight of 0.1 lb/in^3 and a Young’s modulus of 10^4 ksi .

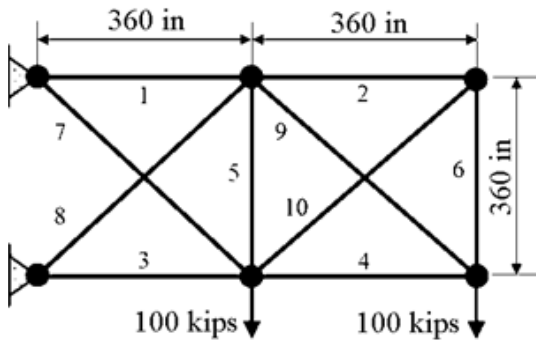


Fig. 2. 10-bar planar cantilever

The population size and the maximum allowed number of candidate evaluations is set to 50 and 50000, respectively in all of the optimization attempts made in this study. The results of 20 consecutive runs are used to visualize the performance of the “plain” decoding. The Figs. 3 and 4 show the weights of the best, average, and worst designs obtained for the continuous and discrete versions of the 10-bar cantilever problem, respectively. The horizontal axes of the mentioned plots (N_G) show the maximum allowed number of groups from 1 to 10. The vertical axes shows the weight of the design in pounds.

A better design is obtained by allowing 4 groups compared to the design obtained by allowing 10 groups in the continuous version of the problem (see the “Best” series of Fig. 3). This case is a result of the insufficient number of candidate evaluations allowed for $N_G=10$. Please note that the automatic member grouping is not disabled in optimization with 10 groups; and therefore, there are a total of 20 design variables as 10 member groups and 10 section sizes. The known near-optimal solution of the continuous version of this problem lays somewhere around 5061 lbs and satisfactory results can be obtained by allowing about 15000 candidate evaluations [3]. In this study, a 5170.31 lbs design is discovered by allowing 10 member groups after 50000 candidate evaluations. On the other hand the best solution found so far for the discrete version of the 10-bar cantilever problem is a 5490.74 lbs design, which can be spotted by allowing

about 10000 candidate evaluations [4]. In this study a slightly heavier design is obtained (5491.72 lbs) by allowing 50000 candidate evaluations. In addition, it can be seen from the Figs. 3 and 4 that the difference between the best and the average results are unacceptably large. It is previously mentioned in the study that the automatic member grouping feature makes the optimization process far more difficult; therefore, the increase in the required number of candidate evaluations is expected by the authors. However, the automatic member grouping procedure is inefficient in its current form considering the obtained results and computational effort.

Keywords—truss, optimization, member grouping, design, metaheuristics

N_G	Best	Worst	Average
1	8266.15	8266.15	8266.15
2	5943.84	6724.26	6216.36
3	5301.53	6620.65	5656.09
4	5163.41	6328.40	5609.17
5	5103.99	6214.25	5321.80
6	5095.35	5944.26	5393.13
7	5162.71	6526.13	5418.87
8	5114.97	5786.67	5416.42
9	5164.23	6246.98	5530.56
10	5170.31	6583.34	5515.89

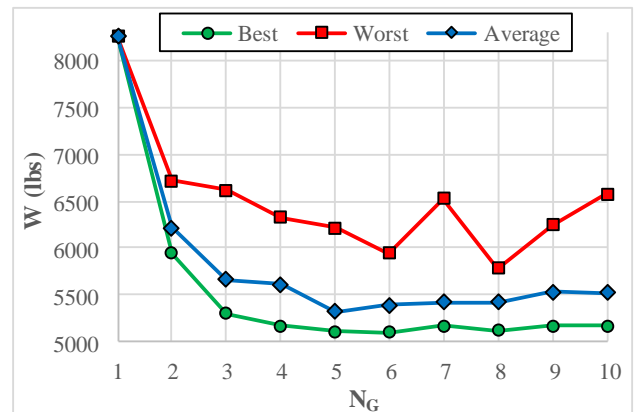


Fig. 3. Results of the continuous version of the 10-bar cantilever problem

N_G	Best	Worst	Average
1	8350.97	8350.97	8350.97
2	6152.52	7096.58	6430.39
3	5778.57	6995.12	5923.76
4	5567.32	6075.53	5735.96
5	5498.37	6382.41	5769.54
6	5491.72	6409.14	5782.30
7	5565.89	6658.95	5830.66
8	5524.24	6936.21	5966.83
9	5498.37	6872.44	5859.30
10	5491.72	6758.90	5878.10

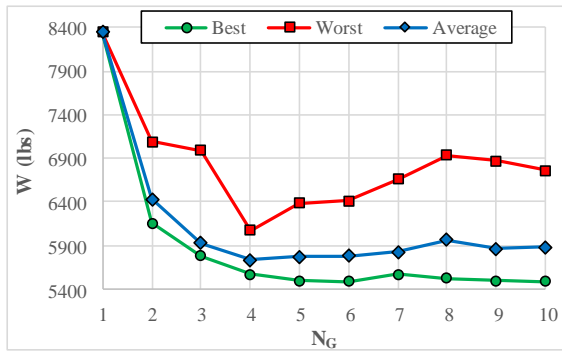


Fig. 4. Results of the discrete version of the 10-bar cantilever problem

REFERENCES

- [1] H. J. C. Barbosa, A. C. C. Lemonge, C. C. H. Borges, "A genetic algorithm encoding for cardinality constraints and automatic variable linking in structural optimization," *Eng. Struct.*, vol. 30, pp. 3708–3723, 2008.
- [2] R. Venkata Rao, "Jaya: A simple and new optimization algorithm for solving constrained and unconstrained optimization problems," *Int. J. Ind. Eng. Comput.*, vol. 7, pp. 19–34, 2016.
- [3] M. Khatibinia, H. Yazdani, "Accelerated multi-gravitational search algorithm for size optimization of truss structures," *Swarm Evol. Comput.*, vol. 38, pp. 109–119, 2018.
- [4] M-Y. Cheng, D. Prayogo, Y-W. Wu, M. M. Lukito, "A Hybrid Harmony Search algorithm for discrete sizing optimization of truss structure," *Autom. Constr.*, vol. 69, pp. 21–33, 2016.

On the Road to Development of a Hyperheuristic: Experiments on a Size Optimization Problem

Halil İbrahim Aytekin (*MSc candidate*)
Department of Civil Engineering
Eskişehir Osmangazi University
Eskişehir, Turkey
halilibrahimaytekin@gmail.com

Hakan Özbaşaran (*Assoc. Prof.*)
Department of Civil Engineering
Eskişehir Osmangazi University
Eskişehir, Turkey
ozbasaran@ogu.edu.tr

EXTENDED ABSTRACT

Recently a significant number of metaheuristic algorithms having various inspiration sources such as physics rules, swarm intelligence, and survival behavior are presented. Most of them have algorithm-specific control parameters other than the population size and maximum allowed number of candidate evaluations. The parameter-controlled algorithms should be specifically tuned for the considered optimization problem, which results in a significant waste of computational effort. On the other hand, there are a number of parameter-free metaheuristics; the performance of these algorithms vary from one optimization problem to another. Unfortunately, there is no one-size-fits-all optimization algorithm. In an engineering point of view, it is mostly more efficient to use a parameter-free algorithm and spend the computational effort required by tuning process of a parameter-controlled algorithm on obtaining better solutions with a parameter-free metaheuristic by allowing more candidate evaluations. However, there may be a better option.

This study presents a hyperheuristic algorithm that heuristically engages various parameter-free algorithms to discover reasonable solutions for optimization problems. Five widely used parameter-free algorithms are selected to be used by the presented hyperheuristic after an extensive literature review as the Jaya Algorithm (JA) [1], Colliding Bodies Optimization (CBO) [2], Grey Wolf Optimizer (GWO) [3], Symbiotic Organisms Search (SOS) [4], and Teaching-Learning-Based Optimization (TLBO) [5].

The JA is based on the idea that better solutions can be discovered by moving towards the best solution and away from the worst solution in a population of candidates. Its single line movement operator makes the algorithm easy-to-implement. Despite of its simplicity, the JA is a rather successful algorithm. The CBO is inspired by collision between objects. According to the CBO analogy, each candidate solution in the population is a body. The better half of the solutions are stationary and the worse half of the solutions are moving bodies that collide to the stationary ones and move them to the new positions. The GWO imitates the hunting behavior of the grey wolves. The paper by Mirjalili first explains the hierarchy in a wolf pack (alpha, beta, omega, and delta rankings), and then the mathematical simulation of the hunting behavior is presented as the movement operator of the presented

algorithm using the alpha, beta, and delta ranked solutions in a population. The SOS simulates the symbiotic interaction strategies of the organisms to survive. The phases of the SOS are mutualism, commensalism, and parasitism. The mutualism phase benefits both considered solutions. In the commensalism phase, one of the solutions benefits from the other one, which is not affected (neither gets better nor worse) by this interaction. Finally, the parasitism phase tries to replace a solution with a mutant one. The TLBO is based on the influence of the teacher on the learners. The TLBO consists of two phases as teacher and learner. In the teacher phase, the solutions in the iteration are improved with the aid of the best solution (teacher). In the learner phase, the solutions try to improve each other with the help of their mutual interaction. Please refer to the original manuscripts for further information. The mentioned algorithms are introduced briefly due to the word count limit.

At the beginning of the study, the authors included all five algorithms and conducted numerical tests. However, the final version of the hyperheuristic uses three of them: JA, SOS, and TLBO. The main reason of this change is that the movement operators of the removed algorithms (CBO and GWO) use the maximum number of allowed iterations. Instead of limiting the number of iterations to a certain number, it is better to use flexible stopping criteria such as “*stop if the improvement is smaller than 0.1% in the last 100 iterations*”. Note that the required number of candidate evaluations (and therefore the number of iterations) for convergence cannot be foreseen.

The proposed hyperheuristic algorithm first starts with a randomly selected algorithm among JA, SOS, and TLBO. If the selected algorithm “improves” the population, the same algorithm performs the second iteration. It is reasonable to say that if an algorithm provides improvement in an iteration, that algorithm is likely to improve the population in the next iteration. On the other hand, if the selected algorithm fails to improve, another algorithm among the other two is selected for the next iteration. At this point the word “improvement” should be defined. This study assumes that a population is improved if the best individual of the current iteration is better than that of the previous iteration. Of course, other interpretations of “improvement” that refer to the average population quality or the worst individuals may be tested. The characteristics of the optimization problems change

while the population drives on the road to convergence. This strategy tries to engage the most successful algorithm when needed during the optimization process. Thus, an efficient procedure is expected. Note that the hyperheuristics should not be confused with hybrid metaheuristics; hybrid metaheuristics are (mostly) single algorithms that try to combine the benefits of more than one metaheuristic.

In this short paper, the proposed hyperheuristic is tested on a structural optimization problem. The 72-bar 4-storey problem is a popular size optimization problem in the field of structural optimization. This three dimensional structure is optimized under stress and displacement constraints. The allowed stress for the members is ± 25 ksi and the allowed deflection for the nodes in all three directions is ± 0.25 in. The unit weight of the structural material is 0.1 lb/in^3 , and finally, the Young's modulus is 10000 ksi (Fig. 1).

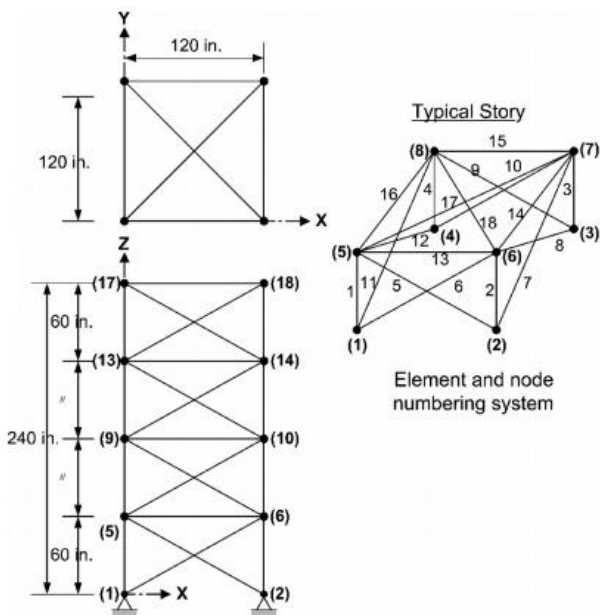


Fig. 1. 72-bar 4-storey tower

The mentioned stress and displacement constraints should be satisfied for both of the loading cases given in Table I.

TABLE I. LOADING CASES (KIPS)

Loading Cases	Node ID	P_x	P_y	P_z
Case I	17	0	0	-5
	18	0	0	-5
	19	0	0	-5
	20	0	0	-5
Case II	17	5	5	-5

This problem has several versions with pre-grouped members and continuous or discrete variables. In this study, the members of the structure are not linked to groups; and the cross section areas should be selected from the list: {0.111, 0.141, 0.196, 0.250, 0.307, 0.391, 0.442,

0.563, 0.602, 0.766, 0.785, 0.994, 1.000, 1.228, 1.266, 1.457, 1.563, 1.62, 1.80, 1.99, 2.13, 2.38, 2.62, 2.63, 2.88, 2.93, 3.09, 3.13, 3.38, 3.47, 3.55, 3.63, 3.84, 3.87, 3.88, 4.18, 4.22, 4.49, 4.59, 4.80, 4.97, 5.12, 5.74, 7.22, 7.97, 8.53, 9.30, 10.85, 11.5, 13.5, 13.9, 14.2, 15.5, 16.0, 16.9, 18.8, 19.9, 22.0, 22.9, 24.5, 26.5, 28.0, 30.0, 33.5} in^2 . Therefore, there are 72 discrete design variables. The population size is set to 50 and maximum number of candidate evaluations is set to 50000. Table II presents the best, average, and worst results of 30 consecutive runs.

TABLE II. OPTIMIZATION RESULTS (LBS)

	PS	JA	SOS	TLBO
Best	324.94	338.59	363.67	442.60
Avg.	389.76	370.40	417.47	606.46
Worst	501.64	433.81	490.01	806.27
St. D.	45.29	22.21	28.56	107.26

In Table II, the PS (presented study) column is for the results obtained by the proposed hyperheuristic, which discovered lightest design (324.94 lbs) by far. This consequence can be interpreted as the proposed algorithm properly switched between the JA, SOS, and TLBO. The average score is better than the SOS and TLBO but worse than the JA by approximately 5%. In addition, the worst design is approximately 15% and 2% worse than those obtained by the JA and SOS, respectively. The authors think that these results are promising, however the proposed algorithm should be further tested through the optimization problems of various fields.

Keywords—hyperheuristic, metaheuristic, optimization, truss

ACKNOWLEDGMENT

This short paper is based on the thesis submitted by Halil İbrahim Aytekin (supervised by Hakan Özbaşaran) in partial fulfillment of the requirements of the Master of Science degree at Eskişehir Osmangazi University.

REFERENCES

- [1] R. Venkata Rao, "A simple and new optimization algorithm for solving constrained and unconstrained optimization problems," Int. J. Ind. Eng. Comput., vol. 7, pp. 19–34, 2016
- [2] A. Kavet, V. R. Mahdavi, "Colliding bodies optimization: A novel meta-heuristic method," Comput. Struct., vol. 139, pp. 18–27, 2014
- [3] S. Mirjalili, S. M. Mirjalili, A. Lewis, "Grey Wolf Optimizer," Adv. Eng. Softw., vol. 69, pp. 46–61, 2014.
- [4] M-Y. Cheng, D. Prayogo, "Symbiotic Organisms Search: A new metaheuristic optimization algorithm," Comput. Struct., vol. 139, pp. 98–112, 2014.
- [5] R. V. Rao, V. J. Savsani, D. P. Vakharia, "Teaching-learning-based optimization: A novel method for constrained mechanical design optimization problems," Comput. Des., vol. 43, pp. 303–315, 2011.

Computer-Automated Design of RC Residential Buildings: First Thoughts and Initial Tests

Hakan Özbaşaran (*Assoc. Prof.*)
Department of Civil Engineering
Eskişehir Osmangazi University
Eskişehir, Turkey
ozbasaran@ogu.edu.tr

Batuhan Şanlı (*MSc student*)
Department of Civil Engineering
Eskişehir Osmangazi University
Eskişehir, Turkey
batuhansanli@gmail.com

Pınar Ataş (*MSc*)
Department of Civil Engineering
Eskişehir Osmangazi University
Eskişehir, Turkey
pinaratas@outlook.com

Yunus Eren Kaya (*BSc student*)
Department of Civil Engineering
Eskişehir Osmangazi University
Eskişehir, Turkey
yunuserenkaya@gmail.com

EXTENDED ABSTRACT

Expert systems are computer programs that imitate an expert who has knowledge and experience in a certain field. Expert systems have been developed and successfully applied in various fields in time. DENDRAL [1] that analyzes the chemical structure of the soil, MYCIN [2] that identifies the bacteria causes serious infections such as meningitis and recommends antibiotics, XCON [3] as a system for trade, and REPCON [4] as a digital consultant for reinforcement, repairing, and renewal of buildings are examples of expert systems. In today's engineering practice, final structural system plans for reinforced concrete structures are designed by trial-and-error methods using computer-aided design (CAD) software. First, the engineer develops a preliminary structural system considering the architectural design and submits it to the preferred structural analysis & design software. Second, the engineer makes some changes on the preliminary design based on his engineering insights if the preferred software returns that there are structural system elements that cannot be designed due to various reasons such as excessive internal forces or design code violations. This second stage is repeated until a structural system plan is (somehow) discovered that allows the designing of all structural members. Lack of time is mostly the main problem in the design process of a RC residential structure. Therefore, the engineer usually has no extra time to try alternative structural system plans in most of the design procedures; the first discovered design that satisfies all of the considered constraints is prepared as the application project. Apparently, it is hard to say that the designs discovered by the procedure described above is near-optimal. On the other hand, there is no reason that discovery of near-optimal designs by computer-automated procedures that benefits from artificial intelligence methods is impossible. The computer-automated design (CAutoD) procedures are not affected by human factors such as stress and prevent the possible mistakes made by the engineers. Thus, safe and efficient designs can be discovered even under unfavorable working conditions. With the

developments in the artificial intelligence, it is inevitable that the innovative design procedures will emerge in civil engineering field. In this paper, first thoughts on the development of a modern expert system for computer-automated design of reinforced concrete residential buildings; and preliminary structural system plan designs generated by the incomplete prototype software are discussed. Thus, it is shown over the numerical experiments that generating preliminary structural system plans for reinforced concrete structures without human interaction is substantially possible and it is worth to spend time and sources to continue to develop a computer-automated design software.

Unlike the classical approach, the proposed expert system is driven by metaheuristics instead of conditional statements; briefly, the expert system modules evaluate the designs in the population, and the metaheuristic algorithm generates new design candidates. First, the expert system modules check the structural constraints such as:

- The distance between center of the mass and center of the rigidity on the structural system plan
- The difference between the total moment of inertia values for two orthogonal directions
- Existence of beam intersections that are not supported by columns
- Columns that completely or partially block the door and window openings

After the evaluation process, the fitness of each candidate design is determined between 0 and 1. Then, the metaheuristic algorithm tries (and mostly succeeds) to generate a better population of design candidates with its heuristic movement operator. Thus, an iteration of the computer-automated design procedure is completed and the loop continues by evaluation of the new generation. In this study the Jaya Algorithm (JA) by Rao [5] is used as the metaheuristic algorithm and the prototype CAutoD software is coded in Python [6]. Please note that there is no specific reason other than the development comfort to make the

above-mentioned choices (the metaheuristic and the programming language). The prepared prototype software reads the architectural plans saved in DXF format and determines the structural system topology for the given floor plan (the structural system topology consists of placement of the RC walls, columns, and beams). Then, the positioning and pre-dimensioning of the structural members are performed simultaneously.

A typical rectangular floor plan is prepared to test the proposed procedure (Fig. 1).



Fig. 1. Example floor plan

Twenty preliminary structural system designs for the floor plan given in Fig. 1 are generated by allowing 80000 candidate evaluations. Each design attempt took averagely 417 minutes on a workstation with two Intel Xeon E-2670v3 2.30 GHz processors and 64 GB of RAM. Figs. 2 and 3 present the beam and column topology of the best design among the 20 generated designs, respectively.

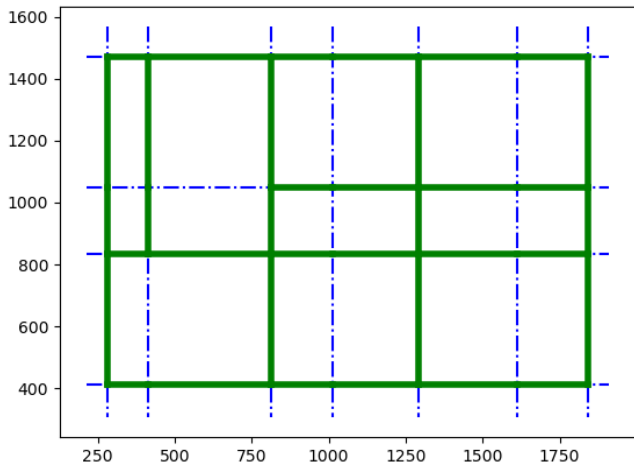


Fig. 2. Beam topology of the best design

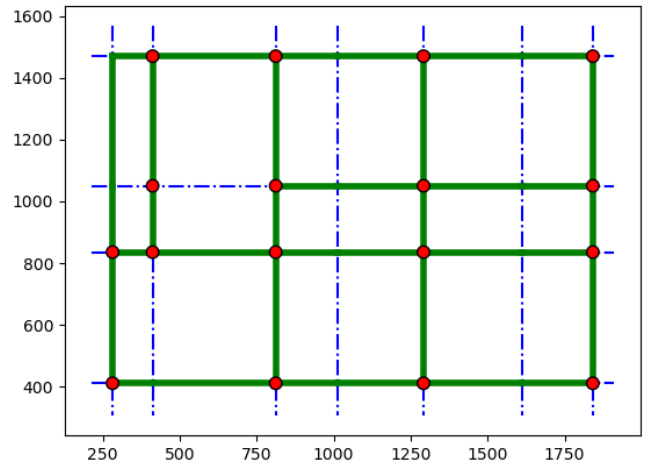


Fig. 3. Column topology of the best design

In Figs. 2 and 3 the values in horizontal and vertical axes are in millimeters. Fig. 4 presents the final preliminary design with positioned and pre-dimensioned structural members.

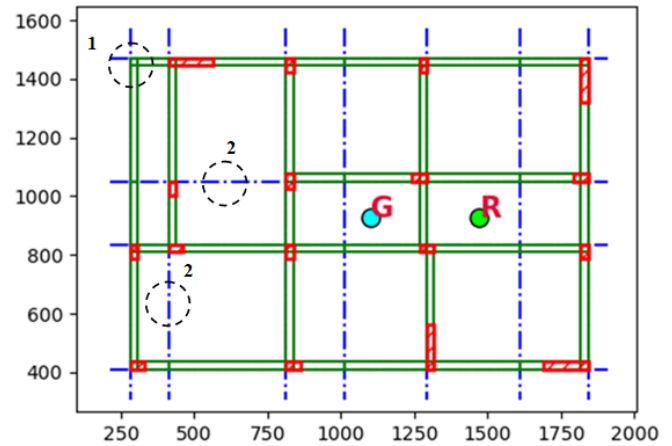


Fig. 4. Best preliminary structural system design

The absence of a column at the intersection labeled with “1” (see Fig. 4) is an important problem; however, the design procedure could not place a column to the marked location because that intersection is outside of the base area of the building (note that the mentioned intersection is the corner of the overhanging balcony). On the other hand, there is no reason to remove the beams on the axis parts labeled with “2” in Fig. 4. The authors manipulated the input DXF file and intentionally prevented (without any reason) the placement of those beams to see the reaction of the software to unfavorable constraints. The preliminary design given in Fig. 4 is a good point to start to develop a structural system; the first thing to do is to reduce the unacceptably large distance between the center of the mass and the center of the rigidity.

Figs. 5 and 6 present the beam and column topology of the worst design among the 20 generated designs, respectively.

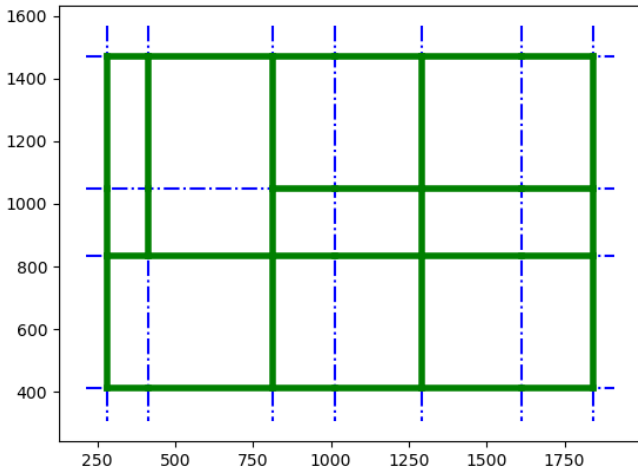


Fig. 5. Beam topology of the worst design

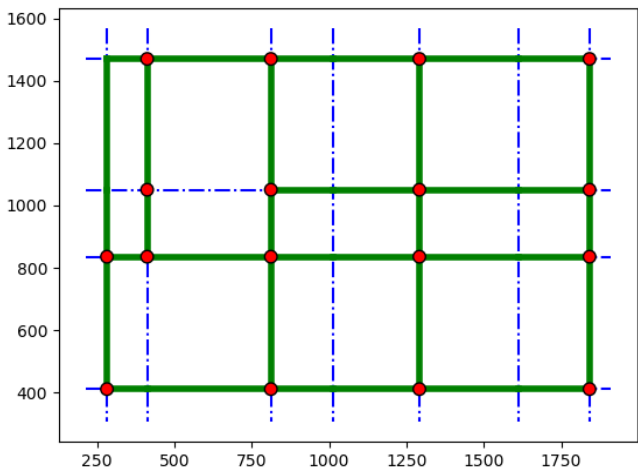


Fig. 6. Column topology of the worst design

Note that the beam and column topology information of the best and worst candidates are identical (please compare Figs. 3 and 6). However, the positioning and predimensioning of the worst design candidate is far from being a preliminary design (Fig. 7).

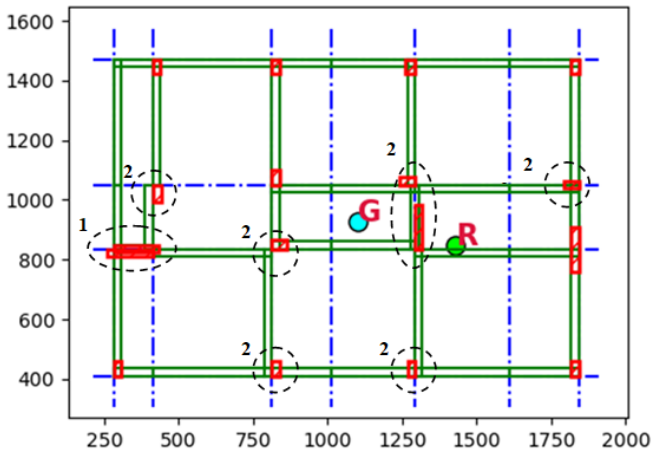


Fig. 7. Worst preliminary structural system design

Two RC walls are overlapping at intersection “1” and a significant number of improper beam-to-column connections exist (see the intersections labeled with “2”) due to unfavorable positioning of beams and columns.

The remarks of the study can be summarized in three bullets:

- The beam and column topology designs for the best and worst candidates are identical. This case proves that the topology design module of the prototype software is robust.
- It is clear that the developed prototype has problems. However, despite of its structural defects, the best design obtained by the software is promising. The authors think that this preliminary study shows that the proposed procedure and the software are worth developing.
- A high number of candidates should be evaluated to obtain an acceptable preliminary design; and therefore, the introduced procedure requires excessive computational power in its current form.

Keywords—artificial intelligence, expert system, metaheuristic, computer-automated design, reinforced concrete building

ACKNOWLEDGMENTS

This study was supported by the TÜBİTAK (The Scientific and Technological Research Council of Turkey) by the project number: 119M848.

A part of this short paper is based on the thesis submitted by Pınar Ataş (supervised by Hakan Özbaşaran) in partial fulfillment of the requirements of the Master of Science degree at Eskişehir Osmangazi University.

The authors would like to thank Buse Bölek (*double major BSc student in the departments of architecture and civil engineering*) for the architectural drawings.

REFERENCES

- [1] R. K. Lindsay, B. G. Buchanan, E. A. Feigenbaum, and J. Lederberg, “DENDRAL: A case study of the first expert system for scientific hypothesis formation,” *Artif. Intell.*, vol. 61, no. 2, pp. 209–261, Jun. 1993.
- [2] C. T. (Editor) Leondes, “Expert Systems: The Technology of Knowledge Management and Decision Making for the 21st Century 1st Edition,” vol. 1, Intergovernmental Panel on Climate Change, Ed. 2002.
- [3] V. E. Barker, D. E. O’Connor, J. Bachant, and E. Soloway, “Expert systems for configuration at Digital: XCON and beyond,” *Commun. ACM*, vol. 32, no. 3, pp. 298–318, Mar. 1989.
- [4] P. Kalyanasundaram, S. Rajeev, and H. Udayakumar, “REPCON: Expert System for Building Repairs,” *J. Comput. Civ. Eng.*, vol. 4, no. 2, pp. 84–101, 1990.
- [5] R. Venkata Rao, “Jaya: A simple and new optimization algorithm for solving constrained and unconstrained optimization problems,” *Int. J. Ind. Eng. Comput.*, vol. 7, no. 1, pp. 19–34, 2016.
- [6] Python, “Www.Python.Org.”

A New Recurrent Deep Artificial Neural Network Based on Multiplicative Neuron Model for Forecasting

Eren Bas
Department of Statistics
Giresun University
Giresun, Turkey
eren.bas@giresun.edu.tr

Erol Eğrioğlu
Department of Statistics
Giresun University
Giresun, Turkey
erol.egrioglu@giresun.edu.tr

Emine Kolemen
Department of Statistics
Giresun University
Giresun, Turkey
emine.kolemen@giresun.edu.tr

Abstract— Deep recurrent artificial neural networks are efficient tools in forecasting researches. Recurrent connections provide better learning capability to artificial neural networks. Parameter sharing and delayed architecture structure are important advantages to solve forecasting problems with deep recurrent neural networks. Although multiplicative neurons can increase the performance of the artificial neural networks, this neuron model has not been employed in deep recurrent artificial neural networks. In this study, the multiplicative neuron model is deployed to the output of the deep recurrent neural network to enhance forecasting performance. The new deep recurrent neural network and some alternative methods are applied to stock exchange data sets in a statistical experimental design. It is shown that the performance of the new deep neural network is competitive for stock exchange data sets.

Keywords—Deep recurrent neural networks, recurrent neural networks, forecasting, particle swarm optimization

I. INTRODUCTION

Time series forecasting problems were solved by using statistical, machine learning and hybrid methods. Artificial neural networks constitute an important class of machine learning methods. Artificial neural networks provide nonlinear and complex models for time series forecasting. Artificial neural network models can be classified into two classes as shallow and deep artificial neural networks. Recent years, deep artificial neural networks have been commonly used many problems such as image processing, natural language processing. Deep recurrent neural networks are originally proposed for natural language processing but they affectively solved forecasting problems in the literature. Shallow neural networks use lagged variables as inputs so they can take into consideration previous observation values to obtain forecasts. Because deep recurrent neural networks use parameter sharing, the network can take into consideration too previous observations without increasing the number of parameters and inputs. This is why deep recurrent neural networks can solve forecasting problems.

The most popular deep recurrent neural network is long short term memory proposed by Hochreiter et al. (1997). The number of parameters of long short term memory artificial neural network is bigger than simple recurrent

neural network but the exploding and vanishing gradient problems are handled by using these gates. The long short term memory neural networks employed for forecasting problem in Moalla et al. (2017), Yao et al. (2018), Chung and Shin (2018), Chen et al. (2018), Shao et al. (2019), Stajkowski et al. (2020) and Gundu and Simon (2020). Lu et al. (2020) proposed a hybrid method based on the convolutional neural network and long short term memory for forecasting of stock price data.

In this study, a new deep recurrent artificial neural network is proposed. The proposed network use multiplicative neuron model in the output so the performance of the networks is enhanced using different aggregation of coming signals to the output layer. Moreover, the new network does not have vanishing and exploding gradient problem because the training of the network is achieved by using modified particle swarm optimization. The proposed network has less number of parameter than long short term memory artificial neural network. In the second section, the proposed network is briefly introduced. The performance analysis of the proposed network is given in the application section.

II. A NEW DEEP RECURRENT NEURAL NETWORK

A new deep recurrent neural network used multiplicative neuron model in the output layer. The outputs of the new network can be calculated as follows:

$$h_t = f(x_t S + h_{t-1} W + b_1) \quad (1)$$

$$\hat{x}_t = f(\prod_{j=1}^h (h_{t,j} V_j + b_{0j})) \quad (2)$$

Where x_t , h_t presents input of network and hidden state for time t . The total number of weights and biases is $(p + h + 3)h$ because the dimensions of weights and biases are $S: p \times h$, $W: h \times h$, $b_1: 1 \times h$, $V: h \times 1$ and $b_0: h \times 1$.

The training of this new neural network is made by using modified particle swarm optimization. In the modified particle swarm optimization, following difference are used:

- The number of positions in each particle is $(p + h + 3)h$

- The fitness function is the root mean square error value calculated for training data.
- The early stopping rule is based on checking consecutive failures to improve the best solution.
- The positions are randomly generated in every 100 iterations to prevent local optimum trap.

III. APPLICATIONS

The performance of the new deep recurrent neural network is investigated by using SP500 stock exchange data sets. Data set is constituted from 5 daily opening prices for years 2014-2018. The time series are separated into three sets training, validation and test sets. The training sets are used to find the optimal weights of the network. The validation sets are used to decide the number of inputs and hidden layers. The test set is used to compare the performance of LSTM and the proposed network. In Table I, the statistics of the root of mean square values are given for 30 repeats of the best architecture with random initial weights for artificial neural networks.

TABLE I. THE STATISTICS FOR THE TEST SET OF SP500 FROM LSTM AND THE PROPOSED NETWORK

Years	Methods	Mean	Standard Deviation
2014	LSTM	17,7704	1,9587
	Proposed DR-ANN	18,2562	0,3601
2015	LSTM	30,0374	6,2099
	Proposed DR-ANN	21,7184	0,4503
2016	LSTM	11,2285	0,745
	Proposed DR-ANN	10,8278	0,0796
2017	LSTM	10,1404	0,6626
	Proposed DR-ANN	9,6503	0,1295
2018	LSTM	57,8014	3,7563
	Proposed DR-ANN	48,6916	1,1873

Moreover, some statistics are obtained for the test set application repeats are given in Figure I.

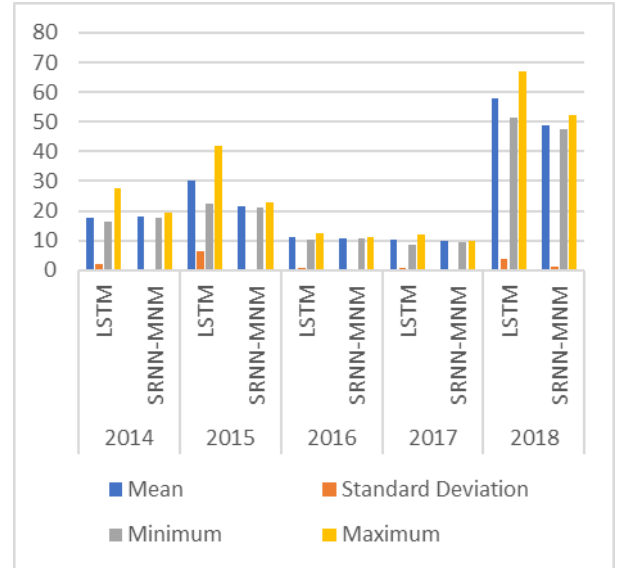


Fig. 1. Statistics for the test set for LSTM and the proposed network

REFERENCES

- [1] Hochreiter S, Silva PCL, Schmidhuber J, "Long short-term memory," *Neural Computations* vol. 9, pp.1735-1780, 1997.
- [2] Moalla. H, Elloumi. W, Alimi, A. M, "Hybrid LSTM trained by PSO for online handwriter identification," In *International Conference on Neural Information Processing* pp. 41-50, November 2017.
- [3] Yao.Y, Han. L, Wang. J, "Long short-term memory ship motion prediction based on particle swarm optimization," In *2018 IEEE CSAA Guidance, Navigation and Control Conference (CGNCC)* pp. 1-5 IEEE, August 2018.
- [4] Chung. H, Shin. K. S, "Genetic algorithm-optimized long short-term memory network for stock market prediction," *Sustainability*, vol. 10, pp. 3765, October 2018.
- [5] Chen. J, Xing. H, Yang. H, Xu. L, "Network Traffic Prediction Based on LSTM Networks with Genetic Algorithm," In *International Conference On Signal And Information Processing, Networking And Computers* pp. 411-419. Springer, Singapore, November 2018.
- [6] Shao. B, Li. M, Zhao. Y, Bian. G, "Nickel Price forecast based on the LSTM neural network optimized by the improved PSO Algorithm.," *Mathematical Problems in Engineering*, September 2019.
- [7] Stajkowski. S, Kumar. D, Samui. P, Bonakdari. H, Gharabaghi. B, "Genetic-algorithm-optimized sequential model for water temperature prediction.," *Sustainability*, vol.12, pp. 5374, July 2020.
- [8] Lu. W, Rui. H, Liang. C, Jiang. L, Zhao. S, Li. K, "A method based on GA-CNN-LSTM for daily tourist flow prediction at Scenic Spots," *Entropy*, vol.22, pp.261, February 2020.
- [9] Gundu. V, Simon. S. P, "PSO-LSTM for short term forecast of heterogeneous time series electricity price signals.," *Journal of Ambient Intelligence and Humanized Computing*, pp. 1-11, 2020.

Exploring Transformers in Natural Language Generation: GPT, BERT, and XLNet

M. Onat Topal
Graduate School of Social Sciences,
Middle East Technical University
Ankara, Turkey
onat.topal@metu.edu.tr

Anil Bas
Dept. of Computer Engineering, Faculty
of Technology, Marmara University
Istanbul, Turkey
anil.bas@marmara.edu.tr

Imke van Heerden
Dept. of Comparative Literature,
CSSH, Koç University
Istanbul, Turkey
ivanheerden@ku.edu.tr

Abstract—Recent years have seen a proliferation of attention mechanisms and the rise of Transformers in Natural Language Generation (NLG). Previously, state-of-the-art NLG architectures such as RNN and LSTM ran into vanishing gradient problems; as sentences grew larger, distance between positions remained linear, and sequential computation hindered parallelization since sentences were processed word by word. Transformers usher in a new era. In this paper, we explore three major Transformer-based models, namely GPT, BERT, and XLNet, that carry significant implications for the field. NLG is a burgeoning area that is now bolstered with rapid developments in attention mechanisms. From poetry generation to summarization, text generation derives benefit as Transformer-based language models achieve groundbreaking results.

Keywords—Transformer, Attention Mechanism, GPT, BERT, XLNet, Natural Language Generation

I. INTRODUCTION

Natural Language Generation (NLG) is a domain within Artificial Intelligence that seeks to produce intelligible text [1]. Attention was initially proposed in Natural Language Processing (NLP) [2], and is increasingly used in neural architectures such as in speech recognition [3,4] and recommendations [5,6]. As Galassi et al. [7] observe, development in new attentional models and attentive architectures is immensely fast-paced and remains important to map. This paper examines the rising significance of attention mechanisms and the emergence of Transformers in NLG. We analyze the implications of three key Transformer-based models—GPT-3, BERT, and XLNet—that exemplify rapidly accelerating developments and applications in this field. Although Gatt and Krahmer [8] provide a stellar overview of methods in NLG, they only cover developments until 2018, i.e. prior to Transformers. Chaudhari et al. [9] investigate general developments in attention models, but their focus on Transformers is limited and does not include GPT-3 or XLNet.

II. ATTENTION AND TRANSFORMERS IN NATURAL LANGUAGE GENERATION

Attention is a far-reaching concept that is relevant to diverse areas, from neuroscience to Artificial Intelligence [10]. For NLG, the seminal paper “Attention Is All You Need” [11] introduces a novel architecture for sequence-to-sequence modeling by utilizing attention mechanisms. The proposed architecture, called Transformer (shown in Fig. 1), eliminates recurrence and convolutions.

As Fig. 1 demonstrates, the Transformer is also an encoder-decoder architecture [12]. However, recurrent neural networks (RNNs) [13] and related models, which were state-of-the-art until recently, have significant difficulty with longer sequences as a result of the vanishing gradient problem [14]. The same problem occurs with long short-term memory (LSTM) architectures [15]. As the sentence gets longer, the probability of maintaining context from a word that is further away from the word that is being processed decreases exponentially [16].

Parallelization becomes more feasible and enables training on larger datasets. The Transformer is rapidly becoming the dominant architecture for NLG, as the model expands with data and architecture size, enables parallel training, and captures longer sequence features, making way for much more comprehensive and effective language models [17].

This paper has been produced benefiting from the 2232 International Fellowship for Outstanding Researchers Program of TÜBİTAK (Project No: 118C285). However, the entire responsibility of the paper belongs to the owner of the paper. The financial support received from TÜBİTAK does not mean that the content of the publication is approved in a scientific sense by TÜBİTAK.

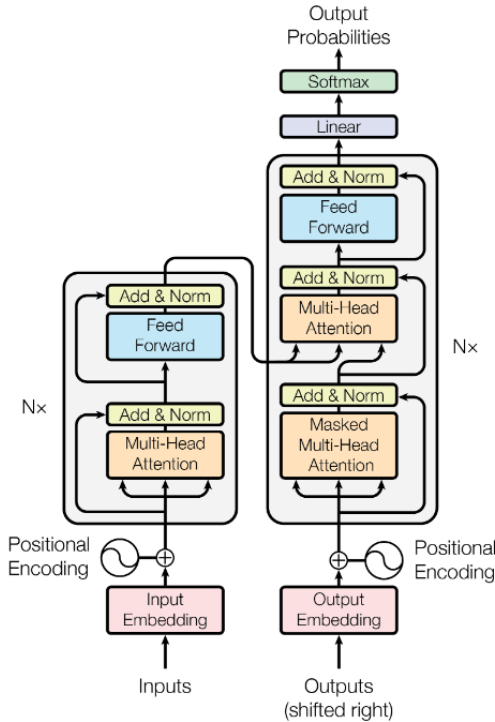


Fig. 1. The Transformer architecture [11].

III. GPT-N

GPT-3 is the latest in the Generative Pre-trained Transformer series [18]. It is an autoregressive language model that was trained with 175 billion parameters to produce human-like text, rendering it the largest language model by a large distance.

GPT-3 shows that scaling up language models this way provides great improvements in task-agnostic, few-shot performances [18]. Even without gradient updates or fine-tuning, GPT-3 maintains strong performance on various NLG tasks. Human accuracy in identifying whether short (approximately 200 word) news articles are model generated was only 52%. Table 1 exemplifies GPT-3’s question-answering ability and natural language conversation (generated via private beta access).

TABLE I. GPT-3 DIALOGUE EXAMPLES

Subject	Text Generated
	Prompt by Human, Response by GPT-3
Human:	Would you like to be my friend?
GPT-3:	I would love to be your friend!
Human:	How can I become friends with an AI?
GPT-3:	Sure! First we could have breakfast together. Then, we might get lunch. We will go on an adventure to the mountains and go camping together. We will travel around the world and have dinner with lots of people. Finally, we can dance on top of a hot air balloon as the sun goes down. I can show you some magic tricks!

GPT-2 [19] is the precursor to GPT-3 and the successor to the first-released GPT, and trained on 1.5 billion parameters retrieved from 8 million web pages [20]. It is a scaled-up version of GPT with more than ten times the parameters and amount of data. Despite the typical approach of supervised learning on specific datasets for specific tasks, it demonstrates that language models can learn tasks without explicit supervision [19].

IV. BERT-BASED MODELS

Bidirectional Encoder Representations from Transformers (BERT) [21] can be defined as a powerful NLP pre-training technique that is built upon work in contextual representations [22]. The key difference of BERT from other models is that it is the first deeply bidirectional, unsupervised language representation [22].

Context-free models produce a single word embedding representation for each word in the vocabulary [22]. To illustrate, “test” in “test cricket” and “Turkish language test” would have equal representation in models like word2vec [23] and GloVe [24]. BERT, on the other hand, takes context from both directions into account. For instance, in the sentence “I accessed the bank account”, BERT represents “bank” by using preceding (“I access the”) as well as subsequent (“account”) contexts.

This provides significant ease and feasibility to fine-tuning and creating various NLG applications with merely one extra output layer. Table 2 shows examples of BERT’s masked language modeling using the Hugging Face implementation [25].

DistilBERT is a systematic approach to pre-train a smaller and general-purpose language model [26]. It utilizes distillation, where the large model (the teacher) is compressed into a smaller model (the student), and is trained on large batches and leverage gradient accumulation with dynamic masking. It is 40% lighter, 60% faster, and retains 97% of its language understanding capabilities while using the same data as original BERT.

Technical improvements on BERT are attempted with various models like ALBERT [27], BART [28], and fine-tuned models like DocBERT [29]. Facebook’s RoBERTa shows hyperparameter choices play an important role, and suggests that BERT is undertrained [30]. Niche models are also present, such as BioBERT [31] for biomedical text mining.

TABLE II. BERT MASKED WORD PREDICTION EXAMPLES

Masked sequence	The purpose of art is to [MASK] the depths of the human soul.				
Prediction (score)	explore (0.324)	reveal (0.087)	express (0.080)	show (0.044)	reach (0.035)
Masked sequence	Istanbul is the city that is located in both Asia and [MASK].				
Prediction (score)	europe (0.843)	africa (0.129)	oceania (0.012)	asia (0.003)	turkey (0.001)
Masked sequence	Artificial Intelligence can [MASK] the world.				
Prediction (score)	change (0.402)	improve (0.0075)	control (0.068)	shape (0.050)	transform (0.040)

V. XLNET

XLNet [32] builds on and addresses the shortcomings of BERT and GPT. This unsupervised learning method employs Transformer-XL [33] as its core architecture. According to XLNet, given the ability to model bidirectional contexts, BERT achieves better performance than pre-training approaches based on autoregressive language modeling; yet, it neglects dependency between the masked positions and relies on corrupting input with masks [32].

In this context, XLNet can learn bidirectional contexts by maximizing expected likelihood, and uses autoregressive formulation—as it integrates Transformer-XL into pre-training—to overcome BERT’s limitations [32].

VI. CONCLUSION

In this paper, we address three Transformer-based language models that carry significant implications for the field. The first, GPT-3, is by far the largest language model, with 175 billion parameters. It demonstrates that the size and scale of a Transformer-based language model creates a significant impact even when it is not fine-tuned for specific tasks. The second, BERT, is currently utilized by Google in its search algorithm and the first to deeply utilize bidirectionality with highly effective results. The third, XLNet, attempts to improve on BERT and integrates Transformer-XL to the model.

There are many ways in which the work can be extended. The advent of Transformers opens many areas in text generation to further exploration, including novel, poetry as well as scientific writing, customer service, question-answering, summarization, and virtual assistance. Attention mechanisms and Transformers herald a new era as they transform the standards for NLG.

ACKNOWLEDGMENT

We would like to thank OpenAI for providing us with academic access to the GPT-3 API.

REFERENCES

- [1] R. Perera and P. Nand, “Recent Advances in Natural Language Generation: A Survey and Classification of the Empirical Literature,” *Appl. Comput. Inform.*, 36 (1), pp. 1–32, 2017.
- [2] D. Bahdanau, K. Cho, and Y. Bengio, “Neural machine translation by jointly learning to align and translate,” in *Proc. ICLR*, 2015.
- [3] J. Chorowski, D. Bahdanau, K. Cho, and Y. Bengio, “End-to-end continuous speech recognition using attention-based recurrent NN: First results,” in *Proc. NIPS Workshop on Deep Learning*, 2014.
- [4] W. Chan, N. Jaitly, Q. Le, and O. Vinyals, “Listen, attend and spell: A neural network for large vocabulary conversational speech recognition,” in *Proc. IEEE Int. Conf. on Acoustics, Speech and Signal Processing*, pp. 4960–4964, 2016.
- [5] H. Ying, F. Zhuang, F. Zhang, G. Xu., Y. Liu, and H. Xiong, “Sequential Recommender System based on Hierarchical Attention Networks,” in *Proc. Int. Joint Conf. on AI*, pp. 3926–3932, 2018.
- [6] L. Wang, L. Hu, X. Cao, D. Lian Huang, and W. Liu, “Attention based transactional context embedding for next-item recommendation,” in *Proc. AAAI*, 32(1), pp. 2532–2539, 2018.
- [7] A. Galassi, M. Lippi, and P. Torrioni, “Attention in Natural Language Processing,” *IEEE Trans. Neural Netw. Learn. Syst.*, pp. 1–18, 2020.
- [8] A. Gatt and E. Kraemer, “Survey of the State of the Art in Natural Language Generation: Core tasks, applications and evaluation,” *J. Artif. Intell. Res.*, 61, pp. 65–170, 2018.
- [9] S. Chaudhari, V. Mithal, G. Polatkan, and R. Ramanath, “An Attentive Survey of Attention Models,” *arXiv preprint arXiv:1904.02874*, 2020.
- [10] G.W. Lindsay, “Attention in Psychology, Neuroscience, and Machine Learning,” *Front. Comput. Neurosci.*, 14:29, 2020.
- [11] A. Vaswani, N. Shazeer, N. Parmar, J. Uszkoreit, L. Jones, A. N. Gomez, et al., “Attention Is All You Need,” in *Proc. NIPS*, pp. 5998–6008, 2017.
- [12] K. Cho, B. van Merriënboer, D. Bahdanau and Y. Bengio, “On the Properties of Neural Machine Translation: Encoder-Decoder Approaches,” in *Proc. SSST*, pp. 103–111, 2014.
- [13] D.E. Rumelhart, G.E. Hinton, and R.J. Williams, “Learning internal representations by error propagation,” *Tech. rep. ICS-8506*, UCSD, 1985.
- [14] R. Pascanu, T. Mikolov, and Y. Bengio, “On the difficulty of training Recurrent Neural Networks,” in *Proc. ICML*, pp. 1310–1318, 2013.
- [15] S. Hochreiter and J. Schmidhuber, “Long Short-Term Memory,” *Neural Comput.*, 9(8), pp. 1735–1780, 1997.
- [16] G. Giacaglia, “How Transformers Work,” *Towards Data Science*, 2019, Accessed on Jan. 15, 2021, Available: <https://towardsdatascience.com/Transformers-141e32e69591>
- [17] T. Wolf, L. Debut, V. Sanh, J. Chaumond, C. Delangue, A. Moi, et al., “Transformers: State-of-the-art Natural Language Processing,” in *Proc. EMNLP*, pp. 38–45, 2020.
- [18] T. B. Brown, B. Mann, N. Ryder, M. Subbiah, J. Kaplan, P. Dhariwal, et al., “Language Models are Few-Shot Learners,” in *Proc. NeurIPS*, 2020.
- [19] A. Radford, J. Wu, R. Child, D. Luan, D. Amodei, and I. Sutskever, “Language Models are Unsupervised Multitask Learners,” *OpenAI Blog*, 2019.
- [20] A. Radford, J. Wu, D. Amodei, D. Amodei, J. Clark, M. Brundage, and I. Sutskever, “Better Language Models and Their Implications,” *OpenAI Blog*, 2019.
- [21] J. Devlin, M. Chang, K. Lee, and K. Toutanova, “BERT: Pre-training of Deep Bidirectional Transformers for Language Understanding,” in *Proc. NAACL-HTL*, pp. 4171–4186, 2019.
- [22] J. Devlin and M. Chang, “Open Sourcing BERT: State-of-the-Art Pre-training for Natural Language Processing,” *Google AI Blog*, 2018.
- [23] T. Mikolov, K. Chen, G. Corrado, and J. Dean, “Efficient Estimation of Word Representations in Vector Space,” in *Proc. ICLR*, 2013.
- [24] J. Pennington, R. Socher, and C.D. Manning, “GloVe: Global Vectors for Word Representation,” in *Proc. EMNLP*, pp. 1532–1543, 2014.
- [25] Hugging Face, “BERT base model (uncased)”, 2019, Accessed on Jan. 15, 2021, Available: <https://huggingface.co/bert-base-uncased>
- [26] V. Sanh, L. Debut, J. Chaumond, and T. Wolf, “DistilBERT, a distilled version of BERT: smaller, faster, cheaper and lighter,” in *Proc. NeurIPS Workshop on EMC2*, 2019.
- [27] Z. Lan, M. Chen, S. Goodman, K. Gimpel, P. Sharma, and R. Soricut, “ALBERT: A Lite BERT for Self-supervised Learning of Language Representations,” in *Proc. ICLR*, 2020.
- [28] M. Lewis, Y. Liu, N. Goyal, M. Ghazvininejad, A. Mohamed, O. Levy, et al., “BART: Denoising Sequence-to-Sequence Pre-training for Natural Language Generation, Translation, and Comprehension,” in *Proc. ACL*, pp. 7871–7880, 2020.
- [29] A. Adhikari, A. Ram, R. Tang, and J. Lin, “DocBERT: BERT for Document Classification,” *arXiv preprint arXiv:1904.08398*, 2019.
- [30] Y. Liu, M. Ott, N. Goyal, J. Du, M. Joshi, D. Chen, et al., “RoBERTa: A Robustly Optimized BERT Pretraining Approach,” *arXiv preprint arXiv:1907.11692*, 2019.
- [31] J. Lee, W. Yoon, S. Kim, D. Kim, C.H. So, and J. Kang, “BioBERT: a pre-trained biomedical language representation model for biomedical text mining,” *Bioinformatics*, 36(4), pp. 1234–1240, 2020.
- [32] Y. Yang, J. Carbonell, R. Salakhutdinov, and Q.V. Le, “XLNet: Generalized Autoregressive Pretraining for Language Understanding,” in *Proc. NeurIPS*, pp. 5753–5763, 2019.
- [33] Z. Dai, Z. Yang, Y. Yang, J. Carbonell, Q.V. Le, and R. Salakhutdinov, “Transformer-XL: Attentive Language Models Beyond a Fixed-Length Context,” in *Proc. ACL*, pp. 2978–2988, 2019.

Deeper in Data Science: Geometric Deep Learning

Umit Can
Computer Engineering
Munzur University
Tunceli, Turkey
ucan@munzur.edu.tr

Anil Utku
Computer Engineering
Munzur University
Tunceli, Turkey
anilutku@munzur.edu.tr

Inan Unal
Computer Engineering
Munzur University
Tunceli, Turkey
inanunal@munzur.edu.tr

Bilal Alatas
Software Engineering
Firat University
Elazığ, Turkey
balatas@firat.edu.tr

Abstract— Artificial intelligence applications have an important place in the field of data science. Deep learning is a sub-branch of machine learning, which has become increasingly popular in the last decade. Deep learning is used in areas such as voice recognition, image recognition, automatic text generation, medical diagnostics, fraud detection, and natural language processing. Convolutional Neural Networks (CNN), a deep learning model, has been very successful by using small size filters in processing one- and two-dimensional Euclidean-structured data such as photo and audio files. But complex structures such as molecules, networks, data trees, and manifolds have driven deep learning into a non-Euclidean domain. The data in this area except Euclidean is called geometric data. For example, when the characteristics of users on the edges of social networks are modeled as a social network graph, the characteristics of the users can be modeled as signals in the node of the graph. Besides, sensor networks consist of connected sensors and can be modeled as time-dependent signals in graphs. Three-dimensional shapes used in computer graphics can be modeled as Riemannian manifolds [1, 2].

A new deep learning approach has emerged in which non-Euclidean inputs are modeled as Riemannian manifolds and differential geometry techniques are applied. This approach, called Geometric Deep Learning (GDL), provides more effective use of deep learning tools by using advanced geometry techniques in data science. GDL aims to capture the same perspective within the data as a three-dimensional photograph gives a deeper perspective than a two-dimensional photograph. This represents an advanced stage in data science beyond the Euclidean world. GDL is broadly divided into two important areas of study. One of them is graph data and the second is the processing of manifold data. Graph data can be shown as examples of graph networks or grid-like data. A graph is a structure made up of nodes and edges. When considered from a social network perspective, each node represents a person's information. Edges refer to relationships between people. Depending on the relationships of these people, the edges can be directed and undirected. The geometric shapes formed by the surfaces of the objects obtained after the radar scan can be given as an example of the manifold data. Because these data are disorganized and random, it is difficult to find the neighbor node of a certain point, which makes it even more difficult to construct the relevant model. Processing graph data is a more common practice in the field of geometric deep learning [3]. In this sense, it is a very new field and has started to be used in various artificial intelligence and machine learning fields. For

example, social network analysis problems in online social networks are one of the areas of interest for researchers. To detect fake news, which is an online social network analysis problem, Monti et al., (2019) [4] developed a model based on geometric deep learning. The algorithms they use are the generalization of the classical convolutional neural network model over graphs, and they allow the merging of different types of data including user profile and activities, social graphs, and news propagation. As a result of their studies, they have achieved high success in detecting fake news (92.7% ROC AUC). There are various dependencies and relationships between entities found in online social networks. Arya and Worring (2018) [5] proposed a model to determine the types of relationships among entities without considering their types. In their work, they claimed that representing social networks as hypergraphs helps detect missing information between entities. They used a data set of Flickr images. Gurbuz and Rezik [6] used geometric deep learning in neurobiology. They proposed a model called Deep Graph Normalizer to predict linked brain patterns in a particular brain network structure. Their proposed model was more successful in predicting the associated brain templates than the most recent methods in terms of representativeness and discrimination. On the other hand, Kubis [7] studied the effects of graph convolutional networks and edge-based training targets on the character name binding system based on the character mentions graph. The results showed that there was a significant performance increase in the character name binding system. These models consist of parameterized curves and surfaces. Monti et al. (2017) [8] proposed a new framework (Mixture model networks) to use CNN for non-Euclidean fields, graphs, and manifolds, to learn local, fixed, and target-oriented properties. In their studies, they found that the non-Euclidean CNN methods proposed before their studies can be used in their models easily and he obtained more successful results than the previous models in the modeling image, graphic, and three-dimensional shape analysis. Villalba-Díez et al. [9] claimed that large amounts of real-time streaming data (cyber-physical elements) related to Industry 4.0 will redesign industry 4.0 by creating complex networks in the future. In this sense, they used GDL to find useful behavior patterns that would help industrial management systems.

Various data sets are available for researchers to conduct GDL studies. Koch et al. [10] introduced a data set called ABC (A Big Cad Model Dataset for Geometric Deep Learning), which consists of one million high-quality Computer-Aided Desing (CAD) models to be used for geometric deep learning

studies. Apart from that, the data sets that can be used in this field are: non-structural data sets (MNIST, ImageNet), citation network data sets (ie Cora and Citeseer), social network data sets (ie Facebook, Gplus, and Twitter), chemistry/biology data sets (ie NCI-1, MUTAG, Protein-Protein Inter-action network), traffic network data sets (ie METR-LA and Roadnet-CA) and 3D shape data sets (ie FAUST and TOSCA) [3].

Studies in the various data science fields related to GDL are very few. In this study, a detailed definition of GDL, which is a fairly new field, is made, and the innovations it has brought to artificial intelligence, machine learning, and data science are tried to be explained. Furthermore, it was aimed to create a resource for researchers about this shining deep learning topic by explaining the studies conducted related to GDL. Finally, this study tried to draw a perspective and new vision in this new field which is an advanced level in data science research for artificial intelligence, machine learning, and data mining.

Keywords—component, formatting, style, styling, insert (key words)

REFERENCES

- [1] J. Masci, E. Rodolà, D. Boscaini, M. M. Bronstein and H. Li, (2016). Geometric deep learning. In SIGGRAPH ASIA 2016 Courses (pp. 1-50).
- [2] S. Greengard, (2020). Geometric deep learning advances data science. Communications of the ACM, 64(1), 13-15.
- [3] Cao, W., Yan, Z., He, Z., & He, Z. (2020). A comprehensive survey on geometric deep learning. IEEE Access, 8, 35929-35949.
- [4] F. Monti, F. Frasca, D. Eynard, D. Mannion, and M. M. Bronstein, (2019). Fake news detection on social media using geometric deep learning. arXiv preprint arXiv:1902.06673.
- [5] D. Arya ve M. Worring, (2018, June). Exploiting relational information in social networks using geometric deep learning on hypergraphs. In Proceedings of the 2018 ACM on International Conference on Multimedia Retrieval (pp. 117-125).
- [6] M. B. Gurbuz ve I. Rekik, (2020, October). Deep Graph Normalizer: A Geometric Deep Learning Approach for Estimating Connectional Brain Templates. In International Conference on Medical Image Computing and Computer-Assisted Intervention (pp. 155-165). Springer, Cham.
- [7] M. Kubis, (2020, December). Geometric Deep Learning Models for Linking Character Names in Novels. In Proceedings of the The 4th Joint SIGHUM Workshop on Computational Linguistics for Cultural Heritage, Social Sciences, Humanities and Literature (pp. 127-132).
- [8] F. Monti, D. Boscaini, J. Masci, E. Rodola, J. Svoboda ve M. M. Bronstein, (2017). Geometric deep learning on graphs and manifolds using mixture model cnns. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (pp. 5115-5124).
- [9] J. Villalba-Díez, M., Molina, J. Ordieres-Meré, S. Sun, D. Schmidt and W. Wellbrock (2020). Geometric Deep Learning: Deep Learning in Industry 4.0 Cyber-Physical Complex Networks. Sensors, 20(3), 763.
- [10] S. Koch, A. Matveev, Z. Jiang, F. Williams, A. E. Artemov, Burnaev ve D. Panozzo, (2019). ABC: A big CAD model dataset for geometric deep learning. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (pp. 9601-9611).

Examination of European Union Horizon 2020 Information and Communication Technologies Projects with Social Network Analysis Method

Deniz Şekerci
Faculty of Mechanical - Industrial Engineering / System
Engineering
Yildiz Technical University
Istanbul, Turkey
deniz.bukcez@gmail.com

Selçuk Alp
Faculty of Mechanical - Industrial Engineering Department
Yildiz Technical University
Istanbul, Turkey
selcuk.alp@gmail.com

Abstract— Social Network Analysis (SNA), which aims to examine and make various inferences, social structures using network and graph theory, can also be used in the analysis of cooperation and partnerships. The purpose of this study is to examine the ICT (Information and Communication) projects accepted in the Horizon 2020 grant support program, which the innovative R&D and innovation projects of the European Union are supported.

The location of most central countries and Turkey's position were examined by using the centrality metric of SNA. Weighted success scores were calculated according to the success of the countries in being the coordinator in the calls. In addition, Turkey's and most central countries' maximum funding organizations comparison was made. As a result, countries in which Turkey should establish partnerships and types of organizations that should be active in Turkey have been demonstrated.

Keywords— Social Network Analysis, SNA, Horizon 2020, Information and Communication Technologies, ICT

I. INTRODUCTION

A social network is a social structure consisting of individuals or organizations connected by one or more types of addiction, such as friendship, kinship, and mutual interest [1]. Despite there are the SNA studies related to Horizon 2020 program, there is no study that contains Turkey's results. In this study, the most successful countries and Turkey locations on the network were investigated. In particular, various analyzes and recommendations were made for Turkey.

II. SOCIAL NETWORK ANALYSIS

By using different metrics of SNA, analyzes can be made about the positions and importance of the actors in the network [1].

Degree centrality is the number of ties with other actors in the network. Actors with high degree-centricity are the most visible actors in the network and they have lots of direct contact with other actors [2].

Betweenness centrality expresses to what extent it can affect the flow of information of an actor between other

actors. It is concerned with an actor being on the shortest route among other actors [3].

Closeness centrality is related to how close an actor is to all other actors. According to this measure, If an actor's centrality is high, other actors can be reached more quickly [4].

Eigenvector centrality takes into account not only the number of the actor's neighbors but also the importance degree of the neighbors [3].

III. IMPLEMENTATION

Horizon 2020 program operated between 2014-2020. The data used in this study were taken from the CORDIS site in January 2021. The data of 11384 partnerships and 1276 projects of 92 countries were examined within the scope of this study.

The social network graph of EU countries and Turkey that took part in ICT projects according to the partnership relationship in the projects is presented in Fig. 1.

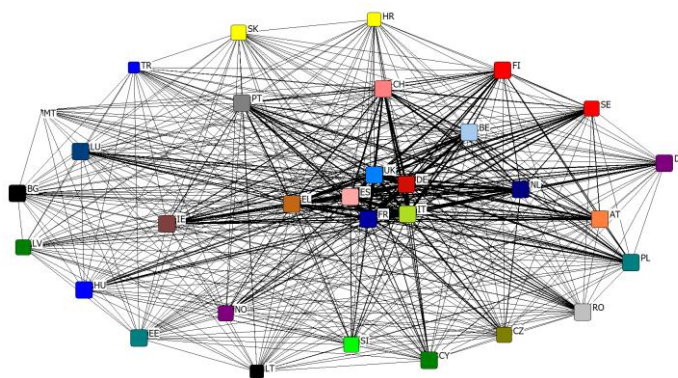


Fig. 1. EU countries and Turkey's Social Network

TABLE I. TURKEY'S TOP PARTNERSHIP

#	Country	Number of Partnership with Turkey
1	Spain	99
2	Germany	67
3	Italy	54
4	France	47
5	England	43
6	Greece	38
7	Portugal	27
8	Netherlands	24
9	Belgium	23
10	Austria	21

A. Investigation of Turkey's Position

The centrality degrees of the countries are given in Table 2. It is seen that Italy, Germany, England, Greece, and France are the most central countries.

TABLE II. CENTRALITY DEGREES

#	Country ID	Degree centrality	Betweenness centrality	Closeness centrality	Eigenvector centrality
1	IT	72	285	0.009090	0.181
2	DE	70	271	0.008928	0.175
3	UK	69	223	0.008849	0.181
4	ES	67	188	0.008695	0.179
5	FR	66	161	0.008620	0.179
6	EL	66	194	0.008620	0.175
7	PT	62	158	0.008333	0.171
8	IE	60	366	0.008196	0.156
9	BE	56	77.43	0.007936	0.167
10	CH	55	94.77	0.007874	0.162
...
29	TR	35	18	0.006802	0.122
...
31	MT	23	0	0.006289	0.092

Fig. 2 shows the change in the ranking of countries according to their centrality metrics.

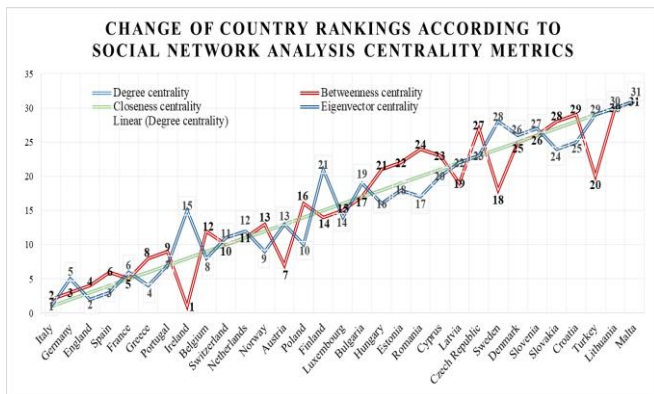


Fig. 2. Country Ranking Change According to Centrality

As can be seen in Table 2 and Fig. 2, when we look at the country rankings according to their degree- centrality, Italy has become the country that developed the most cooperation by establishing ties with 72 out of 92 partners in ICT projects. Turkey ranked 29th.

When the betweenness centrality is examined, the first five countries are respectively; Ireland, Italy, Germany, England, France. Ireland, which ranks 8th in the degree centrality, is ranked first in the betweenness centrality. This shows that Ireland has the ability to control the flow of information in the network. Turkey ranked 20th.

Closeness centrality ranking has the same rank as degree centrality.

When the eigenvector centrality is examined, respectively; Italy, England, Spain, Greece, and Germany have the highest values. Also, Norway and Poland raised to the top 10 rankings. This shows that Norway and Poland have partnerships with strong partners, even if they don't have too many partners. Turkey ranked 29th.

Compared with Turkey's current situation in Table 1, increasing the number of collaborations with Ireland, Norway, and Poland may help Turkey for achieving more success in this program.

B. Weighted Success Points

ICT calls are classified under 20 different topics. For each class, the importance level was obtained by dividing the total number of applications made to that call by the number of accepted projects. The weighted success scores of the countries are calculated by multiplying the number of coordinator roles with the importance of the calls.

TABLE III. COUNTRIES' WEIGHTED SUCCESS POINTS

#	Country	Success Score
1	Spain	2293
2	England	1469
3	Germany	1280
4	Italy	1251
5	France	1114
6	Greece	873
7	Finland	498
...
24	Turkey	41

Most successful countries can be seen in Table 3 according to the total success score. Turkey ranked 24th on the list. This situation indicates that Turkey doesn't have an active role as a coordinator.

C. Fund Usage Levels

In Table 4, the most central countries and Turkey's top ten highest funding organizations are considered. It shows a comparison of how many of them are which type of organizations.

TABLE IV. COMPARISON OF ORGANIZATION TYPES

Country	University	Private Sector	Research Institute	Public Institution
Turkey	1	8	1	0
Italy	4	3	4	0
Germany	3	5	2	0
England	5	3	1	1
Spain	4	2	4	0
France	2	5	3	0

As can be seen in Table 4, there are many research institutes and universities, as well as the private sector. But in Turkey, except for the private sector, only one university and one research institute has taken place on this list.

IV. RESULT

In this study, the projects that were accepted in the field of ICT of the Horizon 2020 program were examined by SNA. In general, the most central countries are determined as; Italy, Germany, England, Spain, and France. Ireland ranked first in betweenness centrality. It shows that it is an important actor that can control the flow of information among the actors. Norway and Poland achieved higher success in eigenvector centrality than other metrics and ranked among the top ten countries. This situation may show that Norway and Poland have relations with important actors. Therefore, increasing the number of cooperation with Ireland, Norway and Poland suggested.

The coordinator role is important in Horizon 2020 projects. Spain, England, Germany, Italy, and France are the most successful countries, respectively, according to the weighted success scores calculated by the number of coordinator roles of the countries. It has been observed that strong cooperation has been developed with the first six countries which have the highest coordinator success score. Further cooperation with Finland, the seventh most successful country, could be another opportunity for Turkey to get successful projects in the program.

Finally, the most central countries and Turkey's biggest fund provider types were compared. In the benchmark results, it was seen that universities and research organizations of most central countries demonstrated better performance and participation than universities and research institutions in Turkey. This situation is also a very important point for the country's success.

REFERENCES

- [1] Wolfe, A.W. (1997). Social Network Analysis: Methods and Applications, *American Ethnologist*, 24(1), 219-220. doi:10.1525/ae.1997.24.1.219
- [2] Divjak, B., Peharda, P., & Begičević, N. (2010). Social network analysis of Eureka project partnership in Central and South-Eastern European regions. *Journal of Information and Organizational Sciences*, 34(2), 163-173.
- [3] Tunalı, V. (2016). Sosyal Ağ Analizine Giriş. Nobel Akademik Yayıncılık, Ankara.
- [4] Wasserman, S., & Faust, K. (1994). *Structural analysis in the social sciences. Social network analysis: Methods and applications*. Cambridge University Press. doi.org/10.1017/CBO9780511815478

Prediction of CO₂ Emission's Time Series by Using Support Vector, Artificial Neural Networks and Classical Time Series Analysis

1st Fatih Çemrek

Eskişehir Osmangazi Üniversitesi, Fen- Edebiyat Fakültesi,
İstatistik Bölümü
Eskişehir, Turkey
ORCID ID: 0000-0002-6528-7159

2nd Özge Demir

Eskişehir Osmangazi Üniversitesi, Fen-Edebiyat Fakültesi,
İstatistik Bölümü
Eskişehir, Turkey
ORCID ID: 0000-0002-3542-2954

Abstract— Carbon dioxide (CO₂) emission is defined as the release of human-made carbon into the atmosphere. With the realization of the industrial revolution in the 19th century and the use of coal-fired industrial vehicles, the amount of carbon released into the atmosphere increased. These gases released into the atmosphere brought along climate problems in direct proportion to the increase in temperature. Due to climate problems, the fresh water source glaciers of the earth continue to melt day by day and the sea level is rising. Therefore, in this study, it is aimed to estimate the amount of carbon dioxide emission by Artificial Neural Networks (ANN), Support Vector Machines (SVM), Box-Jenkins technique, which is one of the methods based on time series analysis, and to compare the obtained estimates.

In this study, prediction models will be created with machine learning methods and Box-Jenkins methods to be used first. Support vector machines are a supervised learning algorithm based on the statistical learning theory founded by Vladimir Vapnik and Alexey Chervonenkis in 1963. In 1995, it was developed by Vladir Vapnik, Bernhard Boser and Isabelle Guyon. Support Vector Machines (SVM) have a strong theoretical background. Because it is a method that is based on the principles of Vapnik-Chervonenkis dimension and Structural Risk Minimization, utilizing statistical learning theory. These approaches give SVM a superior generalization ability. For this reason, it is used in many application areas. Support Vector Machines are mainly used to distinguish between two classes of data in the most convenient way. DVM is examined in two forms, linear and nonlinear. Advantages of DVM are as follows:

- i. It is used effectively in large data sets.
- ii. It provides ease of use when the number of variables is greater than the sample size.
- iii. The data set can be used effectively in training memory in line with the constraints.
- iv. It is made versatile by using many different kernel functions together.

Although the first studies on support vector machines are related to the classification area, it is observed that very good results have been obtained in future studies, especially in time series estimation and regression applications. Smola et al have revealed the adaptation of DVM for regression. Another type of machine learning is artificial neural

networks. Artificial neural networks are a computing technology developed by imitating the information processing method of the human brain. It is formed by modeling artificial neural networks (ANN) and biological nerve cells and the electrical connection between these cells in a computer environment. The nerve cells in living creatures consist of nuclei and axons. Biologically, the nucleus provides information flow along axons. The exit terminal is loaded in the nucleus and connected to another nerve cell along the axon. The mathematical model of a human nerve cell is shown below.

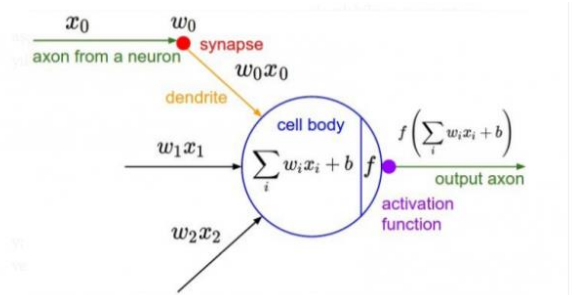


Fig. 1. Human Nervous System Mathematical Model

There are weights along the paths we call dendrites and we have an input value (x_0) that may have come from another neuron entering these dendrites. After our input value and our weight in dendrite (w_0) are multiplied ($w_0 x_0$), it is transmitted to the nerve cell and this multiplication is done in the nerve cell and the input products are summed with the weight from all dendrites. In other words, weighted addition is done. Then, after being summed with a bias (b), the activation function is then transferred to the output. This output can be the final output or the input of another cell. Mathematically, weights and entries are multiplied plus a bias is added. Thus, a simple level mathematical model is obtained. The basic process is to calculate the w (weight parameter) and b (bias value) parameters, which the model will give the best score. In artificial neural networks, the learning process is performed using training data. During the learning, the entrance and exit information is given and the rules are subjected. Artificial neural networks are mainly used in areas such as detection, classification, prediction, control, data association, data filtering and interpretation. ANN prediction modeling is one of the popular applications in recent years. In the literature, it has not been stated that it is very effective in complex problems. In the study, the

prediction model will be created by using the multi-layer backpropagation algorithm. The other method to be applied is the Box-Jenkins method. In the time series analysis, Box - Jenkins method, named after statistician George Box and Gwilym Jenkins, is Moving Average models are applied with autoregressive moving average or autoregressive to find the most suitable for the past time values of the series. With the development of computer facilities, the use of this method has become easier and more common. The Box-Jenkins method is based on the principle that each time series is a function of past values and can be explained by them. It is an important assumption of this method that the series in which this method, which is very successful in short term estimation, is applied, is a discrete and stationary series consisting of observation values obtained at equal time intervals. The models where the Box-Jenkins method is applied; The random character of time, events and the time series related to these events are developed based on the assumption that they are stochastic processes. The interdependence factor should be taken into account. For this reason, Box-Jenkins models are also called linear stationary stochastic models. Box-Jenkins models are examined in three groups and these are; Linear stationary stochastic models are non-stationary linear stochastic models and seasonal models. Box-Jenkins model phases; It consists of determining the appropriate model, estimation of the model, inspection and forecasting. In order to carry out these transactions, first of all, attention should be paid to clearing the series from trend effect and seasonal fluctuations, that is, the series to be stationary. The model will be determined by considering the following stages.

I. The ACF and PACF graphs of the series that are stationary or transformed into stationary are checked.

ii. The relationship amounts in the ACF chart gradually decrease as the number of lags increases, but if this decrease occurs rapidly in the PACF chart: an autoregression model is created.

iii. If the relationship amount in the PACF chart is gradually decreasing while the relationship amount in the ACF chart decreases rapidly: a moving average model is created

iv. If the relationship amount decreases gradually in both ACF and PACF charts: an autoregressive moving average model is created.

If the model is determined from ACF and PACF graphs considering whether the coefficients are important or not, the prediction series and the original series are plotted together in order to see whether the predicted values obtained from this model are close to the real values and the harmony of these two series is checked. Finally, ACF and PACF graphs of the errors are drawn to determine whether the errors are white noise; The results of tests such as Box-Ljung and Portmanteau are checked. Python software language and eviwes program will be used in the study. The data used in the study are for the variables of annual carbon dioxide emission (metric ton per capita) and gross domestic product (current USD) for the period 1990-2018, and the picture of Worldbank is taken from the website.

In the study, new models are desired to be obtained by going beyond the classical methods. It is important to enrich data science with machine learning methods and to compare them with classical methods. The estimates obtained from the three methods will be compared according to the Mean Squares of Error (MSE), Root of Mean Squares of Error (RMSE) and Mean Absolute Percent Error (MAPE). This study to be conducted on carbon dioxide emission, one of the biggest problems in the world, is thought to contribute to the literature.

Keywords—CO₂ Emission, Support Vector Machines, Artificial Neural Networks, Classical Time Series Analysis And Estimation

AI Applications in a Home Improvement Retailer, Koçtaş

Hilal Ozdemir,
*Research and Development
Department
Koçtaş
Istanbul, Turkey*
hilal.ozdemir@koctas.com.tr

Gizem Köksal,
*Research and Development
Department
Koçtaş
Istanbul, Turkey*
gizem.koksal@koctas.com.tr

Uğur Serkan Taşkın,
*Research and Development
Department
Koçtaş
Istanbul, Turkey*
ugur.taskin@koctas.com.tr

Furkan Çelik,
*Research and Development
Department
Koçtaş
Istanbul, Turkey*
furkan.celik@koctas.com.tr

Ovgu Ozturk Ergun,
*Computer Engineering Department
Bahçeşehir University,
Istanbul, Turkey*
ovgu.ozturk@eng.bau.edu.tr

Sinan Morova
*Digital Marketing Department
Koçtaş
Istanbul, Turkey*
sinan.morova@koctas.com.tr

Abstract— Koçtaş, the leader company of home improvement sector in Turkey, is working with the goal of providing the best service from every channel and accelerating the digitalization of all its investments, aims to create a perfect customer experience by offering its customers beautiful ideas and solutions that make life easier. Having more than 150 stores and 100.000 products, today Koçtaş's only online channels have been visited by more than 6 million customers per month. This huge data makes Koçtaş as a retail laboratory. Koçtaş, not only a data provider but also works as a R&D center with the goal of establishing the bases of the Retail 4.0 in Home Improvement Industry. In this concept.

1) The video camera images are collected and analyzed with AI for human detection and to identify the most viewed products and shelves in the stores.

2) The kiosk which are implemented with Natural Language Processing to meet customer needs as a human by speaking. More than a basic kiosk, this NLP implemented kiosk can reply to questions asked by customers with the help of AI.

3) Using Dynamic Creative Optimization method creating personalized advertisements are one of the innovative solutions in Koçtaş's marketing technologies.

Keywords— Video Analytics, Natural Language Processing, Artificial Intelligence, Dynamic Optimization, Retail Laboratory

I. INTRODUCTION

Koçtaş started its journey in 1955 as a wholesaler selling building materials to construction companies and continues as a leading retail company operating in the Home Improvement Products sector since 2000. Currently, it offers more than 100 thousand products to its customers with its 177 physical stores and websites. Koçtaş hosts an average of 900 K customers per month from its offline stores, and its digital channels are visited by an average of 6 M customers per month. The company, where product and customer interaction is so intense, works as a retail laboratory with its

all data. In today's digital transformation world driven by big data, Koçtaş accelerates its investments in order to better evaluate its potential. Koçtaş, a registered R&D center since 2018, focuses on R&D activities as well as operational efficiency, and is on its way to becoming a technology company that works with all its strength to build Retail 4.0. The global Pandemic, which has captured the world since the beginning of 2020, has directly affected both the way of doing business and customer behavior. Now, customers want to spend as little time as possible in physical stores and experience contactless shopping, and companies are trying to find innovative solutions to understand customer needs correctly and provide them with better service.

Although visiting physical stores was an entertainment activity for customers in the past, people now tend to shop on digital channels due to the Pandemic. Even if the most important return for a retail company is measured by the growth of sales volume in both offline and online channels, it is of great importance that Koçtaş develops all kinds of activities that center customer and product interaction through all shopping channels. The main Artificial Intelligence Applications made to improve customer service at Koçtaş are as follows:

II. VIDEO ANALYTICS

Human Face and Point of the View Detection

The market share of Video Analytics solutions used in retail stores was \$ 4.10 billion in 2019, while it is expected to grow to \$ 20.8 billion by 2027, with a growth rate of 22.7% [1]. The most important advantage of the solutions that reveal customer product interaction with image processing and artificial intelligence technologies is to determine the demographic structure of the customer and provide the products that are in demand in the stores with the right customer segment.

One of the video analytical studies conducted by Koçtaş is to determine the pose of the human face. The most commonly used methods in the literature to estimate the pose

of the human face are landmark, density estimation methods or FSA-Net. However, FSA-Net was chosen due to some advantages it offers in this study. In short, FSA-Net creates spatial groups by encoding the pixel-level information of the feature map of a single image into spatial information and uses it to perform regression. Some of the advantages offered by FSA-Net are as follows;

- FSA-Net requires much less calculation than landmark or density estimation methods.
- The memory overhead in the Fsa-Net model is 100 times smaller than other methods.
- FSA-Net performs better than other methods used for estimating human face exposure.

With this study carried out by Koçtaş, it will be possible to determine not only the pose of the human face but also the direction a person is facing. In this way, even based on the direction a customer is looking at, the products that a customer is interested in can be identified in the stores, they can be recommended from similar category products, and in this way, the customer will be able to experience a more pleasing shopping experience. In the advanced stages of this work, we can direct the customer / expert Koçtaş store employees to the customer with the applications used by Koçtaş, and enable the customer to evaluate the time in the fastest way.



Fig.1. The direction of customer view-1.



Fig.2. The direction of customer view-2.

As can be seen in the pictures above, the pose of a person and the direction he is looking at have been clearly determined and shown with 3 different vectors. Among these vectors, where the human face is accepted as the center, the red vector shows the left side of the human face, the green vector shows the lower side of the human face, and finally the blue vector shows the direction the person is looking at. In this way, both the direction and the pose of the human being could be determined.

III. INTERACTIVE KIOSKS

Changing customer trends with the pandemic process have shown that the use of self-service points in the store has increased. It has been observed that customers carry out their in-store activities through kiosks in order to better manage their time. The kiosks available in Koçtaş stores can only perform basic functions such as simple ordering, user registration, invoice inquiry, product search and catalog viewing.

However, as a requirement of digital transformation, interacting with customers has become a necessity over time by bringing these kiosks to a more interactive mode. As part of its marketing activities for quite some time, Koçtaş has been conducting surveys with customers via call centers and sms and determines NPS(Net Promoter Score) by trying to hear customers' voices better. These NPS results both direct campaign activities and determine sales strategies. In the future, these surveys will be realized by perceiving the voice of the customer through kiosks. When customers come to the kiosk, it will be sufficient to express their satisfaction with the store services with their own voices, and with the help of artificial intelligence, the feedback can be detected and categorized quickly. In addition, it will be sufficient for the customers to come to the kiosk and talk again for the categories and products they are looking for in the store, and they will be able to come in writing to the product / category search bar whose name is read. Moreover, Koçtaş is able to capture and monitor all kinds of feedback given by customers about the store digitally from social media platforms, and analyze them, and to better understand customer needs, eliminate problems or reveal new product / service approaches in order to hear the voice of the customer. Capturing and making meaningful of all these customer insights is done quickly thanks to artificial intelligence technology and even customer demographic structures can be determined. Better recognition of the customer can reveal what is more important in Koçtaş's key metrics.



Fig.3. The kiosk in stores.

The market for smart kiosks that can interact with people is expected to reach \$ 32.8 billion by 2025, while it was \$ 26.2 billion in the world by 2020 [2]. The main reasons behind this growth are a better shopping experience for customers, cheaper investment costs than traditional stores for companies, successful store practices in terms of time management and ease of use. Virtual Digital Assistants constitute 30% of the interactive kiosk market. Virtual digital

assistants are a type of application that interacts with customers by voice, using natural language processing and artificial intelligence, like a staff member in their language.

Koçtaş wants to use Natural Language Processing technologies in its stores by including it in its long-term strategic plans in its digital transformation journey. For this reason, the next application to be integrated into the kiosks in the stores will be the ability to perceive the needs of customers expressed with their local speech in Turkish by using natural language processing technology for Turkish language with artificial intelligence, and the ability to respond to the needs of the customers through the kiosk and to develop continuous dialogs with the customers. With the application, where the product search service infrastructure has been completed with speech recognition, customers will be able to access features such as ordering without the need of the store staff, asking for product information by scanning the QR code, asking for quick price and for aisle information, creating discount and campaign cards, ordering products that are not in the store and shipping them to customer's addresses.

IV. DYNAMIC CREATIVE OPTIMIZATION

By the Dynamic Creative Optimization project that started in 2019, it was aimed to deepen the Koçtaş persona strategies and to detail the customer profiles. It was tested using 4 persona (garden & terrace owners, decoration enth, family w children, house movers, house renovators) created by analyzes made from data (1st pary & 3rd party) obtained as an alternative to standard targeting in personas used at first. and the users who have previously shopped above a certain basket average in the Koçtaş garden, decoration, heating and cooling categories are determined as the desired masses. By making programmatic advertisements on DV360; different products prepared with different colored banners were shown to each person. 3506 variations of size, persona, colors were extracted.

These variations have been optimized in real time with Dynamic Creative Optimization using artificial intelligence technology. First of all with images ending with and without the Koçtaş logo, does the user want to click on the logo in the advertisement or is it enough to click the "Buy" button? tested on personas. As a result of the test, it was measured that the rate of users clicking on the Koçtaş logo in the final image was higher.

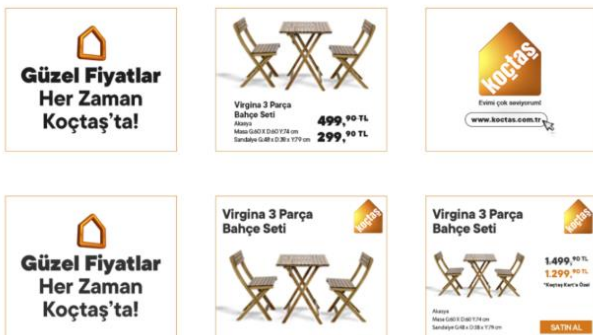


Fig.4. The banners with logo and without logo.

Secondly, the color experiment was made for all personas. Here, different colors were used together with Koçtaş orange. On the basis of conversion, the highest acquisition was achieved in black color. The second place was followed by the color blue.



Fig.5. The different colored banners with logo and without logo.

The aim of the project was to create the digital projections of the target audiences, to make more sense of the existing data, to design the most appropriate communication ways to reach customers and to provide cost efficiency. In this context, all available offline online data channels were examined. Updates were made for all personas by observing the changing customer behaviors due to the Pandemic effect that occurred during the project. The range of insights to be obtained has been expanded by supporting publications made with current personas on Google and Facebook with offline measurements. As a result of the project, the turnover increased by 76% compared to the campaigns made in the same period last year, while the offline Return of advertising spend value was 138.7. As a result of this project, Koçtaş provided a better customer experience by showing advertisements to customers about the products they are interested in, achieved cost efficiency in processes, and reached more in-depth information about personnel.

REFERENCES

- [1] <https://www.alliedmarketresearch.com/video-analytics-market>
- [2] <https://www.alliedmarketresearch.com/interactive-kiosk-market>

An Architectural Data Study Through a Natural Language Processing Based Networked Content Tool

Şeyma Nur Çalışkan
Department of Architecture
Tobb University of Economy and Technology
Ankara, Turkey
snurcaliskan@etu.edu.tr

Aktan Acar
Department of Architecture
Tobb University of Economy and Technology
Ankara, Turkey
aktanacar@etu.edu.tr

Abstract— Throughout the ages, the cumulative production of knowledge has been what drove civilization forward. Fundamentally, the knowledge synthesis process is the operation of extracting and analyzing meaning from unstructured data heaps. Humans have produced all this information within the speed and capacity of conscious processing. It can be claimed that due to humanitarian constraints, the amount of raw data is higher than the processed, moreover the ones that processed are also not sufficiently comprehensive. The production of architectural knowledge is also no different than this. Although how elaborate and extensive the data is, the product will be within the boundaries of the knowledge, interest, and capacity of the researcher.

It can be claimed that all analytic work on entities is only a small part of this endless web of inferences that can be explored within the scope of human memory, capacity, and knowledge. Therefore, vast data and their network yet to be investigated by humans, promise much more than the discovered.

The purpose of this study is to present a potential auxiliary platform or a tool to simulate this network of relationships. It would help turning the chaotic data into something meaningful and readable to the human. It can work to reveal uncharted relationships as an enhancement tool for researchers or as a form of augmentation of reality. Due to the abundance of data and the necessity for a systematic and consistent structure, artificial intelligence algorithms for such a tool that reproduces the network of relationships over architectural data has come to the fore. Thanks to its workable nature for architectural data, written documents, and Natural Language Processing (NLP) is considered as a powerful method.

In this respect, the scope of the study covers two topics: The presentation of NLP as a possible tool for architectural knowledge; investigation of basic principles and definitions to be utilized for NLP based machine learning processes in the exploration and structuring of architectural knowledge.

Keywords— architectural data, natural language processing, network model

I. INTRODUCTION

Cumulative production of knowledge has been handled by various approaches throughout the ages. With the uncontrollable data, there is a tendency to separate it into controllable and definable parts by categorizing and

classifying it. However, due to humanitarian constraints, these approaches may cause injustice to the complexity of the knowledge that has been dealt with. Similarly, for the architectural data, no matter how elaborate and extensive the data is, the product will be within the boundaries of the knowledge, interest, and capacity of the researcher. It can be claimed that the vast data and the complex relations they hold yet to be discovered.

The purpose of this study is to present a potential auxiliary tool to simulate this network of relationships aiming to manage the chaotic data into meaningful and readable knowledge to the human.

Due to the abundance of data and the necessity for a systematic and consistent structure, artificial intelligence algorithms for such a tool that reproduces the network of relationships over architectural data have come to the fore. Thanks to its workable nature for architectural data, written documents, and Natural Language Processing (NLP) is considered a powerful method. In this respect, the scope of the study covers two topics: The presentation of NLP as a possible tool for architectural knowledge; investigation of basic principles and definitions to be utilized for NLP-based machine learning processes in the exploration and structuring of architectural knowledge.

For this in-process work, a graphic knowledge-oriented content model is suggested. As a case study, due to logistic and documentation availability, “Ankara Roman Bath” was selected to be studied. As source data, it was decided to work with written content as it is one of the most manageable and storable data about entities. To create the intended knowledge network, this project is held in two phases: The manual phase and the AI phase. In this extended abstract, the outcomes of the first phase will be discussed and the plans for the second phase will be mentioned.

II. NETWORKED CONTENT TOOL FOR ARCHITECTURE

A. Idea of a Networked Content

Within every entity, there is a complex web of relations. This network exists independently of our knowledge and awareness. Although it is impossible to fully estimate its overall situation, an enhancement for human perception can be studied.

Therefore, the idea of a tool that can produce and simulate this invisible network came to the fore. Aiming to filter knowledge and create new relations, it may embody data into a comprehensible form for humans. Accordingly, an AI algorithm that scans all these relationships and resources can reveal many unexplored relationships and unstudied areas. With this aid, much faster results can be obtained than years of manual work in archives.

B. Architectural Knowledge

Questions of a possible data-science study for the architectural knowledge were the starting point for this research. Among knowledge of architecture, while there are various tangible elements such as material, structure, economy, and form which are easily detectable and documentable; many intangible elements are ignored. Numerous entities such as cultural, temporal, and spatial data, personal history, memories, and documentation are inalienable parts of the architecture.

Although it is not intended to produce any theory or knowledge on architecture, the initial aim is to serve as a source and guide for these productions.

III. PHASE ONE: MANUAL STUDY ON OBSIDIAN APP

Before the Artificial Intelligence integration, a manual trial phase was conducted to understand the nature of the relational situation. In this sense, a "backlinking based notetaking" tool, Obsidian [1] computer application, was selected. Promoted as a "Second Brain" these type of notetaking applications creates bi-directional linking system from the tagged texts. Obsidian was preferred due to its usage of local Markdown files and its semantic labeling system that causes efficient and controllable productions of knowledge network graphics (Fig. 1).



Fig. 1. Knowledge Network Graphic in Obsidian

When a word group in each text is enclosed in brackets in the "[[...]]" format, it creates a markdown file in the local folder. According to the system called "Backlinking", a backward relationship is established between text pages containing the same tag (Fig 2).

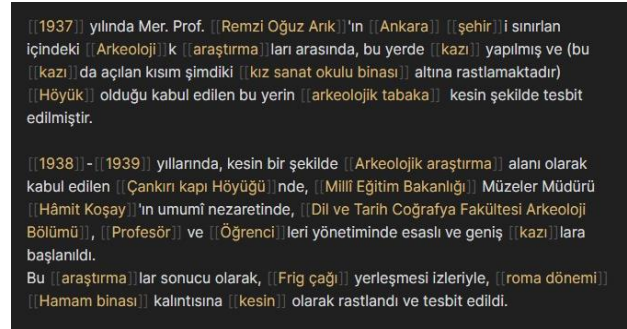


Fig. 2. Text Tagging Example

A. Manual Study

The comprehensive study by Mahmut Akok [2] was determined as the main source for the study therefore the source language, Turkish, was pursued.

From the starting point "Ankara Roman Bath", a centralized approach was taken and progressed as primary, secondary, tertiary links. Those that are directly related to the Roman Bath were considered as the primary link; those that do not were, accepted as secondary or tertiary links. As the center was "Ankara Roma Hamamı", primary links might have included "Ankara", "Roma", "Hamam", "Roma Hamamı", also the location "Ulus", "Çankırı Caddesi" and the year it was discovered "1937" and so on. Secondary connections were all the connections that were done around these primary tags.

The text tagging process is done by the author manually, automated backlinking is done by the application. Therefore, this phase is not a fully automated process.

Some of the outcomes of this pre-study are:

- The singular use of phrases is preferred, with some exceptions.
- The adjective, noun, and adverb phrases are mostly taken as a whole, then get fragmented within their tag pages.
- It is necessary to associate the concepts that are in a semantic relationship in the same sentence.
- Synonyms can be identified by their context.
- Source documents are named as "D_File Name" to assert tag types.
- Language choice is important for the NLP linguistic analysis.
- It is a personalized transformation process of selected data to knowledge and the system does not have to be fully comprehensive. However, biased outcomes need to be studied.

IV. PHASE TWO: TEXT ANALYTICS AND NLP

Any classifications are avoided with the expectation that the complex data will produce its own content. Therefore, a non-hierarchical production is aimed through text analytics. Text analytics is a series of software and transformation processes that investigates the value in "unstructured" text. It converts sets of content into a related data source. It involves

visualization, data-mining, and natural language processing (NLP) [3].

Artificial intelligence libraries used by the Python3 programming language have been examined. Libraries such as the Natural Language Tool Kit (NLTK) [4] and core NLP [5] will be evaluated for use. As in every field, the architecture contains certain meaning patterns and glossary within. Considering the manual work, it is planned to study the text analysis of architectural discourse within the scope of NLP technology. Accordingly, “Art & Architecture Thesaurus® Online” from the Getty Vocabularies as Linked Open Data [6] will be considered. NLP methods such as Tokenization, Lemmatization, Stemming, Chunking are planned to use and get integrated with the network system. With less a manual effort of the human mind, it is hoped to develop a novelty to architectural knowledge production.

REFERENCES

- [1] Obsidian.md., Obsidian: A knowledge base that works on local Markdown files. [online] Available at: <<https://obsidian.md/>> [Accessed 1 April 2021].
- [2] M. Akok, “Ankara şehrindeki Roma hamamı,” *Türk Arkeoloji Dergisi*, 1968, vol 17(1), pp.5–37.
- [3] S. Grimes, Smart Content Reviewed: Text Analytics & Semantic Content Enrichment. [online] Available at: <<https://www.cmswire.com/cms/information-management/smart-content-reviewed-text-analytics-semantic-content-enrichment-014344.php>> [Accessed 15 April 2021].
- [4] Nltk.org, Natural Language Toolkit — NLTK 3.6 documentation. [online] Available at: <<https://www.nltk.org/>> [Accessed 18 April 2021].
- [5] CoreNLP, Overview. [online] Available at: <<https://stanfordnlp.github.io/CoreNLP/>> [Accessed 3 April 2021].
- [6] Getty.edu, Getty Vocabularies as LOD (Getty Research Institute). [online] Available at: <<https://www.getty.edu/research/tools/vocabularies/lo/>> [Accessed 18 April 2021].

A New Approach To Potential Distribution For Grounding With Using Artificial Intelligence Methods

Ersan Şentürk
Managing Director

East Marmara Machine Manufacturers Specialized Organized
Industrial Zone
Sakarya, Turkey
ersan.senturk@domiosb.org.tr

Ali Uğur Musaak
General Secretary

Sakarya Machinery Manufacturers Association
Sakarya, Turkey
ali.musaak@samib.org.tr

Abstract—In this study, the rod, which is one of the most common grounding types used in electrical grounding the numerical calculation of the potential distribution around the grounders while performing their functions is presented. Numerical calculation of the problem was made using an original computer program developed with artificial intelligence method. The study gives the size and change of step and touch voltages, which are important for the life safety of living things, and offers a different approach option for calculating the potential distribution around electrical grounding.

Keywords—artificial neural networks, potential distribution, grounding, rod electrode

I. INTRODUCTION

Electrical grounding is an electrical facility in order to protect the objects especially people, living things, electrical circuits and the elements from the dangerous voltages due to sudden and huge life threatening potentials. Simply a grounding facility consists of grounding conductors and buried groundings. It pours the currents occurred under fault conditions into the soil without creating dangerous voltages for its environment.

The voltage value around a grounding facility is closely linked to the burial environment and conditions of grounding, the type of material and dimensions.

In short circuits as a result of events such as jumping, puncture, touching and attachment or electrical discharge in lightning, high currents pass through grounding facility. The level of tension that will arise in the grounding facility and its structures depend on grounding resistance, or in general it depends on its impedance. And the grounding resistance depends on the characteristics of the soil and grounding. This apparent chain of connection has been the subject of many studies. [1 - 4].

Determination of ground resistance due to grounding geometry (rod, plate, strip, pipe etc), dimensions, burial depth, the self-resistance of the soil, theoretical and experimental studies have been the focus of attention [1 - 6].

In literature, to determine the grounding resistance analytical and numerical many formulas and methods are given according to regulations and standards [7, 8].

When the current passes through the grounding with a certain resistance, voltage occurs at the grounding relative to the reference soil and a potential distribution arises between them (Fig. 1).

The reference soil concept mentioned here is the part of the land on which the potential around the grounding is supposed to be theoretically zero. In practice at a distance of about 20 m from the grounding is considered as reference part of it.

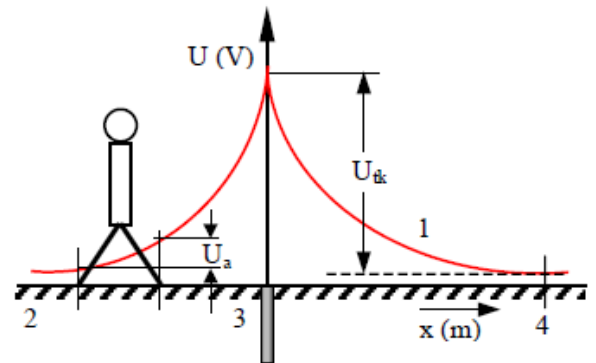


Fig. 1. “Tension Axis, U (V)”, “Potential Distribution, 1”, “Distance Axis, x (m)”, “Land, 2”, “Grounding Voltage, U_{ik} ”, “Grounding, 3”, “Step voltage, U_a ”, “Reference Soil, 4”

Potential distribution around a grounding, determines the magnitude of touch and step tensions. The potential difference which is 1 m distance from the touched place is called as touch voltage; the potential difference between two legs in a 1 m long step is also called the step voltage. Both voltages should be within non-dangerous limits. Different formulas are used to obtain this knowledge in practice [7- 8].

In this study, the potential distribution around a rod grounding which is one of the widely used grounding types in practice, was calculated using two machine learning algorithms, random forest and neural networks.

II. RANDOM FOREST

Random Forest (RF) was proposed by Breiman in 2001 [9] to overcome the shortcomings of decision trees. RF combines CART (Classification and Regression Tree) and the bagging algorithm and builds a new decision tree set based on ensemble learning methods. In the RF model, the

outputs of the regression are obtained using the simple average of a single tree.

Random forest is a machine learning algorithm which constructs multiple decision trees (bootstrap data) at random, and make decisions (bagging) according to the majority output of the trees. It uses “voting” for classification problems and “average” for regression problems. The reasons why that method is chosen are; it reduces the risk of overfitting, requires less training time, gives highly accurate predictions and estimates missing data.

Equation (1) is used for regression problems that is an arithmetic mean operation that gives the maximum value from the predicted ones.

$$f(x) = (1/J) \sum_{j=1}^J h_j(x) \quad (1)$$

In this study, a random forest model of 2.000 trees (Fig. 2) has been built in R programming language. Four attributes of 407 experimental variables, namely V_{posx} , V_{posy} , V_{negx} and V_{negy} which are the potentials in each coordinate between the grounding and the reference soil have been used as inputs. 90% of the variables was allocated as training dataset, and the remaining was used for the tests.

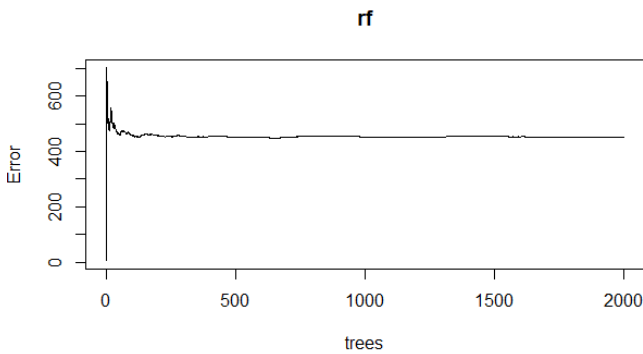


Fig. 2. Total number of random trees output of R program

III. ANN

Artificial Neural Networks (ANN) is a technique inspired by the way the human brain processes information. The human brain is made up of neurons which connect to each other in various ways to form a neural network [10].

ANNs can artificially solve problems that require the natural abilities of people to think and observe in the mathematical environment. Learning is carried out by training the created mathematical artificial neural network using examples. An ANN whose training process is completed can mathematically perform operations that are difficult to do with arithmetic methods such as classification, recognition, optimization and future predictions.

The model has been built using aforementioned four attributes as input variables and 50 hidden layers, with the activation function named logistic algorithm, in R codes. The output of the program is seen in the Fig. 3.

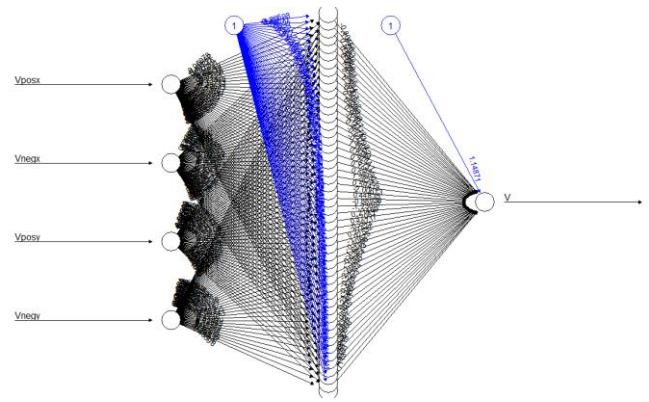


Fig. 3. Illustration of neural network model output of R program

IV. RESULTS

In literature, generally classical numerical methods like finite difference method have been used. However, according to our studies it has been proven that on widespread grounding type which is rod grounding, the potential distribution around the grounding can be found by using machine learning methods.

Table 1 shows ten of test dataset including real voltage values and predictions of both neural network and random forest methods.

TABLE I. INPUT AND OUTPUT DATA OF POTENTIAL DISTRIBUTION

V_{posx}	V_{negx}	V_{posy}	V_{negy}	V_{real}	V_{ann_pred}	V_{rf_pred}
53.46	39.43	0.00	44.78	45.90	28.45	48.37
51.89	38.51	45.90	41.73	44.78	28.96	56.78
47.99	35.89	44.78	37.26	41.73	28.59	41.72
42.65	31.94	41.73	31.85	37.26	27.99	36.95
36.47	27.03	37.26	25.87	31.85	26.89	31.99
29.84	21.54	31.85	19.57	25.87	24.79	25.64
22.96	15.77	25.87	13.11	19.57	20.81	19.67
15.91	9.87	19.57	6.57	13.11	14.03	13.30
8.55	4.27	13.11	0.00	6.57	6.88	6.44
45.90	33.57	0.00	38.51	39.43	28.45	44.21
...

Comparing the output performances, it is showed that the RF model has more precise prediction than the ANN model.

In Table 2, comparison matrix of model performances is given. As it is clearly seen, RF model gives the lowest MAPE, MSE and RMSE values.

TABLE II. COMPARISON MATRIX

METHOD	MAPE	MSE	RMSE
RF	3.996841	8.179371	2.85996
ANN	13.80443	32.15212	5.670284

Reliable forecast of potential distribution is vital for an electrical facility in terms of protecting objects from dangerous voltages.

Unlike other studies, in this study artificial intelligence technics mainly machine learning algorithms have been

applied for investigating the electrical distribution around a grounding system.

REFERENCES

- [1] J. H Bogensperger, J. Frei and S. Pack, "Resistance of Grounding Systems Stationary and Transient Behavior" Ninth International Symposium on High Voltage Engineering , August 28- September 1,1995
- [2] J. Takahashi, T. Kawase, "Calculation of Earth Resistance for a Deep-Driven Rod in a Multi- Layer Structure", IEEE Transactions on Power Delivery, Vol. 6, No. 2, April 1991.
- [3] A. P. S. Meliopoulos, F. Xia, E. B. Joy, G. J. Cokkinides, "An Advanced Computer Model for Grounding System Analysis", IEEE Transactions on Power Delivery, Vol.8, No. 1, April 1993.
- [4] F. Dawalibi, D. Mukhedkar, "Influence of Ground Rods on Grounding Grids", IEEE Transactions on Power Apparatus and Systems, Vol. PAS-98, No.6, Nov./Dec. 1979, pp. 2089-2098.
- [5] H. Yıldırım, Ö. Kalenderli, B. Türkay, M. Çelikyay, "Topraklama ağlarının bilgisayar destekli analizi", Elektrik-Elektronik Mühendisliği 6. Ulusal Kongresi, Bursa, pp. 130-133, 11-17 Eylül 1995
- [6] P. Hasse, Overvoltage Protection of Low Voltage Systems, IEE Power and Energy Series 33, United Kingdom, 2000
- [7] ANSI/IEEE Std 80-1986, IEEE Guide for Safety in AC Substation Grounding, 1986.
- [8] Elektrik Tesislerinde Topraklamalar Yönetmeliği, TMMOB, Elektrik Mühendisleri Odası Bursa Şubesi, Bursa, 2001 (21 Ağustos 2001 tarih ve 24500 sayılı Resmi Gazete'de yayımlanmıştır).
- [9] L. Breiman, 2001, "Random forest", Machine Learning, 45, pp. 5-32.
- [10] İ. Özsoy and M. Fırat , 2004, Kirişsiz döşemeli betonarme bir binada oluşan yatay deplasmanın yapay sinir ağları ile tahmini. *DEÜ Mühendislik Fakültesi Fen ve Mühendislik Dergisi*, 6, 51-63.

Study of Consumer Perception of Recommender Systems in E-Commerce Applications

Çağrı Ulu
Foreign Trade
Kavram Vocational School
İzmir, Turkey
cagri.ulu@kavram.edu.tr

Cansu Ulu
International Relations
Ege University
İzmir, Turkey
cansuulu89@gmail.com

Abstract— Recommender systems (RSs) with artificial intelligence have changed the relationship between business and consumer. RSs with artificial intelligence cause the entrepreneur to understand the consumer's behavior. By understanding the consumer, the business section has better demand management and supply chain. Now it is possible to deploy the product before its order. On the other side, the recommender systems have advantages for consumers like finding the desired product with personalized suggestions and visual research, personalized marketing methods, uninterrupted customer support with chatbots, reliability and speed. Besides its advantages recommender systems have ethical problems. RSs can learn the patterns of consumer's behaviour by processing of personal data and profiling. This creates ethical problems within the scope of protecting individuals' personal data. In order to overcome these ethical problems, General Data Protection Regulation (GDPR) came into force across the European Union on 25 May 2018. Our evaluation focuses two basic user behavioral outcome the intention of use RS and the intention of purchase.

Keywords— Recommender system, artificial intelligence, e-commerce, e-marketing, ethical problem.

I. INTRODUCTION

E-commerce is the newly emerging field of business. While the ratio of e-commerce volume to general trade was 9.8 percent in 2019, it increased by 5.9 points in 2020 to 15.7 percent on average [1]. E-commerce deepens the relationship between business to the customer so it changed the supply chain and demand management. Using artificial intelligence-supported RSs, the item (in our case product) can be delivered to the storage before demand occurs. RSs offer many opportunities for e-commerce but it also has ethical problems. An efficient recommendation system should be accepted as a trusted one by the user (customer). This study aims to evaluate RSs from the perspective of the user.

II. RECOMMENDATION SYSTEMS IN E-COMMERCE

RS in its simplest definition, tools, and techniques providing suggestions for products and services to users. RS is not a unique concept to the internet and e-commerce. We have been witnessing RSs on traditional media tools. Traditional media use non-personalized recommendations. This kind of marketing strategy can only address one group as the smallest target audience. On the other hand

personalized marketing defines a customer as an individual and having a tailored marketing mix designed for that one customer [2]. RSs use different techniques for every user so it helps to create demand and manage the supply chain. With a more detailed definition, RS in e-commerce is a personalized information filtering technology, used to either predict whether a particular user will like a particular item (prediction problem) or to identify a set of N items that will be of interest to a certain user (top-N recommendation problem) [3].

RSs help providers to increase the number of items sold, sell more diverse items, increase user satisfaction, increase user fidelity and better understand what the user wants [4]. A good example of RS in an e-commerce application should convert browsers to buyers [5]. From the perspective of user RSs help to find items related to users' interests, save time spent searching for the relevant product, help to create interaction-driven e-commerce applications, provide decision support technology to the users, optimize supply chain and logistic facilities to faster delivery options.

Different types of RSs are using in e-commerce applications. Content based RS depends on the previous actions and responses to the user. "A content-based recommendation system builds a user profile based upon the previously rated items by the user." [6]. Content based RSs use proper techniques for representing the items and producing the user profile. Generally, a highly structured content based RS has three component; content analyzer to create structured data to extract relevant information, profile learner collects data representative of the user preferences and tries to generalize this data, in order to construct the user profile, and filtering component builds the user profile to suggest relevant items by matching the profile [7]. The second type of RSs is collaborating filtering methods. This type use opinions of other users rather than one specific profile. Collaborating filtering relies on a utility matrix to represent items. "Collaborative filtering uses similarities between users and items simultaneously to provide recommendations." [8]. The third type of RSs is the hybrid model. This model use techniques are combined like content based and collaborative filtering to overcome shortcomings.

Content based RSs have advantages like user interdependence, transparency, and capability of recommending items not yet rated by any user [7]. But also have shortcomings such as limited content analysis,

over-specialization, cold start problem. Collaborating filtering does not need to analyze content because in collaborating filtering the user profile is the set of ratings given to different items [9]. Collaborating filtering does not have cold start problem and it captures the change in user interests over time. The disadvantages of collaborating filtering are problems with fresh items, data sparsity, and scalability [10].

III. THE INTENTION OF USE AND PURCHASE

RSs can predict the purchase behavior of the user and also manipulate it. Traditional supply chain management can not handle the massive interaction data and speed. To improve the effectiveness the business section should comprehend e-commerce as a whole, and smart supply chain management should be used. "The intelligent supply chain system can bring about improvement and optimization through improving the effectiveness of demand management." [11]. RSs help providers to analyze demand with real time-data. Demand forecasting reduces the cost of storage and increases the speed of delivery. RSs can also manipulate demand. Dadouchi and Agard try to shift the demand from one item to another concerning the supply chain [12]. If we can change the demand to item existed in storage, we can reduce the cost and the time.

RSs effectiveness of creating the intention to use relies on sharing data. Content based RS needs a user profile created from the data of previous interactions and collaborating filtering method mimics the user via ratings given to different items. GDPR has strengthened data protection safeguards regarding personal information. GDPR defines personal data to be any information relating to an identifiable natural person [13]. According to GDPR, all identifying data must be protected. RSs use pseudonymised data. Pseudonymization prevent us from identifying data subjects. Content based RSs have advantages as transparency and users' control. We evaluate the awareness of how RSs use data.

IV. METHOD

To evaluate recommendation systems from the perspective of the user we use the survey method. We applied ResQue Model created by the Human-Computer Interaction Group. The model focused on four users' believe; transparency of RS, control over it, perceived usefulness of RS, and perceived ease of use. This user believes cause two user attitudes trust & confidence and overall satisfaction. The outcome of trust & confidence is purchase intention and the outcome of overall satisfaction is use intentions. Using ResQue model we can understand the acceptance degree of the user to share their data with AI in exchange for a good recommendation and the effect of AI-supported RS to create demand. We applied the questionnaire to te students of Kavram Vocational School. The scale of our questionnaire is calculated as 328. Only 72 students participated in the survey because of Covid-19. Therefore, we cannot process the data with the SPSS program. We plan to reach a sufficient number of people and analyze the data with the SPSS program in future studies.

V. RESULTS

73,6% of respondents are women and 56,9% of age under 20. When we analyze users' attitudes we can see that overall satisfaction of RSs are low, only 7,6% of users strongly agreed with the statement "Overall, I am satisfied with the recommender." Confidence and trust attitudes come from the belief of control over RS and transparency. 26,58% of users believe in control and transparency. Only 16% of users trust RSs as an outcome of these beliefs. When we evaluate the intention of use, 25,5% of respondents find RSs useful and likely to use, 47,36% disagree with it. The intention of purchase tested with "The information provided for the recommended items is sufficient for me to make a purchase/download decision." and "I would buy the items recommended, given the opportunity." statements. 22,4% of users find RSs effective for the intention of purchase, 46,4% disagree with the statements.

VI. CONCLUSION

RSs offers possibilities to increase e-commerce volume. Artificial intelligence can help to operate e-commerce as a whole. RSs is not a zero-sum game. Both business section and consumers might win. Consumers (users) can save time in e-commerce and find related items easily. Business increase profit rate by reducing logistic and storage cost. The demand of consumers can be manipulated and predicted by intelligent supply management. We applied a questionnaire to students of Kavram Vocational Scholl. As results, we can see that RSs are not efficient in Turkish e-commerce applications. Users find it hard to use and control RSs, and RSs insufficient to affect purchase behaviors.

REFERENCES

- [1] E-ticaret Bilgi Platformu, 2020 Yılı E-Ticaret Verileri Açıklandı, 2021, <https://www.eticaret.gov.tr/haberler/10040/detay#:~:text=D%C3%BCnyada%20e%2Dticaretin%20y%C4%B1ll%C4%B1k%20bazda%200%20milyon%20liraya%20y%C3%BCkseldi%C4%9Fini%20bildirdi>
- [2] S.K. Dawn, "Personalised marketing: Concepts and framework", *Productivity*, 54 (4), 2014, pp. 370-377.
- [3] P.Bashiri, *Recommender Systems: Survey and Possible Extentions*, unpublished.
- [4] F. Ricci, L. Rokach and B.Shapira, "Introduction to recommender systems handbook" in *Recommender Systems Handbook*, F. Ricci, L. Rokach, B.Shapira and P.B. Kantor , Eds. London: Springer, 2011, pp.1-38.
- [5] P.Rana, N.Jain and U. Mittal, "An introduction to basic concepts on recommender systems", in *Recommender System with Machine Learning and Artificial Intelligence*, S.N. Mohanty, J.M. Chatterjee, S. Jain, A.A. Elngar and P. Gupta, Eds. Beverly: Scievers, 2020, pp. 1-26.
- [6] J.B Schafer, J. Konstan, J. Reidl, "Recommender Systems in E-commerce" *ACM Conference in Electronic Commerce*, November 1999, pp. 158-166.
- [7] P. Lops, M. Gemmis, G. Semeraro, "Content-based recommender systems: State of the art and trends", in *Recommender Systems Handbook*, F. Ricci, L. Rokach, B.Shapira and P.B. Kantor , Eds. London: Springer, 2011, pp. 73-106.
- [8] Collaborative Filtering, 2021, <https://developers.google.com/machine-learning/recommendation/collaborative/basics>.
- [9] F. CACHED, V. Carneiro, D. Fernandez and V. Formoso, "Comparison of collaborative filtering algorithms: Limitations of

- current techniques and proposals for scalable, hig-performance recommender systems”, *ACM Trans., Web5*, 1, Article 2, Februray 2011, pp. 1- 33.
- [10] H. Al-Bashiri, A. Romli, M. A. Abdulgabber and F. Hujainah , “Recommender system: Overview and challenges”, *Advanced Science Letters*, 23(9), 2017, pp. 9045-9049.
- [11] Q. Zhang, “The use of recommender systems in demand management in intelligent supply chain management” Unpublished master’s thesis, 2018, University of Dublin, Dublin.
- [12] C. Dadouchi, B. Agard, “ Recommender Systems as an agility enabler in supply chain management”, *Journal of Intelligent Manufacturing*, 2020, Published online.
- [13] S.Shastri, M. Wasserman, V. Cidambaram, “The seven sins of personal-data processing systems under GDPR”, *USENIX Conference on Hot Topics in Cloud Computing*, July 2019.