1ST INTERNATIONAL DATA SCIENCE & ENGINEERING

PROCEEDINGS & ABSTRACTS BOOK

Karabük University Safranbolu Science and Arts Academy

Symposium on Proceedings&Abstracts Book of the International Data Science & Engineering (IDSES'19)

Editör / Editor
Prof. Dr. Filiz ERSÖZ

Karabük University Publishing, 41
ISBN: 978-605-9554-38-1
http://www.idses.org/ info@idses.org

PREFACE

Dear Colleagues,

On behalf of the Local Organizing Committee I am pleased to welcome our distinguished delegates and guests to the IDSES'19 – 1st International Data Science and Engineering Symposium (IDSES'19) held during May 2-3, 2019 in Karabük, TURKEY. IDSES'19 is organized and sponsored by Karabük University.

The Symposium provides the ideal opportunity to bring together academicians who work in this field, data scientists, data miners, data engineers and researchers who want to improve themselves. The main goal of this event is to provide an international scientific forum for the exchange of new ideas in data science fields. With the increase of global competition and the development of technology, the training of experts in this field gained importance with the studies carried out in the field of data science and engineering.

Data discipline and engineering discipline have emerged to give meaning to data stacks, to analyze data stacks and to transform them into information. The implementation of data science and engineering methods enables administrators to make effective and quick decisions to increase operational efficiency as well as to keep the pulse of the society, employees and institutions.

I would like to thank the program committee members for their support at shaping the Symposium program and the research community for their valuable contributions to the Symposium.

Thank you very much for participating in IDSES'19.

With my warmest regards and respect,

Prof.Dr. Filiz ERSÖZ

Chair of IDSES'19
On behalf of Organizing Committee
May 02-03, 2019 Karabük-TURKEY

Sempozyum Onursal Başkanı / Honorary Head of Symposium

Prof. Dr. Refik Polat, Rector of Karabük University

Sempozyum Başkanı / Chair

Prof. Dr. Filiz ERSÖZ, Director, Institute of Natural and Applied Sciences, Karabük University

Düzenleme Kurulu / Organizing Committee

Prof. Dr. Ahmet OZTÜRK (Uludağ University)

Prof. Dr. Erkan IŞIĞIÇOK (Uludağ University)

Prof. Dr. Mehmet KABAK (Gazi University)

Prof.Dr. Necmi GÜRSAKAL (Fenerbahçe University)

Assoc. Prof. Dr. Canan HAMURKAROĞLU (Karabük University)

Asst. Prof. Dr. Nadi Serhan AYDIN (İstinye University)

Asst. Prof. Dr. Taner ERSÖZ (Karabük University)

Asst. Prof. Dr. Turgut ÖZSEVEN (Tokat Gaziosmanpaşa University)

Bilim Kurulu /Hakemler/ Scientific Committee

Prof.Dr. Ahmet ÖZTÜRK (Uludağ University)

Prof. Dr. Ali GÜNGÖR (Karabuk University)

Prof. Dr. Ali KÖSE (Marmara University)

Prof. Dr. Ali OZDEMIR (Dokuz Eylül University)

Prof. Dr. Ayşe NALLI (Karabük University)

Prof. Dr. Bülent ÖZ (Osmaniye University)

Prof. Dr. Cemalettin KUBAT (Sakarya University)

Prof. Dr. Cevriye GENCER (Gazi University)

Prof. Dr. Dilek ÖNKAL (Bilkent University)

Prof. Dr. Elisabeth PEREIRA (Aveiro University)

Prof. Dr. Emel Kızılkaya AYDOĞAN (Erciyes University)

Prof. Dr. Erdal KILIÇ (Ondokuz Mayıs University)

Prof. Dr. Erkan IŞIĞIÇOK (Uludağ University)

Prof. Dr. Filiz ERSÖZ (Karabük University)

Prof. Dr. Gerhard Wilhelm WEBER (Poznan University)

Prof. Dr. Harun TAŞKIN (Sakarya University)

Prof. Dr. Ineta GEIPELE (Riga University)

Prof. Dr. İhsan YÜKSEL (Kırıkkale University)

Prof. Dr. İsmail Ragıp KARAŞ (Karabük University)

Prof. Dr. José António FILIPE (Lisboa University)

Prof. Dr. Kerim ÇETİNKAYA (Karabük University)

Prof. Dr. Massimo CANALICCHIO (CIA Agricoltori Italiani Umbria)

Prof. Dr. Mehmet KABAK (Gazi University)

Prof. Dr. Mehmet ÖZALP (Karabük University)

Prof. Dr. Metin DAĞDEVİREN (Gazi University)

Prof. Dr. Mücahit COŞKUN (Karabük University)

Prof. Dr. Necmi GÜRSAKAL (Fenerbahçe University)

Prof. Dr. Patrick De CAUSMAECKER (Ke Leuven University)

Prof. Dr. Raif BAYIR (Karabük University)

Prof. Dr. Serpil CULA (Başkent University)

Prof. Dr. Stan URYASEV (University of Florida)

Prof. Dr. Süleyman DÜNDAR (Karabük University)

Prof. Dr. Şenol ALTAN (Gazi University)

Prof. Dr. Şeref SAGIROĞLU (Gazi University)

Prof. Dr. Tatjana TAMBOVCEVA (Riga Technical University)

Prof. Dr. Tülay İlhan HAS (Karadeniz Technical University)

Prof. Dr. Veysi İŞLER (Hasan Kalyoncu University)

Assoc. Prof. Dr. Aslıhan TÜFEKCİ (Gazi University)

Assoc. Prof. Dr. Canan HAMURKAROĞLU (Karabük University)

Assoc. Prof. Dr. Kürşat İŞLEYEN (Gazi University)

Assoc. Prof. Dr. Laura Lotero VELEZ (Universidad Pontificia Bolivariana)

Assoc. Prof. Dr. Ömür TOSUN (Akdeniz University)

Assoc. Prof. Dr. Sadia SAMAR ALI (King Abdul-Aziz University)

Assoc. Prof. Dr. Safiye TURGAY (Sakarya University)

Assoc. Prof. Dr. Semra BORAN (Sakarya University)

Assoc. Prof. Dr. Sezgin IRMAK (Akdeniz University)

Assoc. Prof. Dr. Tolga GENÇ (Marmara University)

Assoc. Prof. Dr. Ufuk COŞKUN (Karabük University)

Assoc. Prof. Dr. Yusuf Tansel İÇ (Başkent University)

Asst. Prof. Dr. Ali Osman ÇIBIKDİKEN (Necmettin Erbakan University)

Asst. Prof. Dr. Cumhur GÜNGÖROĞLU (Karabük University)

Asst. Prof. Dr. Çağrı KOÇ (University of Social Sciences)

Asst. Prof. Dr. Emin Taner ELMAS (İskenderun Technical University)

Asst. Prof. Dr. Hande KÜÇÜKÖNDER (Bartın University)

Asst. Prof. Dr. Linda KAUŠKALE (Riga University)

Asst. Prof. Dr. Murat ALAN (Karabük University)

Asst. Prof. Dr. Mükerrem Bahar BAŞKIR (Bartın University)

Asst. Prof. Dr. Nadi Serhan AYDIN (İstinye University)

Asst. Prof. Dr. Nilesh WARE (Pune University)

Asst. Prof. Dr. Taner ERSÖZ (Karabük University)

Asst. Prof. Dr. Turgut ÖZSEVEN (Tokat Gaziosmanpaşa University)

Contents

Factor Analysis of Distribution Tails: Applications in Finance 1

Identification of Green Supply Chain Management (GSCM) Barriers in the Indian Context 4

How Should Data Science Education Be? 5

Data Analysis and Kansei Engineering 8

Methodology for Building A Security System for Banking Information Resources 11

LiBerated Social Entrepreneur. Using Business Metrics: Migport Refugee Big Data Analytics. With a Note on Ability and Disability 13

Forecasting and Technical Comparison of Inflation in Turkey with Box-Jenkins (ARIMA) Models and Artificial Neural Networks 16

The Importance of Data Mining for Businesses 17

Factors Affecting the Adoption of Social Networks for Academic Purpose in Jordanian Universities 23

An International Framework for a More Sustainable Agriculture: Digital Farming, Transfer of Innovative Knowledge, Training and Certification of Performances 25

Facebook Games Applications 33

A Note on the Examination of Portugal's Hotels Performance - a Proposal for a New Perspective's Approach 44

A Comparative Case Study on Time Series Prediction 53

Optimal PID-like Fuzzy Logic Controller Design for Ball and Beam System 56

Adopting Machine Learning Algorithms for Cloud-Based Application Categorization 57

A New Nonparametric Test For Testing Equality of Locations Against Umbrella Alternatives 60

Smart Agriculture Applications with IoT 64

Teknoloji Kabul Modeli Kullanarak Netflix Platformu Kullanma Maksadının Belirleyicileri 72

Meta-Heuristic Methods Used in Optimization of SVM Learning Parameters 75

Determination Similarities of Basic Financial Indicators of Enterprises Included in the NASDAQ Index Using by Hierarchical Clustering Distance Methods 87

RFM Model for Segmentation in Retail Analytics: A Case Study 90

Siber Tehdit İstihbaratı Alanında Makine Öğrenmesi Algoritmalarının Kullanılması 93

Analysis of Non-Risked Provinces; Unemployment and Traffic Accidents 96

Use of Grid Search in Hyper-Parameter Selection for Time Series Analysis: A Case Study with Ad Mediation Software 97

Marketing and Data Analytics; Increasing Importance of Marketing 107

Data Analytics and Importance in Health Sector 111

Program Development for Cost Calculation in Different Hole Drilling Operations 124

Multi-Objective Optimization of Hard Turning: Non-Dominated Sorting Genetic Algorithm-II Approach 126

CO^2 Emission and Energy Consumption for Different Climate and Building Materials 137

Bitcoin Price Forecasting with Multivariate Long Short Term Memory (LSTM) Deep Learning Method 140

Covering-Based Generalized IF-Rough Set Models For A Selecting HVAC System 141

Determination of Production Defects in Iron and Steel Sector by Data Mining 143

Connected Employee Platform and A Case Study in A Global Company 146

Investigation of the Effects of Normal Distribution or Nonnormal of Data on Machine Capability Analysis 150

Portfolio Selection with the Possibilistic Mean – Variance Model: An Application on the Borsa Istanbul 152

Building Digital Assistant (ChatBot) with SAP Conversational Artificial Intelligence 153

Map Ranking, Map-Reduce and Application in Big Data Analysis 163

Benchmarking of OECD Countries in Views of Value-Added Manufacturing Using DEA 165

Recent Literature of Benchmarking of Countries 166

Benchmarking Analysis of OECD Countries 169

A Study on Method Prediction for a Better Directed Treatment of Warts 175

Estimation of the Demand for the Blood Bank Using Hybrid PCA-ANFIS Method 187

Analysing of Multivariate Processes with Machine Learning Algorithms 189

Comparison of Two Different Social Groups on Twitter with Network Analysis 191

Analysis of Earthquake Awareness in Education By Data Mining 209

Tersanelerde Yalın 6 Sigma ve Uygulanabilirliği 213

Smart and Green Supply Chain Applications in Enterprises 215

Bitcoin Price Prediction by Using Artificial Neural Networks and Time Series 219

Analysis of Presence of Bank Branches According to Settlement in Turkey with Data Mining 227

Artificial Bee Colony Algorithm for Container Loading Problem 229

Image Size Scaling and Feature Transformation Function Application for Image Processing in Machine Learning 231

Determination of Weights With Fuzzy AHP in the Job Evaluation Process 232

Cost Estimation in the Iron and Steel Industry 234

Determination of Socio-Economic Factors Affecting Forest Fires (A Case Study of Forest Regional Directorate of Antalya) 237

Group Acceptance Sampling Plans Based on Time Truncated Life Tests For Compound Weibull-Exponential Distribution 259

Stochastic Approach to a Buffer Stock Problem 260

Prediction of Air Permeability of Denim Fabrics Using Articifial Neural Networks 262

Numerical Investigation of Cutting Forces in Turning of C23000 Brass Alloy 263

Evaluation of Critical Factors in Industry 4.0 Transition Processes by R'WOT Analysis 272

Determination of Variables That Affect the Satisfaction Levels of Visiting Tourists By Logistics Regression Analysis 274

Demand Estimation of Wood Quantity Used in Wood Industries 288

Deep Learning Based Abnormality Detection Application in Enterprise Network Traffic 300

Lean Production in Iron and Steel Production Line 303

A Road Map for The Big Data Implementation in the Judicial System 305

Supplier Selection Using an Intuitionistic Fuzzy Evaluation System 306

Environmental Impact Assessment of Two Alternative Wastewater Neutralization Chemicals in Textile Industry Wastewater Treatment Plant 307

Empati Çıkar Yöntemiyle Akademik Çalışma Gruplarının İncelenmesi 310

Application and Comparison of Biclustering Methods in Detecting Crime Regions 311

Kronik Böbrek Hastalığının Makine Öğrenmesi Teknikleri ile Sınıflandırılması 313

Survey on Dynamic Bayesian Network Software Tools 316

Determination of Factors Affecting Employee Productivity 318

Factor Analysis of Distribution Tails: Applications in Finance

Stan URYASEV

Director of the RMFE Lab in Industrial & Systems Engineering

University of Florida, USA

Abstract: A popular risk measure, Conditional Value-at-Risk (CVaR), is called Expected Shortfall (ES) in financial applications. The paper developed algorithms for implementation of linear regression for estimating CVaR as a function of some factors. Such regression is called CVaR (Superquantile) regression. The main statement of the paper: CVaR linear regression can be reduced to minimizing the Rockafellar Error function with linear programming. The theoretical basis for the analysis is established with the Quadrangle Theory of risk functions. We derived relationships between elements of CVaR Quadrangle and Mixed-Quantile Quadrangle for discrete distributions with equally probable atoms. The Deviation in CVaR Quadrangle is an integral. We presented two equivalent variants of discretization of this integral, which resulted in two sets of parameters for the Mixed-Quantile Quadrangle. For the first set of parameters, the minimization of Error from CVaR Quadrangle is equivalent to the minimization of Rockafellar Error from the Mixed-Quantile Quadrangle. Alternatively, a two-stage procedure based on Decomposition Theorem can be used for CVaR linear regression with both sets of parameters. This procedure is valid because the Deviation in the Mixed-Quantile Quadrangle (called Mixed CVaR Deviation) coincides with the Deviation in CVaR Quadrangle for the both sets of parameters. We illustrated theoretical results with a case study demonstrating the numerical efficiency of the suggested approach. The case study codes, data and results are posted at the website. The case study was done with the Portfolio Safeguard (PSG) optimization package which has precoded Risk, Deviation, and Error functions for the considered Quadrangles.

Keywords: Quantile, VaR, Quadrangle, CVaR, Conditional Value-at-Risk, Expected Shortfall, ES, Superquantile, Deviation, Risk, Error, Regret, Minimization, CVaR Estimation, Regression, Linear Regression, Linear Programming, Portfolio Safeguard, PSG

REFERENCES

[1] Acerbi, C. and D. Tasche (2002): On the Coherence of Expected Shortfall, Journal of Banking and Finance 26, 1487–1503.

[2] Adrian, T., and Brunnermeier, M. (2016): CoVaR. American Economic Review 106 (7), 1705–1741.

[3] Basset G.W. and H-L Chen (2001): Portfolio Style: Return-based Attribution Using Quantile Regression. Empirical Economics 26, 293-305.

[4] Carhart M.M. (1997): On Persistence in Mutual Fund Performance. Journal of Finance 52, 57-82.

[5] Case Study (2014): Style Classification with Quantile Regression. http://www.ise.ufl. edu/uryasev/research/testproblems/financial_engineering/style-classification-with-quantile-regression/

[6] Case Study (2016): Estimation of CVaR through Explanatory Factors with CVaR (Superquantile) Regression.h ttp://www.ise.ufl.edu/uryasev/ research/testproblems/ financial_engineering/ on-implementation-of-cvar-regression/

[6] Huang W.Q. and S. Uryasev (2018). The CoCVaR Approach: Systemic Risk Contribution Measurement. Journal of Risk. V.20(4), April 2018, DOI:10.21314/JOR.2018.383, 75-93.

[7] Koenker, R. and G. Bassett (1978): Regression Quantiles, Econometrica 46, 33–50.

[8] Koenker, R. (2005): Quantile Regression. Cambridge University Press.

[9] Portfolio Safeguard (2018): http://www.aorda.com

[10] Rockafellar, R.T. and S. Uryasev (2000): Optimization of Conditional Value-At-Risk. The Journal of Risk, 2(3), 21-41.

[11] Rockafellar, R. T. and S. Uryasev (2002): Conditional Value-at-Risk for General Loss Distributions, Journal of Banking and Finance 26, 1443-1471.

[12] Rockafellar, R. T., Uryasev, S., and M. Zabarankin (2008): Risk Tuning with Generalized Linear Regression. Mathematics of Operations Research, 33(3), 712–729.

[13] Rockafellar, R. T., and S. Uryasev (2013): The Fundamental Risk Quadrangle in Risk Management, Optimization and Statistical Estimation. Surveys in Operations Research and Management Science, 18, 33–53.

[14] Rockafellar, R.T., Royset, J.O., and S.I. Miranda (2014): Superquantile Regression with Applications to Buffered Reliability, Uncertainty Quantification and Conditional Value-at-Risk. European J. Operations Research, 234, 140-154.

[15] Rockafellar, R.T., and J.O. Royset (2018): Superquantile/CVaR Risk Measures: Second-order Theory. Annals of Operations Research, 262, 3-29.

[16] Sharpe W.F. (1992): Asset Allocation: Management Style and Performance Measurement. Journal of Portfolio Management (Winter), 7-19.

Identification of Green Supply Chain Management (GSCM) Barriers in the Indian Context

Mohd. Azmi Khan[1], Salma Ahmed[2], Sadia Samar Ali[3]

[1,2]Aligarh Muslim University, Department of Business Administration Aligarh, India

[3] King Abdul-Aziz University, Department of Industrial Engineering Jeddah, Saudi Arabia

Abstract: Due to the new trends in global warming, environmental consciousness has become a greater concern among the organizations and industries globally. Green Supply Chain Management (GSCM) has received an increased attention from academia and industries in last few years. Green Supply Chain Management (GSCM) has emerged as an innovative organizational strategy to reduce environmental impacts of supply chain activities as well as efficient usage of energy and material. In today's developing economies, customers are becoming more conscious of the environment and governments are making stringent environmental laws, so the industries need to reduce the environmental impact of their supply chain activities. There are many barriers which directly and indirectly affect the implementation of GSCM in an organization. In this paper twenty barriers have been identified through extensive literature review and expert opinion of academicians. These barriers are found to exist in all organizations irrespective of industry domain. Due to the presence of various barriers, Indian organizations are struggling to implement GSCM in their operations. By removing the barriers, Indian industries can focus on cleaner production by adopting Green Supply Chain Management (GSCM) in their operations.

The Objective of the study was to identify the GSCM barriers in Indian context. The research methodology was perusing literature in GSCM and validating by experts opinion. Literature was perused irrespective of industry domain. The study concludes by narrowing on twenty barriers which play a prominent role in the Indian context.

Keywords: Green Supply Chain Management, Barriers, Indian Context.

How Should Data Science Education Be?

Necmi GÜRSAKAL[1], Fırat Melih YILMAZ[2], Ecem ÖZKAN[2], Deniz OKTAY[2]

[1]Fenerbahçe University, Istanbul, Turkey

[2]Uludağ University, Bursa, Turkey

Abstract: The interest in data science is increasing in recent years. Data science, including mathematics, statistics, big data, machine learning, and deep learning; can be considered as the intersection of statistics, mathematics, and computer science. Although the debate continues about the core area of data science, the subject is a huge hit. Universities have a high demand for data science. They are trying to live up to this demand by opening postgraduate and doctoral programs. Since the subject is a new field, there are significant differences between the programs given by universities in data science. Besides, since the subject is close to statistics, most of the time, data science programs are opened in the statistics departments, and this also causes differences between the programs.

Data science education has to be more project-based since up to now, there is no core knowledge of data science like other sciences. It is probably the hypercorrect choice to learn this job in a university which is intertwined with industry and provides plenty of opportunity for internships. In this article, we will summarize the data science education developments and give curriculum examples from the world at the undergraduate and graduate level. Regarding these examples, every university thinks data science as he wants and the names and the contents of these programs really differs.

Keywords: Data Science, Data Product, Recommendation System.

REFERENCES

[1] S. Gutierrez, Data Scientists at Work, Apress, 2014.

[2] N. Casey, "Can AI fix education? We asked Bill Gates", 2016.[Online]. Available: https://www.theverge.com/2016/4/25/11492102/bill-gates-interview-education-software-artificial-intelligence [Accessed: 13- April- 2019]

[3] Atlantic, "Artificial Intelligence Promises a Personalized Education forAll",2018.[Online]. Available:https://www.theatlantic.com/sponsored/vmware-2017/personalized-education/1667/ [Accessed: 13- April-2019]

[4] S. Lior, "Is Data Science an Academic Discipline?",2017.[Online]. Available: https://www. datasciencecentral.com /profiles /blogs/is-datascience-an-academic-discipline [Accessed: 13- April- 2019].

[5] P. L. Jennifer, "Maslow's Hierarchy of Data Science: Why Math and Science Still Matter", 2019, [Online]. Available: https://www.datasciencecentral.com/profiles/blogs/maslow-s-hierarchy-of-data-science-why-math-and-science-still [Accessed: 13- April- 2019].

[6] E. Colson, "Why Data Science Teams Need Generalists", HBR, 2019, [Online]. Available: https://hbr.org/2019/03/why-data-scienceteams-need-generalists-not-specialists [Accessed: 13- April-2019].

[7] N. Thomas, "How It Feels to Learn Data Science in 2019", Towards Data Science, 2019. [Online]. Available: https://towardsdatascience.com/how-it-feels-to-learn-data-science-in-2019-6ee688498029. [Accessed: 15- April- 2019].

[8] DJ Patil, Building Data Science Teams, O'Reilly Media, 2011.

[9] S. Steve, "Enable Deeper Understanding with Great Data Storytelling",2017.[Online]. Available: https://tdwi.org/articles/2017/03/09/enable-deeper-understanding-with-great-data-storytelling. aspx [Accessed: 15- April- 2019].

[10] S.C. Hicks, and R. A. Irizarry, "A Guide to Teaching Data Science.", American Statistical Association, 2018. Pp.382–91.

[11] A. Stone, "Will Data Scientists Have a Big Impact on Education?", 2017. Online]. Available: https://www.govtech.com/education/k-12/Will-Data-Scientists-Have-a-Big-Impact-on-Education.html. [Accessed: 15- April- 2019].

[12] W. Vorhies, "Getting a Data Science Education", Data Science Central, 2015.[Online]. Available:https://www.datasciencecentral.com/ profiles/ blogs/getting-a-datascience-education [Accessed: 17- April- 2019].

[13] A. Woodie, "Universities Get Creative with Data Science Education", Datanami, 2018.[Online]. Available: https://www.datanami.com/2018/08/23/universities-get-creative- withdata-science-education/ [Accessed: 19- April- 2019].

[14] National Academies of Sciences, Engineering, and Medicine. Data Science for Undergraduates: Opportunities and Options, The National Academies Press. 2018.

[15] Park City Math Institute (PCMI) Undergraduate Faculty Group, "Curriculum Guidelines for Undergraduate Programs in Data Science", Annual Review of Statistics, 2016. pp.1–26

[16] T. Akimichi, "A new era of statistics and data science education in Japanese universities", Japanese Journal of Statistics and Data Science, 2018, pp. 109–116.

[17] Zhang Jilong, Fu Anna, Wang Hao, Yin Shenqing, "The Development of Data Science Education in China from the LIS Perspective", The International Journal of Librarianship,2017, pp. 19–26.

[18] MNM, "20 Universities for pursuing Master of Science in Data Science (On-Campus) in the USA — 2018", Towards Data Science,2018. Online]. Available: https://towardsdatascience.com/20-universities-for-pursuing-master-of-science-in-data-science-oncampus-in-the-usa-2018-9970d5d25bd5 [Accessed: 22- April- 2019]

[19] University of Waterloo, "Department of Statistics and Actuarial Science", 2017-2018. [Online]. Available: https://uwaterloo.ca/statistics-and-actuarial-science/sites/ca.statisticsand-actuarial-science/files/uploads/files/datascience-2017-2018.pdf [Accessed: 22- April- 2019]

[20] Data Science Degree Programs, "30 Best Master's In Data Scıence Degree Programs 2019". 2019. [Online]. Available: https://www.datasciencedegreeprograms.net/rankings/masters-datascience/ [Accessed: 22- April- 2019]

[21] J. Brooks, "Why so many data scientists leaving their jobs", 2018. [Online]. Available: https://www.kdnuggets.com/2018/04/why-datascientists-leaving-jobs.html [Accessed: 22- April- 2019]

Data Analysis and Kansei Engineering

Mustafa Umut ÖZTÜRK, Ahmet ÖZTÜRK

Department of Econometrics, Uludağ University, TURKEY

Abstract: One of the most important steps in establishing a successful business is to do accurate data analysis. The correct analysis of the data and the correct information as a result of the analysis reveal the wishes, feelings, emotions, and demands of the users. With data analysis useful information should be discovered, those who are useless should be cleaned and what should have done in the next process. Several types of data analysis methods can be done based on the data collected from Kansei survey. These analyses play an important role in the process of Kansei Engineering. There are several types of statistical analysis that are developed to use in Kansei studies such as *variance analysis*, *linear regression analysis*, *flow data analysis*, *principal component analysis*, *quantification theory I*, *factor analysis*, *cluster analysis*, *rough set theory*. The purpose of data analysis is to synthesis statistical data or Kansei words with the product properties and therefore to be applied in the design context.

Kansei engineering is a method used to convert a customer's ambiguous imagine product into detail design. Kansei Engineering starts from decision of strategy as design domain as well as target. It is collected the Kansei words related to the product domain. The word Kansei, which is used in design and other research areas. It means the feeling of beauty and pleasant emotions reflected by any object and its desire in Japanese. Kansei words form the basis of Kansei engineering. In a way, Kansei engineering is a product development method which can measure a customer's feeling, values and affective needs and translate them into concrete product solutions.

Since 1980's Kansei Engineering has expanded greatly and become a significant discipline both in the industrial and the academic world. Furthermore, Kansei Engineering developed as an efficient research discipline, providing many innovations and market success in the industrial world. It is founded by Mitsou Nagamachi, a professor at Hiroshima University. He is considered to be the father of Kansei Engineering. The term Kansei Engineering itself was used the first time in 1986 by Yamanota, president of Mazda Automotive Corporation at Michigan University.

The main aim of this study is to explain Kansei engineering and model which is rarely seen in Turkish literature and to reveal its relationship with data analysis. Besides, the future of importance of data, Big data, data analysis and Kansei Engineering will be discussed.

Keywords: Data, Data Analysis, Big Data, Kansei, Chise, Kansei Engineering**.**

REFERENCES

[1] Data Analysis – Wikipedia, https//en. Wikipedia. Org/wiki/Data analysis.

[2] Doğan, K.- Arslantekin, S., Büyük Veri: Önemi, Yapısı ve Günümüzdeki Durum (Big Data: Its Importance, Structure and Current Status), Dil Tarih Coğrafya Dergisi, 56.1, 2016, s.15.

[3] Erlag, H.- Kör, H., Büyük Veri ve Büyük Verinin Analizi (Big Data and Its Analysis), International Conference on Science and Technology, October 3-6, 2016, Ankara.

[4] Harari, Yuval Noah, 21 Yüzyıl için 21 Ders, (Türkçesi, Selin Seral, 21 Lessons fort he 21st Century) Kollektif Kitap, İstanbul,2018, s.84.

[5] Le'vy, Pierre," Beyond Kansei Engineering: The Emancipation of Kansei Design", International Journal of Design Vol.7, No:2, 2013, p.83.

[6] Lokman, Anitawati Mohd- Nagamachi, Mitsuo," Validation of Kansei Engineering: Adoption in E- Commerce Web Design", Kansei Engineering Internatonal Vol.9 No:1,2009, pp.24.

[7] Lokman, Anitawati Mohd," Desing and Emotion: The Kansei Engineering Methodology", Design: the Kansei Engineering Methodology, Vol.1, Issue1, 2010, pp.1, https:// aniwati.uitm. edu.my/mypapers/21MJOC10-Designand Emotion-The KE methedology.pdf.

[8] Lokman, Anitawati Mohd," Design and Emotion: The Kansei Engineering Methedology", Vol.1, Issue, 2010, pp.3. https://Aniwati-uitm.edu.my/mypapers/21-MJOC 10 -Design and Emotion- KE Methodology.Pdf.

[9] Lokman, Anitawati Mohd- Mitsuo Nagamachi," Validation of Kansei Engineering Adoptatation in E- commerce Web Design", Kansei Engineering International Vol.9 No.1, 2009, pp.24.

[10] Marco, Lluis- Tort, Almogro Xavier - Llabre's, Martorell, "Statistical Methods in Kansei Engineering: a Case of of Statistical Engineering", ENBIS 11, September 2011, p.2.

[11] Nagamachi, Mitsuo, Home Applications of Kansei Engineering in Japon: An Overview, 2016:15.4, p.209, https//www. Research gate.net/publication/ 311707844, Home application of Kansei Engineering in Japon.

[12] Okamoto, Ricordo Hirata, Nagamachi, Mitsou, Ishihara, Shigekazu," Satisfing Emotional Needs of the Beer consumer Trough Kansei Engineering", https://www.research gate.net//publication/314041244.Satisfying, emotional needs of the beer –consumer through Kansei Engineering,2004, pp.1-8.,

[13] Öztürk, Ahmet, Kalite Yönetimi ve Planlaması (Quality Management and Planning), 2. Baskı Ekin Kitabevi, Bursa,2013, p.4.

[14] Seung Hee Lee, Akira Harada-Pieter Jan Stappers, Pleasure with Products: Design based on Kansei. https://www.researchgate.net/publication/228396068-Pleasure-with-product-Design-based-on-Kansei.

[15] Shaari, Nazlina, Methods of Analzing Images based on Kansei Engineering, International Journal of Computer Science and Electronic Engineering (IJCSEE) Volume 1, Issue3, s.17.

[16] Shafieyoun, Zhabis- Maicocchi, Marco, Flow Kansei Engineering Qualifying conscious and unconscious behaviour to gain optimal experience in Kansei engineering, International Conference on Kansei Engineering and Emotion Research, KEER 2014, Linköping, June 11-13, 2014, pp.621.

[17] Schütte, Simon, Designing Feeling into Products Integrating Kansei, Engineering Methodology in Product Development, Thesis No.946, Institute of Technology, Linköping, 2002, p.23.

[18] Schütte, Simon, Engineering Emotional Values in Product Design: Kansei Engineering in Development, Linköpin University, Linköpin,2005, pp.45. https://pdfs. Semantic scholar.org/52ao/5 fed.774a1.

[19] Sridhar, Jay, What is the Data Analysis and Why is it important?, Feb.12,2018,https://make useofcom/tag/what-is-data-analysis/

Methodology for Building A Security System for Banking Information Resources

Serhii Yevseiev

Simon Kuznets Kharkiv National University of Economics, UKRAINE

Abstract: The revolutionary changes of the last decade in the banking sector have led to the integration of information and computer networks into a single information and cybernetic space, which has led to the creation of automated banking systems that have substantially expanded the spectrum of electronic services of state and commercial banks of the world. As a result, threats to such a national information resource of the state as the banking information resources under which the banking information refers. Threats to the security of banking information resources have become signs of hybridization. Manifestations of hybridity as a result of the simultaneous impact of threats to information security, cybernetic security and information security on banking information resources have led to the emergence of synergies, the negative manifestations of which require a radical revision of the concepts of the construction of existing security systems. Thus, it becomes clear that there is a need for a radical revision of the current methodological principles for building a security system for banking information resources both for Ukraine and for the world as a whole.

Keywords: Security of Banking Information Resources, Automated Banking System, A Synergetic Model of Threats to the Security of Banking Information Resources, A Classifier of Threats to the Security of Banking Information Resources, Information Security, Cybersecurity, Security to Information, Investments, Emergent Properties, Synergistic Effect, Hybrid Crypto Codes on Damaged Codes, Elliptic Codes, Methodology.

REFERENCES

[1] Hryshchuk R., Yevseiev, S. Shmatko A. Construction Methodology of Information Security System of Banking Information in Automated Banking Systems: Monograph, 284 P., Vienna.: Premier Publishing S. R. O., 2018.

[2] Yudin, O., "Informacion Of Security. Regulatory and Legal Regulations", 640p., NAU, Kiyev, 2011. 3. Grischuk, R., Danik, Yu., "The Foundations of CyberSec", ZhNAEU, Zhytomyr, 636 p., 2016.

[3] Evseev, S., Korol, O., Kots, G., "Analysis of the Legislative Framework for the NSMEP Information Security Management System", East European Journal of Advanced Technologies, v.77, no. 5/3, p. 48–59, 2015.

[4] Evseev, S., "Methodology for Evaluating the Security of Information Technologies of Automated Banking Systems of Ukraine", Science and Technology Journal "Information Security", v. 22, no. 2, p. 297-309, 2016.

[5] Briones, P., Chamoso, A., Barriuso, A., "Review of the Main Security Problems with Multi-Agent Systems used in E-commerce Applications", ADCAIJ, Regular Issue, v. 5, no. 3, pp. 55-61, 2016.

[6] Simpson, W., "Securing Information Systems in an Uncertain World Enterprise Level Security", Systemics, Cybernetics and Informatics, v. 14, no. 2, pp. 83-90, 2016.

[7] Evseev, S., Korol, O., Rzayev, H., Imanova, Z., "Development of a Mac-Alice Modified Asymmetric Crypto-Code System on Shortened Elliptic Codes", East European Journal of Advanced Technologies, v. 4, no. 82, p. 18-26, 2016.

[8] Evseev, S., Korol, O., Kots, G., "Construction of the hybrid systems", East-European Journal of Advanced Technologies, v. 4, no. 88, p. 4-21, 2017.

[10] Evseev, S., "The Use of Flawed Codes in Crypto-Code Systems", Information Processing Systems, V. 151, no. 5, 2017.

[11] Scheau, C., Arsene, A., Dinca, G., "Phishing and E-Commerce: An Information Security Management Problem", Journal of Defence Resources Management, v.7, no. 1 (12), pp. 129-140, 2016.

[12] Alhothaily, Ab., Alrawais, A., Song, T., Lin, B., Cheng, X., "QuickCash: Secure Transfer Payment Systems", Sensors, no. 17, pp.1-20, 2017.

[13] Yusupova, O., "Transaction Security When Using Internet Banking", Financial Analytics: Problems and Solutions, no. 35, p. 26–40, 2016.

[14] Evseev, S., Kots, G., Lekarev, E., "Developing A Multi-Factor Authentication Method Based on The Crypto-Code System", East European Journal of Advanced Technologies, v. 6/4, no. 84, p. 11-23, 2016.

[15] Evseev, S., Korol, O., Kots, G., Minukhin, S., Kholodkova, A., "Eastern European Journal of Advanced Technologies, v. 5 / 9, no. 89, p. 19–36, 2017.

LiBerated Social Entrepreneur. Using Business Metrics: Migport Refugee Big Data Analytics. With a Note on Ability and Disability

Gerhard-Wilhelm WEBER[1], Berat KJAMILI[2], Dominik CZERKAWSKI[1]

Poznan University of Technology[1], POLAND; Middle East Technical University[2], TURKEY

Abstract: LiBerated Social Entrepreneurship in Developing and Emerging Countries consists of a social entrepreneur using business metrics, to sustain social impact. We study differences between developing and developed countries, introducing a new OR approach to development. Commercial entrepreneurs are generally oriented to business metrics like profit, revenues and return. Instead, social entrepreneurs are non-profits or a blend with for-profit goals, generating Return to Society. In DCs, a social entrepreneurship has been uncommon. We introduce a mid-way as LiBerated Social Entrepreneur, where social businesses should be sustainable. We apply Game and Max-Flow - Min-Cut Theories, Schumpeter's creative destruction and Adam Smith's diversification model for our business plan. As a result, B. Kjamili started Migport, formerly Q-Zenobia: a mobile application that runs as a "refugee portal", supported by "Refugee Big-Data Analytics": refugees submit data to the application via "questionnaire" and search for opportunities, verified news privatized based on their answers. The idea of both-sided help with benefit generated by D. Czerkawski is an extension of B. Kjamili's conception. Nshareplatform (NSP) will create a friendly public space for people with disabilities, understanding they needs. It tries to facilitate better communication between "two worlds"- Ability and Disability and personalizes an assistant (Special person helping people with disabilities). Multivariate Adaptive Regression Splines (MARS), Conic MARS (CMARS) and its robust version RCMARS have shown their potential for Big-Data and, recently, Small-Data. With that toolbox, we aim to further support our joint and novel project.

Keywords: Social Entrepreneur, Start-up, Business Canvas Model, Ability, Disability, OR, Data mining, Analytics

REFERENCES

[1] Ankara Department Agency (2017). TechAnkara Proje Pazarı 2017'de sergilenecek 100 proje belirlendi. Retrieved from: http://www.ankaraka.org.tr/tr/techankara-projepazari-2017de-sergilenecek-100-proje-belirlendi_3810.htm

[2] B. Kjamili and G.-W. Weber, The Role of LiBerated Social Entrepreneur in Developing Countries: A mid-way, in Societal Complexity, Data Mining and Gaming; State-of-the-Art 2017, Greenhill & Waterfront, Europe: Amsterdam, The Netherlands; Guilford, UK North-America: Montreal, Canada, 2017. ISBN /EAN 978-90-77171-54-7.

[3] B. Kjamili, G. W. Weber (2014). IFORS: Opening Doors to International Students, Retrieve from: http://ifors.org/newsletter/ifors-dec-2014.pdf.

[4] Brooks A. C. (2009). Social Entrepreneurship: A Modern Approach to Social Value Creation (1st ed.). New Jersey, Pearson Education.

[5] Business Dictionary, Inc (2017). Definition of Entrepreneurship Retrieved from http://www.businessdictionary.com/definition/entrepreneurship.html.

[6] C. Schinckus (2015). The Valuation of Social Impact Bonds: An Introductory Perspective with Peterborough SIB. Elsevier B.V.

[7] M. Young, The Technical Writer's Handbook. Mill Valley, CA: University Science, 1989.

[8] Egirisim (2017). Startup Istambul 2017'de İkinci Aşamaya Geçen 50 Girişim. Retrieved from: https://egirisim.com/2017/10/21/startupistanbul-2017de-ikinci-asamaya-gecen-50-girisim/

[9] Empowerment Plan, Inc (2017) Story. Retrieved from http://www.empowermentplan.org/about.

[10] Entrepreneurs' Organization (2018). Qzenobia To Run For Finals İn MIT Enterprise Forum Before GSEA| Retrieved from: https://www.eonetwork.org/turkey/chapterpressreleases/qzenobia-torun-for-finals-in-mit-enterprise-forum-before-gsea.

[11] G. Eusden (2013). Max-Flow, Min-Cut: History and Concepts Behind the Max-Flow, Min-Cut Theorem in Graph Theory. Retrieved from http://web.williams.edu/ Mathematics/sjmiller/public_html/hudson/Eusden_maxflowminut.pdf.

[12] MIT Enterprise Forum (2018). Innovate for Refugees Semifinalists. Retrieved from: https://innovateforrefugees.mitefarab.org/en/site/semifinalists.

[13] Moblobi (2017). Türkiyedeki Mültecilerin Oluşturduğu Pazar Hakkında Güvenilir Veriler Sağlayan Girişim QZenobia ile Röportajımız! Retrieved from: http://moblobi.com/roportajlar/qzenobiaroportaj.html

[14] Nacional Geographic (2006). Nobel Peace Prize Goes to Micro-Loan Pioneers, Retrieved from http://news.nationalgeographic.com/news/2006/10/061013-nobelpeace.html

[15] ODTU Teknokent (2017). TechAnkara Proje Pazarı 2017'den ODTÜ Teknokent Firmaları Ödülle Döndü. Retrieved from: http://odtuteknokent.com.tr/tr/haber/techankaraproje-pazari-2017den-odtu-teknokent-firmalari-odulle-dondu

[16] Osterwalder A., Pigneur Y. (2010). Business Model Generation: A Handbook for Visionaries Game Changers, and Challengers, Paperback (1st ed.). New Jersey, John Wiley & Sons.

[17] P. F. Cuevas (2017). The World Bank's mission: Eradicate Poverty and Boost Shared Prosperity. Ankara, World Bank group Poverty.

[18] Strategyzer, Inc (2017). The Business Model Canvas Website: https://strategyzer.com/canvas/business-model-canvas

[19] Tech.co, Inc (2015). Top 15 Most Entrepreneurial Countries in the World Website: http://tech.co/top-15-entrepreneurial-countries-world2015-06

[20] Acumen (2010). An Introduction to Human-Centered Design, Retrieved from https://www.plusacumen.org/courses/introduction-human-centereddesign

[21] Liviing Whith Disability (2016). The New York Times, Retrieved from https://www.nytimes.com/2016/08/28/opinion/living-withdisability.html

[22] Kaul I., Isabelle Grunberg, Marc A. Stern (1999) Global Public Goods: International. Cooperation in the 21st Century, United Nations Development Program, New York.

[23] Mackelprang, R. and Salsgiver, R. O. (1999) Disability: A Diversity Model Approach in Human. Service Practice, Brooks/Cole Publishing Company, Toronto.

Forecasting and Technical Comparison of Inflation in Turkey with Box-Jenkins (ARIMA) Models and Artificial Neural Networks

Erkan IŞIĞIÇOK, Ramazan ÖZ, Savaş TARKUN

Department of Econometrics, Uludağ University, TURKEY

Abstract: Inflation refers to an ongoing and overall comprehensive increase in the overall level of goods and services price in the economy. Today; inflation, which is tried to be kept under control by the central banks, is trying to ensure price stability, the continuous price changes that arise in all the goods or services that consumers use includes. Undoubtedly in terms of economy, inflation expectations are also ganing importance, except for rhe realized inflation. This situation makes it necessary to predict the future vaules of inflation. In that case, a reliable estimate of the future values of inflation in any country will create an entry in determining the policies that decision-maker units will implement on the economy.

The aim of this article is to predict inflation in the next period by using the Consumer Price Index (CPI) data with two alternative techniques. It is also aimed to examine the prediction performances of these two techniques in comparisons. Thus, the first of the two main objectives of the study is to predict the future values of inflation with two alternative techniques. The second goal is to determine which of these two techniques well compared to statistical and econometric criteria.

In this context, the estimated performance of both techniques was predicted by the 9-month inflation, Box-Jenkins (ARIMA) and Artificial Neural Networks (ANN) in the April – December 2019 period, using CPI data consisting of 207 in the period of January 2002 – March 2019. In the study, Eviews and Matlab programs were utilized.

Keywords: Inflation, Box-JEnkins, ARIMA, Artificial Neural Networks, Prediction (Forecasting), Technical Comparison.

The Importance of Data Mining for Businesses

Filiz ERSÖZ

Department of Industrial Engineering, Karabük University, TURKEY

Abstract: Today, with digitalization, it is possible to read digital data and make the right decisions based on analytical results. Along with big data, the science of data management and analysis is evolving to enable organizations to transform their knowledge into information that will help them achieve their goals. In this study, it is given as an example to increase awareness of big data, data mining, data mining and its applications in various sectors in Turkey.

Keywords: Industry 4.0, Digitalization, Big Data, Data Mining.

Introduction

The importance of data has started all over the world and with the increase of big data every day, data processing has become more important. The data revolution, affecting the entire world and all sectors, has attracted attention with increasing technology and advances in machine learning. In addition, the increasing number of structured and unstructured large data owned by enterprises in recent years has led to the need to make sense of these data. Systems that support data volumes along with big data continued to increase rapidly. Microsoft is involved in business analytics, data science and machine learning, big data systems and platforms or data management.

Big companies are now making great use of technology and digitalization to solve their problems and make predictions for the future. Using big data and management, these companies make sense and analyze data based on digital and artificial intelligence and shed light on future planning. However, SMEs hesitate to adopt data science and these technologies. SMEs do not have enough resources for artificial intelligence or digital applications and analytics. In addition, analytical approaches are approached with suspicion.

Researchers have revealed the term "big data" and data mining (Data analytics) to describe this evolving technology. In other words, it can be defined as a repository of information that

provide predictive results that will solve the business problem or develop the strategies of the business and enable the business to make good decisions. The data size is still Zettabyte size [1] and this data is processed and guides the enterprises.

Data mining, data analytics, and big data are essentially data science-related sciences. Today, it is used in many places in the same function and in the same sense. Big data and analytics have increased rapidly in the service and manufacturing sectors. Data storage technology, data processing technology, data visualization technique, models and algorithms, and in particular the creation of the right decision-making models, offers opportunities for service and manufacturing enterprises [2].

Data science is said to be an industry with a market value of more than $ 9 trillion by 2020, and not only the ability to extract relevant information, but also the ability to make accurate estimates and significantly improve strategies and performance in the industry [3]. The fields associated with data Science (analytics) are given in Figure 1. Data science is said to be an industry with a market value of more than $ 9 trillion by 2020, and not only the ability to extract relevant information, but also the ability to make accurate estimates and significantly improve strategies and performance in the industry [3]. Scientific methods, visualization, statistical modeling and calculation, data technology, data consultancy.

Firstly, Big data & technologies and analytics are used extensively in many sectors such as health, management, banking and finance, manufacturing, insurance, electronic commerce, communication, transportation, defense, fraud detection and education.

Today, organizations are confronted with complex databases due to the development of technology, the increase of databases and information technologies, and the widespread use of information technologies. It is clear that if the data in accordance with the needs of the institutions is managed successfully and effectively, it will offer great advantages and opportunities to institutions and organizations in economic terms. Each action performed in a digital environment leaves behind a data record. In fact, with each step taken and every choice made, new data is created. According to IDC (International Data Corporation) Big Data and Business Analytics Forum data, the registered data volume increased to 16 ZB (1 ZB 1.09 Trillion Gigabytes) in 2016, it is estimated that the data record will be 35 Zettabyte in 2020 and 163 Zettabyte in 2025 (1024 ZB = 1 Yottabyte (YB)).

In recent years, investment projects for transformational information technologies for different technologies in the financial sector, energy sector, healthcare sector, telecom sector and public institutions have started to increase rapidly. These technologies need to be structured in data centers to meet their corporate goals, increase customer satisfaction and respond to business needs. Rapid walk in projects such as national Data center or city hospitals in Turkey, our transition to 4.5 G, continued rapid growth of cloud computing, information technologies in the formation of future data centers, will further increase the importance of big data and data mining (Business Analytics). In addition, with the increasing complexity of the decision-making process and the need for more numerical and textual data, it became difficult to reveal valuable and meaningful information in big data bases. Amounts of data's being in very large quantity (algorithms for processing data such as Zettabyte, Petabyte, Terabyte must be highly measurable), high size of data (tens of thousands can be micro-arrays), data's being very complex, presence of new and advanced applications requires data mining and text mining.

The data collected in the data stacks stored in databases and data warehouses is now very large. The need to uncover meaningful relationships, patterns and trends from big data stacks has increased the importance of processing data in making accurate and strategic decisions. For these reasons, in data mining and text mining application studies, the value of "machine learning" techniques and "statistical analysis and modelling" has also increased in parallel.

Data mining is the process of discovering the rules and patterns associated with each other from big data stacks. It is not just a technique; it is a data approach that accommodates many techniques. Converts all information from data stacks to an easy and understandable structure. Data mining is related to both database techniques and machine learning. Information from data is the extraction of valuable information in short, data mining can be defined as "the way to convert data to qualified information".

Today, data mining is also referred to as "business intelligence" and "Business Analytics" as well. Other definitions are; Knowledge mining from the databases, Knowledge extraction, Data / pattern analysis, Data archeology and Data analytics.

Data mining; It is an interdisciplinary study where machine learning, statistics, database technology, artificial intelligence and visualization are used together. The most important of these

fields is “the science of statistics”. Statistical methods are the basis for the data mining tools and methods that are being used today.

In data mining; with a study aimed at achieving specific results from large and meaningless masses of data; The data is passed through several stages prior to modeling. The first step is to clean the data before modeling. With the determination of outliers and extreme values, after clean and quality data is obtained (Cleaning), combining the data enable to be able to speak the same language. Here, the relevant and important variables for the research topic are selected and size reduction is performed. As a result of the analysis, the transformation of the available data into a format suitable for reuse, evaluation of the importance of data and relations (Evaluation) and presentation of the results to the decision makers (Presentation) are the processes that complete the data mining. Stages of data mining; Starting from the database, the transformation of data into information is given below [4].

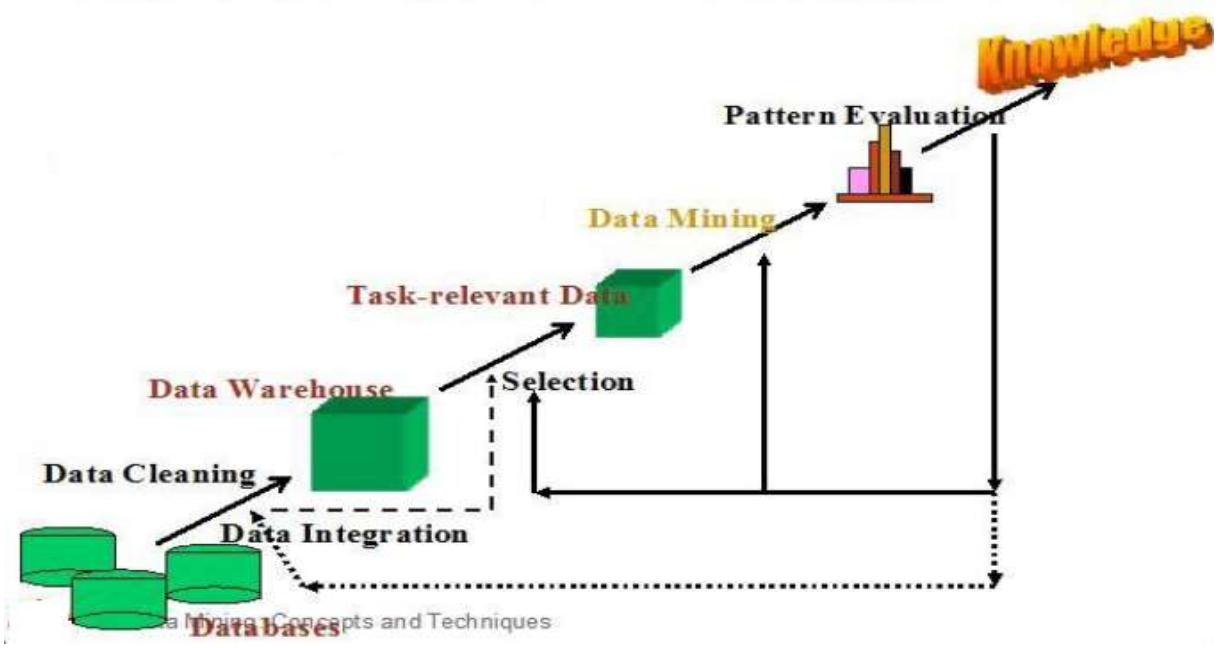


Figure 2. Process of Acces to Information

The data mining cycle is completed by withdrawing information from databases and the results of the analysis with the interpretation of the decision maker. The Data Mining process can be expressed as a step of the process called “Knowledge Discovery in Databases” and “Decision Support System”.

In data mining, business information must be used together with advanced information technologies for the disclosure of information within the database.

Conclusion

Data mining contributes to all sectors. Today, data mining, data analytics and data science, together with concepts such as, but will play an important role in the decision making and roadmaps of production or service businesses. The benefits of data mining to enterprises are briefly given below.

Data mining contributes to all sectors. Today, the concept of data mining is associated with concepts such as data analytics and data science. These concepts will play an important role in the correct decision making and roadmaps of production or service enterprises in the future.

Data mining in enterprises;

- It reveals the valuable information in enterprises by understanding the general computations and probability principles underlying the data in big data stacks, modern machine learning and mining algorithms.
- Analyzes data for scientific and business Analytics with the implementation of many computational and statistical methods.
- Enables the determination and resolution of the appropriate method for the collection and use of data in enterprises.
- By implementing machine learning (software programs for private data mining) solutions enable enterprises to discover new and efficient information.
- Helps the decision maker to make a good decision and to report the information in the enterprises in a clear and understandable way.
- By using data mining and technologies in enterprises that have a big data bound, they solve their problems, reveal their needs, plan and develop their strategies for the future.
- Production or service businesses bring out the need not only to gain customers but also to develop long-standing relationships by optimizing the experiences of their customers.
- It provides efficiency and efficiency in meeting customer expectations through analytical applications in designing, controlling and optimizing business operations in the production of goods or services.

- It helps businesses quickly identify fraud by improving their data and analytical capabilities. Provides continuous monitoring of activities based on their prediction and determination of future activities.
- With forecasting analysis, it can reduce the business risk or out-of-service risk of businesses. In particular, retail and service-based businesses can use predictive analytics to better understand the success of new products or with whom they do business.

REFERENCES

[1] Prof. Dr. Necmi Gürsakal, Büyük Veri. Bursa: Dora Yayınları, 2017.

[2] Ray Y. Zhong, Stephen T. Newman, George Q. Huang, "Big Data for Supply Chain Management in the Service and Manufacturing Sectors: Challenges, Opportunities, and Future Perspectives," Comput. Ind. Eng., vol. 101, pp. 572–591, 2016.

[3] Victor Roman, "4.0 Industry Technologies & Supply Chain" [Online]. Available: https://towardsdatascience.com/4-0-industry-technologies-supply-chain-97c857de14ae.

[4] Han, J. and Kamber, M. (2006) Data Mining Concepts and Techniques. 2nd Edition, Morgan Kaufmann Publishers, San Francisco.

Factors Affecting the Adoption of Social Networks for Academic Purpose in Jordanian Universities

Ala'a Abu Gharrah GHARRAH, Ali ALJAAFREH, Noor AL-MA'AITAH
School of Business, Mutah University, JORDAN

Abstract: Due to the rapid revolution in information technology (IT), teaching methods differ from those used in the past. In recent years, Social Networks (SNs) have become very popular among people. SNs such as Facebook and Twitter can be used in the learning process to stimulate thoughtful discussions on specific classroom topics, and to share learning resources. Despite the Universities have their own eLearning platforms; students are using SNs for the same purpose. The current study attempts to explore the factors affect students' adoption to use SNs for academic purposes.

Keywords: Social Network, Higher Education, Jordan

REFERENCES

[1] Balakrishnan, V. (2017). Key Determinants For İntention to Use Social Media For Learning İn Higher Education İnstitutions. Universal Access İn The Information Society, 16(2), 289-301.

[2] Tan, M., Shao, P., & Yu, P. (2014). Factors Influencing Engineering Students' Use of Social Media In Learning. World Trans. Eng. Technol. Educ, 12(4), 648-654.

[3] Boyd, D. M., & Ellison, N. B. (2007). Social Network Sites: Definition, History, And Scholarship. Journal Of Computer-Mediated Communication, 13(1), 210-230.

[4] Kemp, S. (2018). Digital İn 2018: World's Internet Users Pass The 4 Billion Mark. Retrieved From Https://Wearesocial.Com/Blog/2018/01/Global-Digital-Report2018

[5] Seely Brown, J., & Adler, R. (2008). Open Education, The Long Tail, And Learning 2.0. Educause Review, 43(1), 16-20.

[6] Kaplan, A. M., & Haenlein, M. (2012). Social Media: Back to The Roots And Back To The Future. Journal Of Systems And Information Technology, 14(2), 101-104.

[7] lStats, S. C. G. Browser Market Share Worldwide (Mar 2018 – Mar 2019). Available from: http://gs.statcounter.com/social-mediastats/all/jordan

[8] Selwyn, N. (2009). Faceworking: Exploring Students' Education-Related Use of Facebook. Learning, Media And Technology, 34(2), 157-174.

[9] Tapscott, D., & Williams, A. D. (2010). Innovating the 21st-Century University: It's Time. Educause Review, 45(1), 16-29.

[10] Malkawi, N. M., & Halasa, A. (2016). Exploiting Electronic Social Networks in Educational Process: Study at Universities in Irbid State-Jordan.

[11] Hudson, S., & Thal, K. (2013). The Impact of Social Media on The Consumer Decision Process: Implications For Tourism Marketing. Journal of Travel & Tourism Marketing, 30(1-2), 156-160.

[12] Li, W., & Darban, A. (2012). The Impact of Online Social Networks on Consumers' Purchasing Decision: The Study Of Food Retailers.

[13] Bonilla Polo, P. A., & Osman, M. (2017). Is Social The New Smart? Factors Influencing Students on Their Decision to Use Social Media For Academic Purposes.

An International Framework for a More Sustainable Agriculture: Digital Farming, Transfer of Innovative Knowledge, Training and Certification of Performances

Massimo CANALICCHIO

CIA Agricoltori Italiani Umbria, ITALY

Abstract: Digital farming as we see it has the potential to revolutionize agriculture, and bring significant benefits for farmers and the society overall, as we need new ways to grow more food more sustainably. In this study, Sustainable Farming and Digital Age, Certification as a Model for a Sound Precision Agriculture and Impacts of Precision Agriculture and Certification Needs are described.

Keywords: Digital farming, Sustainable Farming, Certification

Introduction

Smallholder agriculture still dominates the European rural economy, with 86% of EU farms holding an area below 20ha (CEMA, November 2017). Advanced agricultural machinery solutions can help farm holdings, regardless of their size, to operate in a profitable, competitive and sustainable manner. In particular, Precision Agriculture (PA) technologies holds great potential for farmers in this regard. However, available economic evidence shows that there is a strong link between the size of a farm holding and its income, with larger farms tending to have higher income and investment capacity. Precision agriculture, a farming management concept based upon observing, measuring and responding to inter and intra-field variability in crops, or to some aspects of animal rearing, is a new frontier based on use of innovative technologies, such as Global Navigation Satellite Systems (GNSS) and aiming at a rational adoption of decisions and planned agricultural works well balanced between competitiveness and sustainability.

This concept has been simply explained as a way to apply the right treatment in the right place at the right time". As a consequence the world of agricultural machineries is quickly changing with the evolution of modern, competitive and sustainable farming. If this is the scenario for the mechanization of agriculture in the 3rd millennium, the principle to use innovative and

low-cost technologies to improve farming sustainability and competitiveness is a mandatory choice for all farmers. That is why it is necessary to plan efforts and resources to train the majority of farmers and more specifically all the youngsters, on the correct use of farming machinery, the most advanced as well as the less modern but still at work machineries.

The challenge is therefore to combine Open Education Resources (OER) with user friendly quality training materials available online. There is an increasing demand of online courses targeted to work updated competences, as in the case of smart farming can be Internet of Things (IoT) connected to on field good practices. These kinds of training needs are particularly requested for use, management and controls of agricultural machinery and equipment, also not depending of kind of agriculture, organic or conventional, due to a growing awareness to keep safe work place, natural environment and food.

Correct and sustainable use of fertilizers and pesticides is one of the most important farming issues, and training is fundamental to avoid risks and limit pollution as much as possible. The most effective way to spread knowledge and competence on correct use of agricultural machinery implemented with digital technologies is e-learning, since internet gives the opportunity to provide high quality advice to a large number of users.

Sustainable Farming and Digital Age

The digital age in which we live needs responses consistent with global challenges and web opportunities implementing effective quality teaching materials with existing online training and education system such as Massive Open Online Courses (MOOC). Youngsters are very skilled with internet technologies and they are the main target users, but it is also useful to combine well developed e-learning materials, self-paced and also facilitated/instructor led, with work sharing, work-shadowing and internships with the contemporary advice of experts and adult farmers. This approach addresses therefore most of farmers who can be involved in e-learning and combined training.

A tutorial is also needed to explain how to use an online course and how the interactive learning platform can support course delivery and even communication among participants, through forums and social networks. Nevertheless there is a lack of structured training platforms and interactive teaching materials to get adequate competences from a basic use and maintenance up to more detailed skills, based on the European Qualifications Framework (EQF), for farmers,

students, technicians and workers also to induce good pratices, information and training on safe and sustainable use of digital agricultural machinery. Concretely, it means applying new technologies such as data science, advanced sensors in the field and flying drones, digital communication channels, and automation on the field. This way more and more farmers have access to better insights to take more optimal decisions, drive up yield, reduce using pesticide etc.

Current state of researches indicates that to reduce significantly diffuse pollution during pesticides application, the emphasis must be placed on tools and methods for agricultural professionals.

Three key aspects are involved:

• to optimize the agronomic decisions

• to control the precision of the applications

• to record the work performed.

All recent studies show a significant potential for reducing pesticide use: the emphasis should be placed on tools and methods to optimize decisions concerning the use of pesticides and the quality of applications. Guidance systems will be particularly explained as drivers for Precision Agriculture linking farming with technological competences. They can be used by all kinds of equipment (e.g. tractors, combine-harvesters, sprayers, planters…) and as part of a broad range of different agricultural applications.

Guidance systems focus on precise positioning and movement of the machine with the support of a Global Navigation Satellite System (GNSS).

Guidance Systems (GS) enable:

• Field digitalisation
• Automatic steering
• Precise machine movement between plant rows
• Precision drilling and sowing
• Precision spraying
• Mechanical weeding

Certification as a Model for a Sound Precision Agriculture

A certification system based on measurement of good practice performances can have many benefits for all stakeholders offering an added value to manufacturers, a quality guarantee to consumers and an effective tool for farmers to make the proper choice for their investment.

Besides, it offers the public sector a strong and effective tool to identify the best technologies in order to focus better subsidy policies and extension efforts. With this approach we can assess that the certification of digital farming practices as well as technologies can be effective even for education and training at different levels. At present time the main problem is that public extension services as well as many education courses do not offer an up-to-date program concerning digital farming and related technologies. In order to provide farmers with effective knowledge it is highly recommended to certify technologies and offer a ranking based on the improvement they can offer. Furthermore the certification of farming practises and technologies can be useful for a sustainability index. Actually the only well established certification system for these technologies is the ISOBUS providing for a unique dialogue system between tractor and implement.

The weak point of ISOBUS is the cost and the fact that it is common in expensive and large machines and not available on all small scale farming machines, while on the other hand it is a user friendly system. It is well known that digital farming makes every operation in agriculture more effective and based on the real needs of crops. In this frame a reduction of the use of chemicals, water for irrigation and other inputs is to be expected when compared to traditional agriculture. Furthermore even the quality of agricultural products will be better because of the reduction of chemical residuals etc. As an example of index we can compare the use of agricultural machinery in a frame of precision farming rather that in a traditional system. The less chemical being used the higher ranking in the indexing of the process. The same can be done for all other inputs.

Certification is usually based on specific check lists in order to establish analytic parameters that will assess the real performance. In other words a minimum requirement could be a traditional crop protection practice that using a fixed amount of a certain fertilizer or pesticide might provide the best results for the crop.

On this basis every technology that will allow the same result or better with a reduced use of whatever inputs, such as fertilizers, pesticides, fuels, etc. will get a better evaluation and ranking in the certification process. Certification will assess the good final result with a reduced input based on less chemical and/or a less intense use of machines compared to present practices. Certification will then set up an index with a minimum requirement and a progressive better value according to the results.

Precision Agriculture technologies are able to identify clearly the real need of crops at very local level and optimize the inputs only on these locations. In other words fertilizers and pesticides will be used only where it is really needed and not on the whole crop area. Furthermore extension activities will be focused on making farmers aware of the benefits of new technologies and on the use of the certification as a tool to identify the best technology to be used in their farm.

This system will allow farmers to have a clear idea on the effectiveness related with the use of a certain technology in terms of less work inputs, better environmental conditions and expenses.

The aim will be focused on demonstrating that the digital technologies used can have clear benefits:

- Environment: less fertilizers and pesticides will be used,
- Quality of production: less chemicals = less residuals, so more healthy food products,
- Quality of life: humans will be less exposed to chemicals,
- Market: economic benefits for farmers and consumers.

All these parameters can be a framework of reference for certification and value of the certification to a clear index assessment. The same process can be followed to reduce/optimize the parameters (i.e. the use of water for irrigation). In addition their products can have a clear traceability process (i.e. protocols based on DLT) of inputs being used in order to get a higher value on markets. All these issues supporting a more sustainable farming will be clearly identified and measured in relation to a set of indicators certifying a sustainability rating.

Impacts of Precision Agriculture and Certification Needs

Impacts on different regional levels based on the foreseen changes can be evaluated and measured, and the outputs described.

The overall desired impact is to foster enrepreneurial and professional handling of innovative farming means by farmers, strengthening their business and role in the market and in society with effect on more sustainable development in rural areas and effective reduction of water, soil and air pollution and GHG emissions.

Conclusion

Entrepreneurial implementation of this kind of financing is a great challenge. Therefore, it is also highly important to raise awareness for the importance of this topic amongst all rural stakeholders and authorities. There is a strong need for further awareness and recognition at regional and national level. By offering knowledge on existing practical successful experiences of innovative farming based on Precision Agriculture at European level the project will deliver very useful inputs for this important innovative trend for a more sustainable development in rural areas and launching a broad discussion and dialogue as central basis for well led conceptualization and implementation of innovative digital farming, starting from the State-of-the-Art of this very recent development and ongoing experience for different crops and agricultural sectors.

The impact can be evaluated at different levels. At local level:

i) Exchange and development of Precision Agriculture on farms, with specific reference to IoT and Apps for spraying machinery.

ii) Practical case studies inspring feasible introduction of Precision Agriculture on farms

iii) Rising awareness for innovatve trends introducing planning digital farming solutions for agricultural machinery on farms, also improving quality of life in rural areas.

At national level:

i) Further development of the project results in collaboration activities with national authorities and scientific institutions, keeping a strong focus on good practices emerging from the case studies.

ii) Future inputs of practical ways of realizing concepts and training materals showed in case studies in benchmarking from different countries and embedding transnational ideas and experiences in the field of Precision Agriculture to different international and national contexts.

iii) Further national work in this field with analysis of most common problem solving supporting evolving trends for precision agriculture.

At European Level:

i) Awareness on importance and feasibility of Precision Farming introductory planning as important part of sustainable rural development and employment social inclusion as well as new opportunity for economic development in rural areas

ii) Exchange among experts linking theory and practice at European level aiming at further convergence of European nations and regions for planning measures of sustainable farming based on Precision Agriculture.

iv) Development of further relationships at European level involving Turkey in the field of Precision Agriculture for a more sustainable soil, water and air, fostering exchange of experiences.

REFERENCES

[1] https://www.cema-agri.org/publications/8/download

[2]https://www.slideshare.net/ChristianMStracke/ 20180713-mooq-conference-in-athens-mooq-and-the-quality-of-moocs-how-it-started-and-continues-stracke

[3] https://www.gsa.europa.eu/european-gnss/what-gnss

[4] https://www.aef-online.org/the-aef/isobus.html

[5] https://www.worldgovernmentsummit.org/api/publications/document?id=95df8ac4-e97c-6578-b2f8-ff0000a7ddb6

[6] Balsari P. et al. Developing Strategies to Reduce Spray Drift in Pneumatic Spraying in Vineyards: Assessment of The Parameters Affecting Droplet Size in Pneumatic Spraying, Oct. 2017 Science of The Total Environment

[7] http://www.agriprecisione.it/wp-content/uploads/2010/11/ general_introduction_to_ precision agriculture.pdf

[8] https://www.oliverwyman.com/our-expertise/insights/2018/feb/agriculture-4-0--the-future-of-farming-technology.html

[9] European Commission, Internal Market, Industry, Entrepreneurship and SMEs Industry 4.0 in Agriculture: Focus on IoT Aspects, July 2017

[10]https://www.euractiv.com/section/agriculture-food/infographic/farming-4-0-the- future-of-agriculture/

[11] https://ec.europa.eu/eip/agriculture/sites/agri-eip/files/eip-agri_focus_group_on_precision_farming_final_report_2015.pdf

[12] European Parliament, Precision Agriculture in Europe. Legal, Social and Ethical Considerations. European Parliamentary Research Service Author: Mihalis Kritikos

Facebook Games Applications

Babaev Vladimir Yandashevich

Master of Arts in International Business and Management, Lecturer of Tashkent State University of Uzbek Language and Literature Named After Alisher Nava'I, Taskent City, UZBEKİSTAN

Abstract: This article sheds light on how games applications increased its popularity using social network platform such as Facebook. What is more, this piece of writing reflects rapid evolution of particular games such as "Farmville", "Pet society", which using Facebook API. In addition, the article provides information what technologies, policies, protection measures; Facebook takes to protect users' personal information, "OAuth 2.0 protocol", in particular. Additionally, the article provides information concerning benefits Facebook and its users get from using particular games apps, challenges they are facing. Finally, the article gives some recommendations how Facebook and its followers can cope with these challenges.

Keywords: Online Social Networks; Web 2.0; OAuth 2.0 Protocol; Application Programming Interface

Introduction

Due to the rapid development of Web 2.0, popularity of Online Social Networks (OSNs) dramatically increased for the last two years. According Scott Golder et al. in the USA alone, number of undergraduate students using OSNs every day reached number of 90%. (Scott Golder et al. 2007). No wonder that, OSNs such as Facebook is the most visited website in the internet. (Comscore, 2008). Looking at the UK, 10% of all connections to the internet are to OSNs. What is more, popularity of OSNs outweighed even pornography websites (R. Goad, 2009).

Facebook has created a unique digital environment for third party developers to design various applications running on Facebook platform, in order to enhance number of users on its page, which constituted more than 200 million users in western countries alone, while China users hit the number of more than 300 million. (Cosenza, V, 2009). Games applications designers found opportunity to use Facebook as a gaming platform quite lucrative.

Facebook is the Best Platform For Games Apps

Despite the fact that number of SNSs largely contributed to the expansion of games Facebook is the one of the best platforms, where games applications has the biggest impact. For instance, "Farmville", which is simulation farming management created by Zinga Company, designed for Facebook platform, currently has 83,131,550 active users. Whereas, "Mafia Wars" the game in role-play genre created by the latter developer's number of players hit 25,225,819. 5. (Ines Di Loreto, 2010).

Many games designed for Facebook based on browser, and, as a result do not require further installation (Ines Di Loreto, 2010). This fact makes it easy for users to navigate through the game, not spending a lot of time on installation process.

In December 2009, 208 apps were the most downloaded games applications in the Facebook official website. One of the most popular casual games was "Solitaire" apps, with 600,000 active users. Creation of "casual games" is not incidental, these kinds of genres designed specifically for users who position themselves as "casual gamers", rather than "gamers". Taking into consideration, the casual gamers and specific niche that these users comply together, Facebook's engineers designed quite new game category "Just for Fun", which put together players regarding themselves "casual gamers". (Ines Di Loreto, 2010).

These actions taken by Facebook, shows its great adaptability to users' preferences and tastes in games apps. Casual games in its nature do not necessarily have complicated gameplay and designed mainly for entertaining and relaxation. (Ines Di Loreto, 2010). As Ahn, L. pointed out, user spend much of their time on games, which successfully designed. (As Ahn, L, 2006). Bearing his words in mind, number of games designers trying their best to create something great, thus luring more users to install their apps.

Facebook API As Incitement For Third Party Designers

Owing to Facebook Application Programming Interface (API), third party games designers can easily integrate their apps with Facebook platform. Thus, "Texas Holdem Poker" apps are one of the most successful games apps on Facebook with more than 1,7 million users on daily basis (Adonomics, "Top applications.").

Since 2007, thousands of apps were brought on Facebook, using its API. Various apps themes were introduced, such as presenting gifts, casual talking among friends, bodily movements, suggestions, etc. (M. Gjoka, et al. 2008).

Sharabi (A. Sharabi, 2007) identifies apps according to their social intentions:

- "Self-Presentation Tools", which is how people distinguish their authenticity with regard to their best movie or present state of feeling.
- "Collective Identity Formation" is asking people how they define other peoples' authenticity - Pleasant or Unpleasant games apps asking users to choose any adjective to depict their friend.
- "Phatic Communication Tools" is sending "gifts", "hugs", etc. in order to sustain "social contact" among users.

Looking in more details at "Hugs" application (http://apps.facebook.com/huggees), users pick up any kinds of "hugs" which they intend to send "Fuzzy Hug" or "Friendly Hug" , whereas, addressee receive a "hug" together with a picture panda, dog any funny animals. (A. Nazir, 2008).

It is clear that games apps designers, together with Facebook engineers, aim to strengthen relationship among "friends", by using quite unobtrusive communication tool.

The "Oauth 2.0 Protocol" As Additional Safety Measures

Facebook quite masterly uses information management to protect their users, which is their main priority. Thus, Facebook uses "OAuth 2.0 protocol", in order to reduce information transferring between users and third parties' apps designers, who can intentionally violate users' private information, can distribute it to other parties, without user consent. This protocol provides additional protection authentication level, when user trying to install apps on his or her Facebook profile, application that is going to be installed asks users' permission to his or her information access, which is kept on Facebook. (N Wang et al, 2011). Below there is Figure 1 demonstrating how "OAuth 2.0 protocol", operates.

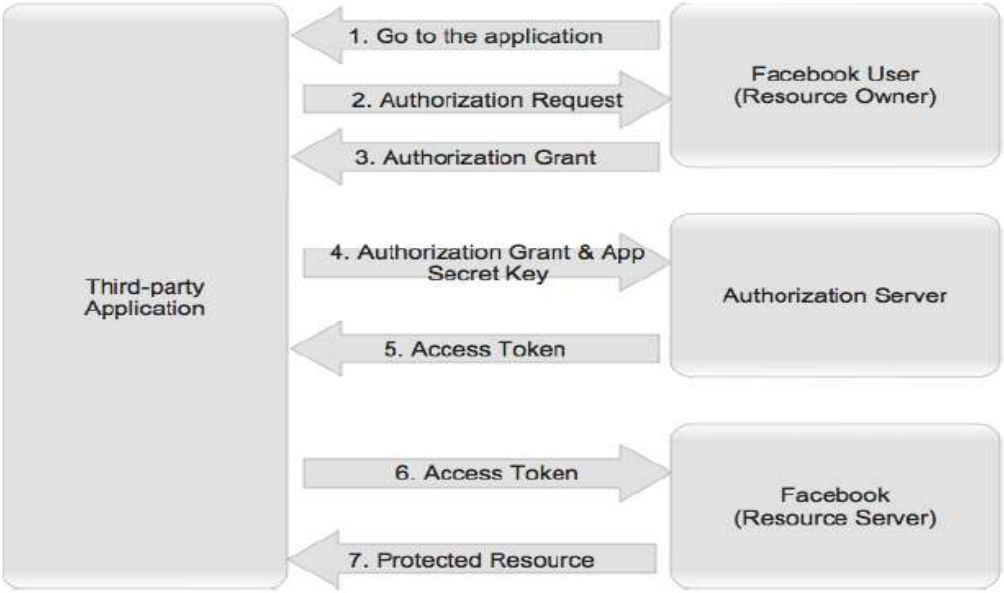


Figure 1. The OAuth 2.0 Protocol 4

Source: N Wang et al, 2011.

This information usually concerns user's name, gender, friends' list, avatar, etc. Such kind of information refers to "basic information", and if application trying to acquire data, which is beyond the "basic information", this application needs to prolong user's consent. Generally, additional consent refers to extended information of user, which include his or her personal contacts, friends' personal information, photos, videos, etc. (N Wang et al, 2011). Figure 2 depicts claim to this consent.

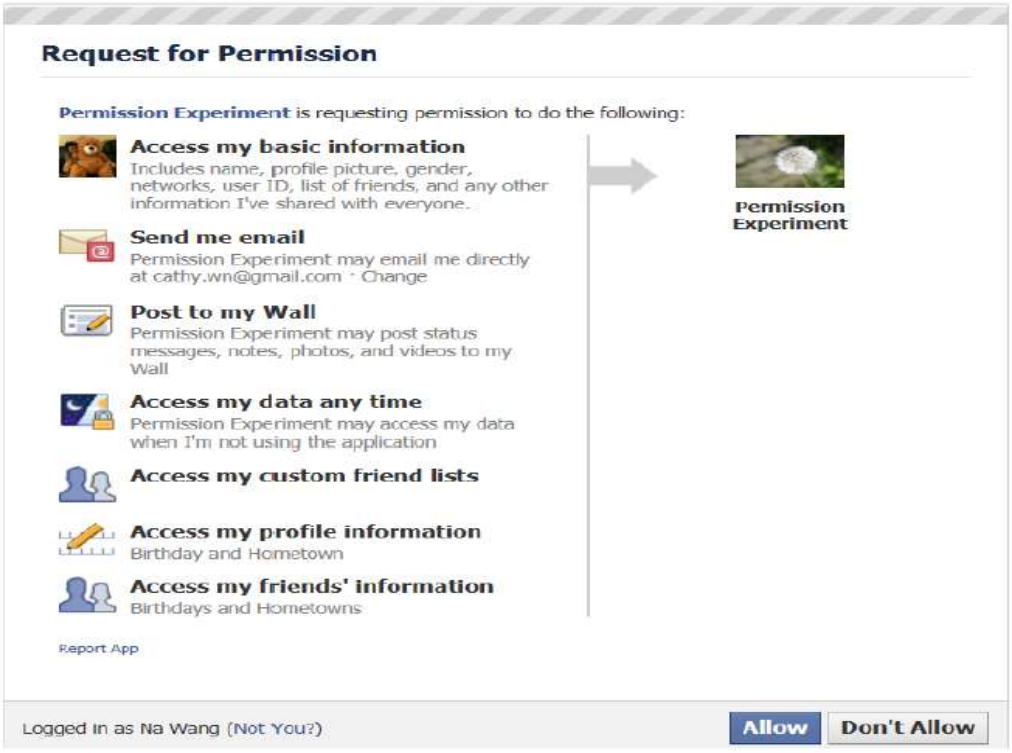


Figure 2. Request for Extended Permission

Source: N Wang et al, 2011.

Information control is user's prerogative Users in turn, can protect their personal information by adapting their "privacy settings", in order to restrict or limit access to their private data to other users in the cyberspace. By changing private settings, user can control general information his or her friend can use to locate him or her on Facebook. Secondly, regulate who will able to inspect their sharing content. Last, but not least, users are able to modify list of people who is blocked. (N Wang et al, 2011). Figure 3 demonstrates the interface of "privacy settings".

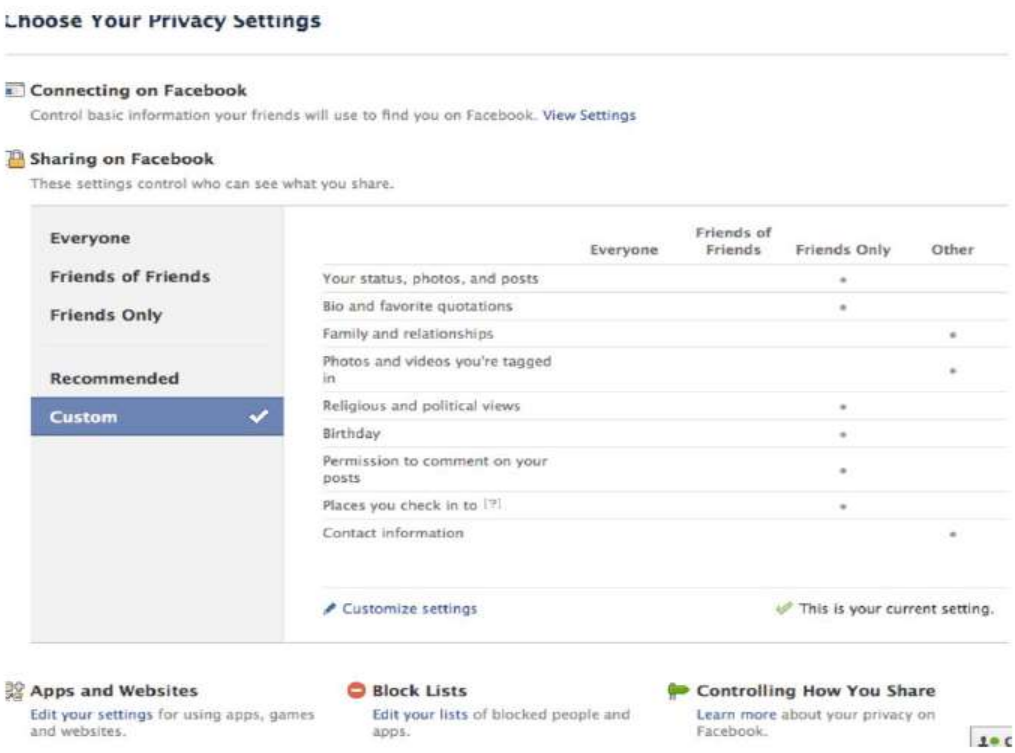


Figure 3. "Privacy Settings"

Source: N Wang et al, 2011.

Bearing in mind the information above, it is clear, that Facebook interface and digital environment are quite flexible. Users can modify, number of settings to their needs and tastes. Undoubtedly, users control and manage their information according to their preference; they choose whom to give access to their personal content, whom to make friends with, whom to blocked. Generally, Facebook grants "carte blanche" to user.

"Playfish's" "Pet Society" As A Great Socializing Tool

Facebook games applications represent not entertainment and relaxation solely, these apps bring people together, consolidate the unity inside the friends' group. Games apps unify not just user's friends, whom he or she can meet on a regular basis but distant relatives, who live in a faraway country, and there is rare possibility to club together more often.

This great opportunity represents "Playfish", company headquartering in UK, founded in 2007 by, Sami Lababidi, Kristian Segerstrale, Shukri Shammas and Sebastien de Halleux, and, and now is in Electronic Arts' possession (www.playfish.com/?page=company).

"Playfish" for quite a short period made a spectacular revenue and attracted millions of users. "Playfish" designs its games mainly for Facebook platform attracting above 11 million users every month. Currently, 5 out of 7 "Playfish" company's games are leading the top 25 games applications based on Facebook platform. "Pet society", which is the biggest success of the "Playfish" company, and the first application that reached 1 million players on Facebook alone. (L. Rossi, 2010).

"Pet society" success phenomenon is quite simple to distinguish. Firstly, this game designed for all ages, genders; no matter do they have their own pets or going to acquire any. Players find this game quite interesting, non-obtrusive, and fun. Figure 4 illustrates screenshot of "Pet society" apps.

Figure 4. "Pet society"

Source: www.free2play.com, 2013

The gameplay is quite simple: players looking after their pets, feeding, playing with them, etc. "Pet society" has two modes "Single" and "Multiplayer". Players within the friends' network can visit each other's pets, doing various activities such as cleaning, giving food, etc. Players, getting coins for each visit of their friends' pets visit. In order to obtain more coins, players should extend their friends' network, more friends, higher revenue, and as a result greater achievement.

With more coins in their pocket, pets' owners are able to buy more stuff for their pets (www.facebook.com/petsociety).

Taking into consideration information mentioned above, it is obvious that players to obtain desired coins, actively reviewing their current friends list, and trying to become friends with as many people as possible. These potential friends, who can largely contribute to player game status, come not just from close circuit, but even people they are working with and ex-schoolmates, whom players did not consider as friends. (L. Rossi, 2010)

Due to games requirements in obtaining coins, players should extend their friends list, to be able to get more coins. Usually players, create separate group of "friends" which are people users interacting during the game process, rather than hanging out together. Eventually, "friends for social gaming" can transform into real life friends.

Owing to unique Facebook digital environment, every action, committed by player, all his achievement during the game process, people they become friends with, instantly display on player's profile. This fact is a challenge for other users, who is involved in the same game, play more often, becoming friends with larger number of people. What is more, player should maintain this created relationship; otherwise, they can lose their superiority in the game. (L. Rossi, 2010)

Facebook demonstrates effective information control and marketing strategies. Firstly, they are cooperating with third party developers, extending their influence area among games designers and technical community. Second thing is that owing to Facebook people are always in touch with each other and it does not matter what brings them together, game or any other occasions. The most important thing is that people gathering, sharing their success, achievements, etc., no matter in a real-life world or in the cyberspace.

Foul Play of Games Apps Designers

However, despite Facebook assurance of users' personal information protection, there some flaws in games apps designed for Facebook platform. According to a Wall Street Journal's finding, number of popular games apps in Facebook, pass users' personal information including age, photos, friends list to third parties, such advertising agencies, data collection centers, thus violating users' privacy, without their awareness of this actions taking place. Among privacy offenders is "Zynga Company", "Farmville" game apps creator, which accused of transferring users' friends'

details to third parties organizations. Currently, “Farmville”, is not accessible for Facebook users. “LOLapps” and “Family tree” apps have been passing users’ ID to “RapLeaf” data center, a Wall Street Journal have found. All companies mentioned denied the fact of privacy violation, and declared all actions were unintentional, and their main priority was and still users’ privacy protection. After findings publication, Facebook shut down number of apps depicted in investigation. As Facebook authorities claimed, they are taking all necessary steps to prevent users’ personal information leakage. Nevertheless, it is still unclear what steps, they are taking and how they can stop this information leakage. It is true, that due to Facebook privacy regulations, games designers restricted players’ personal information transmission to other parties, even though, with user’s approval. However, it should be noted, that Facebook highly relied on varies apps created by third parties’ designers, because apps are great tool in extending Facebook network. Currently, 70 percent of Facebook users utilizing various apps on monthly basis (E Steel, G Fowler, 2010).

Despite, shutting down apps, which is not reliable, trustworthy and do not complying with privacy regulations, Facebook cannot stop them operating all at once. What is more, it is not clear how many other apps and its designers violate users’ privacy.

Conclusion

Undoubtedly, Facebook platform one of the most popular social network in the world, with millions of users all around the world. Facebook is not just the platform where are people are getting together, it is more than that, it is a quite new “digital world”, where every current or potential user can find something extraordinary for themselves, and make something they really want into reality, even impossible from the first sight. It is out of the question that among the vast majority of apps available in Facebook, games apps has its own niche. Games apps for quite a short period become extremely popular among Facebook users. It is evident, that games apps encroach users’ privacy, and there a long way ahead for Facebook to protect personal content millions of its users.

On the other hand, Facebook developing new strategies, policies, to prevent users’ private information dissemination to other parties. Moreover, Facebook grants more options to users in setting up their privacy settings within their profiles. Additionally, Facebook lets people manage their own information, according to their personal preferences. The best strategy for Facebook to prevent its users’ privacy violation is to limit number of games creators using its platform and

decrease collaboration with third parties' apps designers. Finally, to start designing more their own apps themselves, thus ensuring users' privacy protection and less reliability on other apps creators, who intentionally or unintentionally can undermine users' confidence.

It is still unclear, should people share their personal details with any other social networks, in order to maximize protection of their accounts. Nevertheless, it is out of the question, that any social networks and apps cannot substitute face-to-face socializing.

REFERENCE

[1] A. Nazir, S. Raza, and C.-N. Chuah, "Unveiling Facebook: A Measurement Study of Social Network-Based Applications," in IMC '08: Proceedings Of The 8th ACM SIGCOMM Conference on Internet Measurement. New York, NY, USA: ACM, 2008, Pp. 43–56.

[2] A. Sharabi, "Facebook applications trends report," November 19th, 2007. [Online]. Available: http://no-mans-blog.com/2007/11/19/facebookapplications-trends-report-1/

[3] Ahn, L. v. Games with a Purpose. Computer, 39, 6 (2006), 92-94.

[4] "About Playfish Company & Founders" http://www.playfish.com/?page=company

[5] Adonomics, "Top Applications." [Online]. Available: http://adonomics.com/leaderboard.

[6] Cosenza, V. World Map of Social Networks. Vincos Blog, 2009. http://www.vincos.it/world-map-of-socialnetworks.

[7] Comscore. http://www.comscore.com/, Apr 2008.

[8] E Steel, G Fowler- The Wall Street Journal, 2010 Facebook in Privacy Breach - terriau.org, http://terriau.org/blog/postings/20101018%20Facebook%20in%20Online%20Privacy%20Breach%3B%20Applications%20Transmitting%20Identifying%20Information%20-%20WSJ.pdf

[9] Ines Di Loreto, Abdelkader Gouaich. Social Casual Games Success is not so Casual. RR-10017.

[10] 2010, pp.001-011. <lirmm-00486934> http://hal-lirmm.ccsd.cnrs.fr/lirmm-00486934/ document.

[11] L Rossi Playing Your Network: Gaming in Social Network Sites. Available at SSRN 1722185, 2010 - papers.ssrn.com http://www.digra.org/wp-content/uploads/digital-library /09287.20599.pdf

[12] M. Gjoka, M. Sirivianos, A. Markopoulou, And X. Yang, "Poking Facebook: Characterization of Osn Applications," in WOSP '08: Proceedings Of The First Workshop On Online Social Networks. New York, NY, USA: ACM, 2008, Pp. 31–36.

[13] N Wang, H Xu, J Grossklags - Third-Party Apps On Facebook: Privacy And the Illusion of Control Proceedings Of The 5th ACM Symposium, 2011-dl.acm.org, http://personal. psu.edu/nzw109/papers/Wang_chimit_2011.pdf.

[14] R. Goad, "Social Networks Overtake Adult Websites," Hitwise Intelligence Online, 2009.

[15] Scott Golder, Dennis Wilkinson, and Bernardo A. Huberman. Rhythms of Social Interaction: Messaging within a Massive Online Network. In International Conference on Communities and Technologies, 2007.

[16] www.free2play.com

[17] https://www.facebook.com/petsociety

A Note on the Examination of Portugal's Hotels Performance - a Proposal for a New Perspective's Approach

Nuno Mendes FREITAS, José António FILIPE

Information Sciences, Technologies and Architecture Research Center

Instituto Universitário de Lisboa, PORTUGAL

Abstract: The main objective of the present research paper is to give an exploratory vision on a new form of examining hotels' performance of the Portuguese hotels' market. This approach is based on the experience of a consultant in this area and is made in terms of revenue and occupancy, crossing then operational data with financial results. This analysis focuses on a recent period, since 2010 to 2017, involving the years since before and until after the big boom of tourism in Portugal. This paper exposes available data about hotels operational results and compares these results with hotels financial results. In terms of the evaluation of a hotel's performance, the revenue per available room (RevPAR) is regarded by the tourism industry and INE (the Portuguese National Statistics Institute) as the most important measure, once it compares hotels among themselves and also compares regions. An important aspect of this study is highlighting the limitations of RevPAR.

Keywords: Revpar, Financial Results, Hotels Performance

Introduction

The present study is based on the operational and financial data available for the period between 2010 and 2017 regarding to Portuguese hospitality. For several decades, the lodging industry has used RevPAR (revenue per available room) as a key indicator to evaluate a firm's performance and to make investment decisions [1]. RevPAR is one of the most recognized and used performance measures in the hospitality industry, providing general market trends and some revenue indications. There are some pitfalls to be aware of when analyzing a hotel's performance based solely on RevPAR [2]. RevPAR only measures revenues; it ignores expenses and does not

capture or report profitability. Thus, RevPAR is a rather limited measure for measuring the hotel's performance [3]. This indicator is calculated by dividing a hotel's net room revenue (considered after discount and sales taxes and net of breakfast and other meals) by the total number of available rooms and by multiplying a hotel's average daily room rate (ADR) by its occupancy [2].

Hotels, with significant revenue-generating departments rather than the rooms division, by focusing only on room revenue do not provide a complete or accurate picture of the entire hotel property's performance [3].

For the analysis of the hotel's performance, other complementary metrics should be used to measure the financial performance, by crossing the financial result with the number of occupied rooms or available rooms. This will allow hotel decision makers to take into account the costs of the operation in the creation of the room price, and to realize how their operating decisions and cost structure reflect the net revenue. RevPAR is an important indicator when comparing the hotel performance with the one of competitors and with the market. Anyway, it is not the best indicator to analyze the operational performance of hotel units. A greater value in the RevP AR to the market or competition is not necessarily a sign that the hotel is performing better than the others, as each hotel company has different operating structures with different costs.

The calculation of RevPAR

Rooms' revenue per available room (RevPAR) measures the rooms' revenue yield that a property achieves in terms of the rooms available in the property for a given period. It includes the influence of occupancy and ADR.

RevPAR is calculated as follows:

RevPAR = Total Rooms Revenue / Number of Rooms Available

The number of rooms available or the number of the rooms' nights available is based on the calculations of weighted average number of bedrooms in the hotel, multiplied by the total number of nights in the period under the analysis.

Total rooms revenue is the number of rooms' nights sold in the period of analysis.

The Portuguese Tourism and Hotels Market

In this market, there was recently a recession in Portugal, involving the period between 2010 and 2014 most considerably, but having started with the global financial crisis in 2007-2008.

TABLE I. GDP GROWTH RATE – PORTUGAL. UNIT: EURO – MILLIONS

GDP Growth Rate	
2010	1.90
2011	1.83
2012	- 4.03
2013	- 1.13
2014	0.89
2015	1.82
2016	1.93
2017	2.80

Source: BdP.

In recent years, Portugal had a significant growth in terms of tourism. In 2017, 12.7 million tourists visited Portugal, representing 3.4 billion euros in terms of hotels' revenues. Tourism is one Portuguese key sector, having a great importance for the Portuguese economy with a relative weight of 13.17 % in Portuguese GDP in 2017. The number of accommodations in Portugal, according to the National Statistics Institute (INE), is 5840 in 2017.

TABLE II. Travel and Tourism Account as a Percentage of GDP. Unit: Euro - Millions

Year	*% of GDP*
2010	6.31
2011	7.16
2012	7.78
2013	8.78
2014	10.10
2015	10.75
2016	11.57
2017	13.17

Source: INE.

Literature Review

Many observers, particularly those in the financial community, view the hospitality industry generally as being a risky business for owners and investors. This vision can be explained by considering much of the investments risk stems in the fundamentals of this business, namely, room-rate and occupancy fluctuations [4]. Several research papers in the literature show the importance of RevPAR as a financial performance indicator ([5]. Higgins interviewed industry analysts and managers, regarding the importance of RevPAR [6]. These interviews indicated RevPAR as the most widely used measure, internally and externally, for hotels performance analysis and accepted by lodging firms as a benchmark. A limited amount of research discussed and investigated the relationship between the lodging firm performance and RevPAR [7]. . Brown and Dev highlighted two key limitations of RevPAR: it does not include revenue from food and beverage and other departments; and it does not take into account costs that are incurred to provide the requisite service level (e.g., special amenities such as a spa or additional guest-service employees such as a concierge) [8].

Gross Operating Profit per Available Room (GOPPAR) has become popular as an important alternative performance measure because it resolves limitations of RevPAR [9]. It provides a deep indication of a hotel's profitability by taking into consideration management control and efficiency, and eliminating, to a certain extent, the potential advantage of a small hotel on this analysis.

In addition, GOPPAR offers an overall more robust performance measure, especially when comparing the financial performances of hotels with different sizes or in different markets [2].

One of the issues in lodging finance is the ratio of debt to equity [10]. According to [11], the hospitality industry is traditionally confronted with a high need for financial capital to invest in fixed assets such as land, building, and equipment, and since debt is relatively cheaper than equity, it has been widely used as a source of capital to fund investments.

Methodology

a. The Study

The analysis of data in this study has two stages. First, the study is conducted considering the hospitality statistical data issued by the Portuguese National Statistics Institute. The number of

occupied rooms and available rooms is calculated and then RevPAR and ADR from 2010 to 2017 are also calculated.

Second, financial reports of available data, assessed from the Central balance sheet database of the Portuguese Central Bank (Banco de Portugal), are used as the relevant financial information for the examination of the perspective of financial performance per available room and per occupied room.

b. The Data Collection

The study collected data from two main sources.

Financial data are collected from the Portuguese CAE 551 – Hotels and similar accommodations.

The operational data from INE statistics used in this study are the tourism statistics, having been this information collected by INE, from a survey in hotels and other similar establishments.

c. Corporate performance measures

Return on assets (ROA) is the net income divided by total assets. It reflects the ability of a firms' management to generate profits from firms' assets. Hotel managers have to ensure great profits in order to cover the high assets costs and the related fixed costs. These profits have to be coherent with hotel investments. The ROA is retained to evaluate hotel profits enclosing all provided services and activities.

ROA = (Net Income/ Total assets) X 100

Return on equity (ROE) is used to measure profitability of companies, and is the net income divided by total equity, measuring the firm efficiency to generate profits from shareholders investment. The company's objective is the maximization of shareholder's value to have an optimum mix of debt and equity without compromising ROE.

ROE = (Net Income / Total equity) X 100

Data Analysis

RevPAR variability is measured by the mean of the absolute values of the annual changes in RevPAR. Table III presents the mean of the annual measures during the period from 2010 to 2017. It is noted that the RevPAR variability is growing during the period between 2012 to 2017.

TABLE III. REVPAR. UNIT: EURO

Year	Revpar	Variability
2010	28,27	
2011	29,39	4%
2012	28,47	-3%
2013	30,20	6%
2014	33,01	9%
2015	37,60	14%
2016	43,23	15%
2017	50,26	16%

Source: BdP.

When analyzing the ROA, we verify that it is inconsistent over the years under analysis, comparatively to RevPAR that had a slight decrease from 2011 to 2012 and returned to a constant growth from 2012 to 2017.

Assets grew during the years under analysis, and the negative variation was verified in net income.

TABLE IV. ROA

Year	ROA	Variability
2010	2.35%	
2011	0,89%	-62%
2012	0,50%	-44%
2013	1,87%	274%
2014	3,27%	75%
2015	5,15%	57%
2016	5,38%	4%
2017	7,34%	36%

Source: BdP

From 2010 to 2104, ROE has always had a negative value, due to the Net results of hotel companies. Comparing to RevPAR, the variability is not constant. RevPAR shows a growth between 2013 and 2017; ROE only grows in 2016 and 2017.

TABLE V. ROE.

Year	ROE	Variability
2010	-5,26%	
2011	-10,96%	-108%
2012	-15,90%	-45%
2013	-10,82%	32%
2014	-5,21%	52%
2015	1,56%	130%
2016	2,07%	33%
2017	7,61%	268%

Source: BdP.

When analyzing EBITDA or Net Income per occupied room, we observe that Revpar grew from 2012 to 2017 and had a constant positive variation. However, we do not observe the same trend when analyzing the values presented in Table VI and table VII.

TABLE VI. EBITDA PER OCCUPIED ROOM. UNIT: EURO

Year	EBITDA	Variability
2010	12,59	
2011	4,90	-61%
2012	2,84	-42%
2013	10,45	268%
2014	16,65	59%
2015	24,57	48%
2016	25,39	3%
2017	33,84	33%

Sources: BdP.

TABLE VII. NET INCOME PER OCCUPIED ROOM. UNIT: EURO

Year	Net Income	Variability
2010	-9,76	
2011	-17,61	-80%
2012	-22,95	-30%
2013	-14,71	36%
2014	-6,80	54%
2015	2,11	131%

Year	Net Income	Variability
2010	-9,76	
2011	-17,61	-80%
2016	3,07€	45%
2017	11,56	277%

Sources: BdP.

Discussion, Conclusion and Limitations

We can verify that complementary ratios should be used to achieve and control financial results. The used ratios do not measure financial costs, only the operational costs of the hotel.

Another strand for the analysis is related to the capital structure used in hotels. Investments are made using financing, being the financial autonomy always below 30% and for that reason, in some years, the net income was negative. The use of the net income per available room can be a starting point to measure the financial performance.

RevPAR is the dominant, and currently, the most used and accepted measure by tourism and hotels, and, considering that, it is possible to understand how important is to search for alternative performance measures.

REFERENCES

[1] J. Chen, Y. Koh, And S. Lee, "Does The Market Care About Revpar? A Case Study Of Five Large U.S. Lodging Chains," Journal Of Hospitality And Tourism Research, Vol. 35, No. 2, 2011.

[2] E. Younes And R. Kett, "GOPPAR, A Derivative Of Revpar," HVS International, No. March, 2003.

[3] S. Lee, B. Pan, And S. Park, "Revpar Vs. GOPPAR: Property- And Firm-Level Analysis," Annals Of Tourism Research, Vol. 76, No. September 2018, 2019.

[4] J. A. Ismail, M. C. Dalbor, And J. E. Mills, "Using Revpar To Analyze Variability," Vol. 24, No. 4, 2002.

[5] M. Gallagher And A. Mansour, "Analysis of Hotel Real Estate Market Dinamics." American Real Estate Society, 2000.

[6] S. M. Higgins, "Revpar Still King, But GOPPAR on The Rise," Hotel & Motel Management, Vol. 221, No. 1, 2006.

[7] A. R. Elgonemy, "The Pricing Of Lodging Stocks: Reality Check." Cornell Hotel And Restaurant Administration Quarterly, 2000.

[8] J. R. Brown And C. S. Dev, "Looking Beyond Revpar: Productivity Consequences of Hotel Strategies," Cornell Hospitality Quarterly, Vol. 40, No. 2, 1999.

[9] S. Lee, B. Pan, And S. Park, "Revpar Vs. GOPPAR: Property- And Firm-Level Analysis," Annals Of Tourism Research, Vol. 76, No. September 2018, 2019.

[10] A. R. Elgonemy, "Debt-Financing Alternatives," No. June, 2002.

[11] Y. Li And M. Singal, "Capital Structure İn The Hospitality İndustry: The Role of The Asset-Light And Fee-Oriented Strategy," Tourism Management, Vol. 70, No. August 2018, 2019.

[12] BP STATS – Banco De Portugal 2010 To 2017, Lisboa.

[13] INE Instituto Nacional De Estatística 2010 To 2017, Lisboa.

A Comparative Case Study on Time Series Prediction

Anıl ÖZDEMİR, Furkan COŞKUN, Selim BALCISOY

Faculty of Engineering and Natural Sciences, Sabancı University, TURKEY

Abstract: A time series is a sequence collected at consecutive equally spaced points in time. The basic idea behind the time series forecasting is the use of a model to estimate future values based on previously observed ones. Traditionally, statistical methods are used to forecasting time series however, Machine Learning (ML) algorithms have been also proposed as alternatives to statistical methods in past decades. In this paper, we evaluate forecasting performance of different ML algorithms and statistical methods on Turkey automobile sales. Recently, various of work has claimed that traditional statistical methods dominate the ML solutions in terms of time series forecasting. This study discusses different aspects of ML and statistical methods and compare their performance on different time series.

Keywords: Time Series Forecasting, Machine Learning Regression, Statistical Models.

REFERENCES

[1] Brillinger, David R. Time Series: Data Analysis And Theory. Vol. 36. Siam, 1981.

[2] Brown, Robert Goodell. Statistical Forecasting For Inventory Control. McGraw/Hill, 1959.

[3] Whittle, P. (1951). Hypothesis Testing in Time Series Analysis. Almquist and Wicksell. Whittle, P. (1963). Prediction and Regulation. English Universities Press. ISBN 0-8166-1147-5.

[4] Hannan, Edward James (1970). Multiple Time Series. Wiley Series in Probability And Mathematical Statistics. New York: John Wiley And Sons.

[5] Box G, Jenkins G. Time Series Analysis: Forecasting and Control. San Francisco: Holden-Day; 1970.

[6] Brockwell, Peter J., Richard A. Davis, and Matthew V. Calder. Introduction to Time Series and Forecasting. Vol. 2. New York: springer, 2002.

[7] Bauwens L, Laurent S, Rombouts JVK. Multivariate GARCH Models: A Survey. Journal of Applied Econometrics. 2006;21(1):79–109.

[8] Makridakis, Spyros, Evangelos Spiliotis, and Vassilios Assimakopoulos. "Statistical And Machine Learning Forecasting Methods: Concerns And Ways Forward." PloS one 13, no. 3 (2018): e0194889.

[9] Gilchrist, Warren. Statistical Forecasting. Vol. 322. London: Wiley, 1976. [10] Zhang G, Eddy Patuwo B, Hu Y M. Forecasting With Artificial Neural Networks: The State of The Art. International Journal Of Forecasting. 1998;14(1):35–62.

[11] Kim, Kyoung-Jae. "Financial Time Series Forecasting Using Support Vector Machines." Neurocomputing 55, no. 1-2 (2003): 307-319.

[12] Bontempi, Gianluca, Souhaib Ben Taieb, And Yann-Aël Le Borgne. "Machine Learning Strategies For Time Series Forecasting." In European BusinessIntelligence Summer School, pp. 62-77. Springer, Berlin, Heidelberg,2012.

[13] Adya M, Collopy F. How Effective Are Neural Networks At Forecasting And Prediction? A Review And Evaluation. Journal Of Forecasting. 1998;17(56):481–495.

[14] Makridakis, Spyros, and Michele Hibon. "The M3-Competition: Results, Conclusions And İmplications." International Journal of Forecasting 16, no. 4 (2000): 451-476.

[15] Hibon, Michèle, and Herman Stekler. The M-3 Competition: Statistical Tests of the Results. INSEAD, 2003.

[16] Crone SF, Hibon M, Nikolopoulos K. Advances in Forecasting With Neural Networks? Empirical Evidence From The NN3 Competition On Time Series Prediction. International Journal Of Forecasting. 2011;27(3):635–660.

[17] Thiel, Christian, Johannes Schmidt, Arnold Van Zyl, And Eschmid. "Cost And Well-To-Wheel İmplications Of The Vehicle Fleet CO2 Emission Regulation İn The European Union." Transportation Research Part A: Policy And Practice 63. (2014): 25-42

[18] Hansen, James V., James B. McDonald, and Ray D. Nelson. "Time Series Prediction with Genetic-Algorithm Designed Neural Networks: An Empirical Comparison with Modern Statistical Models." Computational Intelligence 15, no. 3 (1999): 171-184.

[19] Goodwin P, Lawton R. On The Asymmetry of The Symmetric MAPE. International Journal of Forecasting. 1999; 15(4):405–408.

[20] Armstrong, J. Scott; Collopy, Fred (1992). "Error Measures for Generalizing About Forecasting Methods: Empirical Comparisons" (PDF). International Journal of Forecasting. 8 (1): 69–80.

[21] Tipping, Michael E. "Sparse Bayesian Learning and The Relevance Vector Machine." Journal of Machine Learning Research 1, No. Jun (2001): 211-244.

[22] Scikit-learn: Machine Learning in Python, Pedregosa et al., JMLR 12, pp. 2825-2830, 2011.

[23] Rokach, Lior; Maimon, O. (2008). Data Mining with Decision Trees: Theory and Applications. World Scientific Pub Co Inc. ISBN 978-9812771711.

[24] Altman, N. S. (1992). "An İntroduction to Kernel And Nearest-Neighbor Nonparametric Regression". The American Statistician. 46 (3): 175–185.

[25] Scho ¨lkopf B, Smola AJ. Learning with kernel: Support Vector Machines, Regularization, Optimization and Beyond. The MIT Press; 2001.

[26] Breiman, Leo. "Random Forests." Machine Learning 45, No. 1 (2001): 5-32. [27] Yan, Xin, And Xiaogang Su. Linear Regression Analysis: Theory And Computing. World Scientific, 2009.

[28] Weisstein, Eric W. "Binet-Cauchy Identity." CRC Concise Encyclopedia Of Mathematics (2nd ed.), CRC Press, ISBN 1584883472 (2003): 228.

[29] Seabold, Skipper, And Josef Perktold. “Statsmodels: Econometric and Statistical Modeling With Python.” Proceedings Of The 9th Python İn Science Conference. 2010.

[30] Hyndman, Rob J., And George Athanasopoulos. "8.9 Seasonal ARIMA Models." Forecasting: Principles And Practice. Otexts. Retrieved 19 (2015).

[31] Pal, Sankar K., and Sushmita Mitra. "Multilayer Perceptron, Fuzzy Sets, Classifiaction." (1992).

[32] Kedem, Benjamin, And Konstantinos Fokianos. Regression Models For Time Series Analysis. Vol. 488. John Wiley & Sons, 2005.

Optimal PID-like Fuzzy Logic Controller Design for Ball and Beam System

O. Tolga ALTINÖZ, A. Egemen YILMAZ

Department of Electrical and Electronics Engineering, Ankara University, TURKEY

Abstract: Ball and beam system (BBS) is a benchmark hardware for designing control action. The structure of the system is based on changing the angle of the beam so that the position of the ball is changed. It is desired to move the ball to a reference position. In this paper Fuzzy Logic Controller (FLC) is applied for this problem. Instead of conventional FLC, the derivative and integral terms are integrated to the FLC, which is called as PID-like FLC. This controller has a constant Fuzzy structure with variable parameters. The performance of the controller is based on these parameters. Therefore, in this study, the parameters of PID-like FLC are optimized by using three optimization algorithms; Genetic Algorithm, Particle Swarm Optimization, and Differential Evolution. The performance of the controller is demonstrated on both simulation and hardware environment. The performance of the optimization algorithm with respect to the obtained performances are compared in this paper.

Keywords: Ball and Beam System, Particle Swarm Optimization, Genetic Algorithm, Differential Evolution, Fuzzy Logic Controller, PID-like.

Adopting Machine Learning Algorithms for Cloud-Based Application Categorization

Çağatay ÇATAL[1], Besme ELNACCAR[2], Özge ÇOLAKOĞLU[2], Bedir TEKİNERDOĞAN[1]

[1]Wageningen University, NETHERLANDS; [2] Istanbul Kültür University, TURKEY

Abstract: Manual categorization of applications in software repositories such as SourceForge is often time-consuming and error-prone. Automation of this process not only simplifies the daily task of administrators but also helps project owners to add their projects into the corresponding subcategory of the repository without any delay. In this study, we propose a cloudbased application categorization system that applies machine learning algorithms to support the classification of applications. The categorization system has a web-based client application to parse, process, and submit the project source code, a web service which automatically performs classification of applications into domain categories, and a cloud-computing platform which hosts the categorization service. Several multi-class classification algorithms have been adopted including, Artificial Neural Networks, Logistic Regression, Decision Jungle, and Decision Forest algorithms to validate the effectiveness of the system in multiple case studies. The case studies were performed on three public datasets generated based on 3286 Java applications of SourceForge repository. Our study shows that the highest accuracy was achieved with Artificial Neural Networks (ANN) algorithm. The resulting prediction model has been transformed into a web service and then, deployed on the Azure cloud platform.

Keywords: Cloud Computing, Software Maintenance, Machine Learning, Application Categorization, End-To-End Cloud System

REFERENCES

[1] E. Kalliamvakou, G. Gousios, K. Blincoe, L. Singer, D. M. German, and D. Damian, "The Promises And Perils Of Mining Github," in Proceedings Of The 11th Working Conference On Mining Software Repositories. ACM, 2014, pp. 92–101.

[2] J. Howison And K. Crowston, "The Perils And Pitfalls Of Mining Sourceforge," in Proceedings of the International Workshop on Mining Software Repositories (MSR 2004). IET, 2004, pp. 7–11.

[3]SourceForge. (2018) Saourceforge. [Online]. Available: https: //sourceforge.net /directory/ os:windows/

[4] C. Catal, S. Tugul, and B. Akpınar, "Automatic Software Categorization Using Ensemble Methods And Bytecode Analysis," International Journal of Software Engineering and Knowledge Engineering, vol. 27, no. 07, pp. 1129–1144, 2017.

[5] K. Tian, M. Revelle, And D. Poshyvanyk, "Using Latent Dirichlet Allocation For Automatic Categorization Of Software," in Mining Software Repositories, 2009. MSR'09. 6th IEEE International Working Conference On. IEEE, 2009, Pp. 163–166.

[6] M. Linares-Vásquez, C. Mcmillan, D. Poshyvanyk, And M. Grechanik, "On Using Machine Learning To Automatically Classify Software Applications into Domain Categories," Empirical Software Engineering, Vol. 19, No. 3, Pp. 582–618, 2014.

[7] S. Kawaguchi, P. K. Garg, M. Matsushita, And K. Inoue, "Mudablue: An Automatic Categorization System For Open Source Repositories," Journal Of Systems And Software, Vol. 79, No. 7, Pp. 939–953, 2006.

[8] C. Mcmillan, M. Linares-Vasquez, D. Poshyvanyk, And M. Grechanik, "Categorizing Software Applications For Maintenance," in Software Maintenance (ICSM), 2011 27th IEEE International Conference On. IEEE, 2011, Pp. 343–352.

[9] S. Vargas-Baldrich, M. Linares-Vásquez, And D. Poshyvanyk, "Automated Tagging of Software Projects Using Bytecode And Dependencies (N)," in Automated Software Engineering (ASE), 2015 30th IEEE/ACM International Conference On. IEEE, 2015, Pp. 289–294.

[10] T. Wang, H. Wang, G. Yin, C. X. Ling, X. Li, And P. Zou, "Mining Software Profile Across Multiple Repositories For Hierarchical Categorization," in Software Maintenance (ICSM), 2013 29th IEEE International Conference On. IEEE, 2013, Pp. 240–249.

[11] S. Ugurel, R. Krovetz, And C. L. Giles, "What's The Code?: Automatic Classification of Source Code Archives," in Proceedings Of The Eighth ACM SIGKDD İnternational Conference on Knowledge Discovery And Data Mining. ACM, 2002, Pp. 632–638.

[12] A. Kuhn, S. Ducasse, And T. Gírba, "Semantic Clustering: Identifying Topics in Source Code," Information And Software Technology, Vol. 49, No. 3, Pp. 230–243, 2007.

[13] P. S. Sandhu, J. Singh, And H. Singh, "Approaches For Categorization of Reusable Software Components," 2007.

[14] M. B. Kelly, J. S. Alexander, B. Adams, And A. E. Hassan, "Recovering A Balanced Overview Of Topics in A Software Domain," İn Source Code Analysis And Manipulation (SCAM), 2011 11th IEEE International Working Conference On. IEEE, 2011, Pp. 135–144.

[15] A. Hindle, D. M. German, M. W. Godfrey, And R. C. Holt, "Automatic Classication of Large Changes into Maintenance Categories," İn Program Comprehension, 2009. ICPC'09. IEEE 17th International Conference On. IEEE, 2009, Pp. 30–39.

[16] J. Escobar-Avila, M. Linares-Vásquez, And S. Haiduc, "Unsupervised Software Categorization Using Bytecode," in Proceedings Of The 2015 IEEE 23rd International Conference on Program Comprehension. IEEE Press, 2015, Pp. 229–239.

[17] A. A. Al-Subaihin, F. Sarro, S. Black, L. Capra, M. Harman, Y. Jia, And Y. Zhang, "Clustering Mobile Apps Based On Mined Textual Features," İn Proceedings Of The 10th ACM/IEEE International Symposium on Empirical Software Engineering And Measurement. ACM, 2016, P. 38.

[18] W. Martin, F. Sarro, Y. Jia, Y. Zhang, And M. Harman, "A Survey of App Store Analysis For Software Engineering," IEEE Transactions On Software Engineering, Vol. 43, No. 9, Pp. 817–847, 2017.

[19] A. T. Nguyen And T. N. Nguyen, "Automatic Categorization With Deep Neural Network For Open-Source Java Projects," in Proceedings Of The 39th International Conference On Software Engineering Companion. IEEE Press, 2017, Pp. 164–166.

[20] A. Sharma, F. Thung, P. S. Kochhar, A. Sulistya, And D. Lo, "Cataloging Github Repositories," 2017.

[21] W. Chen, P. Xu, G. Wu, W. Dou, C. Gao, And J. Wei, "A Hierarchical Categorization Approach For System Operation Services," in Web Services (ICWS), 2017 IEEE Int'l Conf. On. IEEE, 2017, Pp. 700–707.

A New Nonparametric Test For Testing Equality of Locations Against Umbrella Alternatives

Bülent ALTUNKAYNAK, Hamza GAMGAM, Merve BAĞÇACI

Department of Statistics, Gazi University, TURKEY

Abstract: In this study, a nonparametric new test is proposed to test the hypothesis of equality of locations against umbrella alternatives. The Shan test for ordered alternatives is adapted to the umbrella alternatives. This test can be considered as an extension of the sign test and the Wilcoxon signed rank test. By a comprehensive simulation study, the proposed test is compared with the Mack-Wolfe and Hettmansperger and Norton tests in terms of type I error rate and power. The simulation results showed that all tests ensured the Bradley's robustness criteria for type I error rate. The power comparison results indicated that the proposed test gives better results than the other tests.

Keywords: Umbrella Alternative, Nonparametric Test, Monte Carlo, Simulation

REFERENCES

[1] J. T. Terpstra, "The Asymptotic Normality And Consistency of Kendall's Test Against Trend, When Ties Are Present İn One Ranking," (İn En), Indagationes Mathematicae, Vol. 14, No. 3, Pp. 327-333, 1952-01-01 1952.

[2] A. R. Jonckheere, "A Distribution-Free K-Sample Test Against Ordered Alternatives," Biometrika, Vol. 41, No. 1/2, Pp. 133-145, 1954.

[3] V. J. Chacko, "Testing Homogeneity Against Ordered Alternatives," (in En), Ann. Math. Statist., Vol. 34, No. 3, Pp. 945-956, 1963/09 1963.

[4] M. L. Puri, "Some Distribution-Free K-Sample Rank Tests of Homogeneity Against Ordered Alternatives," Communications On Pure And Applied Mathematics, Vol. 18, No. 1-2, Pp. 51-63, 1965.

[5] R. B. May And P. R. Konkin, "A Nonparametric Test of An Ordered Hypothesis For K Independent Samples," Vol. 30, No. 2, Pp. 251-257, 1970.

[6] R. E. Odeh, "On Jonckheere's K-Sample Test Against Ordered Alternatives," Technometrics, Vol. 13, No. 4, Pp. 912-918, 1971/11/01 1971.

[7] J. Cuzick, "A Wilcoxon-Type Test For Trend," Vol. 4, No. 1, Pp. 87-90, 1985.

[8] T. P. Hettmansperger And R. M. Norton, "Tests For Patterned Alternatives in K-Sample Problems," Journal Of The American Statistical Association, Vol. 82, No. 397, Pp. 292-299, 1987/03/01 1987.

[9] F. Beier And H. Buning, "An Adaptive Test Against Ordered Alternatives," (in English), Computational Statistics & Data Analysis, Vol. 25, No. 4, Pp. 441-452, Sep 30 1997.

[10] M. Neuhäuser, P.-Y. Liu, And L. A. Hothorn, "Nonparametric Tests For Trend: Jonckheere's Test, A Modification And A Maximum Test," Biometrical Journal, Vol. 40, No. 8, Pp. 899-909, 1998.

[11] H. Buning And W. Kossler, "Adaptive Tests For Umbrella Alternatives," (in English), Biometrical Journal, Vol. 40, No. 5, Pp. 573-587, 1998.

[12] J. T. Terpstra And R. C. Magel, "A New Nonparametric Test For The Ordered Alternative Problem," (in English), Journal Of Nonparametric Statistics, Vol. 15, No. 3, Pp. 289-301, Jun 2003.

[13] C. H. Chang And C. H. Yen, "A Nonparametric Test For The Ordered Alternative Based On Fast Discrete Fourier Transform Coefficient," (in English), Journal of Testing And Evaluation, Vol. 39, No. 6, Pp. 1131-1143, Nov 2011.

[14] J. T. Terpstra, C. H. Chang, And R. C. Magel, "On The Use of Spearman's Correlation Coefficient For Testing Ordered Alternatives," (in English), Journal Of Statistical Computation And Simulation, Vol. 81, No. 11, Pp. 1381-1392, 2011.

[15] C. H. Chang, C. C. Chin, W. W. Yu, And Y. Y. Huang, "Using The Bernoulli Trial Approaches For Detecting Ordered Alternatives," (in English), Bmc Medical Research Methodology, Vol. 13, Dec 5 2013.

[16] G. G. Shan, D. Young, And L. Kang, "A New Powerful Nonparametric Rank Test For Ordered Alternative Problem," (in English), Plos One, Vol. 9, No. 11, Nov 18 2014.

[17] A. Gaur, "A Class of K-Sample Distribution-Free Tests For Location Against Ordered Alternatives," (in English), Communications in Statistics-Theory And Methods, Vol. 46, No. 5, Pp. 2343-2353, 2017.

[18] G. A. Mack And D. A. Wolfe, "K-Sample Rank-Tests For Umbrella Alternatives," (in English), Journal of The American Statistical Association, Vol. 76, No. 373, Pp. 175-181, 1981.

[19] W. Kossler, "Some C-Sample Rank Tests Of Homogeneity Against Umbrella Alternatives With Unknown Peak," (in English), Journal Of Statistical Computation And Simulation, Vol. 76, No. 1, Pp. 57-74, Jan 2006.

[20] A. Gaur, K. K. Mahajan, And S. Arora, "New Nonparametric Tests For Testing Homogeneity of Scale Parameters Against Umbrella Alternative," (in English), Statistics & Probability Letters, Vol. 82, No. 9, Pp. 1681- 1689, Sep 2012.

[21] N. Z. Shi, "Rank Test Statistics For Umbrella Alternatives," (in English), Communications in Statistics-Theory And Methods, Vol. 17, No. 6, Pp. 2059-2073, 1988.

[22] A. Cohen And H. B. Sackrowitz, "Tests For The Umbrella Alternative Under Normality," (in English), Communications İn Statistics-Theory And Methods, Vol. 25, No. 11, Pp. 2807-2817, 1996.

[23] G. H. Pan, "Distribution-Free Tests For Umbrella Alternatives," (in English), Communications in Statistics-Theory And Methods, Vol. 25, No. 12, Pp. 3185-3194, 1996.

[24] B. A. Hartlaub And D. A. Wolfe, "Distribution-Free Ranked-Set Sample Procedures For Umbrella Alternatives in The M-Sample Setting," (in English), Environmental And Ecological Statistics, Vol. 6, No. 2, Pp. 105-118, Jun 1999.

[25] A. J. Hayter And W. Liu, "A New Test Against An Umbrella Alternative and The Associated Simultaneous Confidence İntervals," (İn English), Computational Statistics & Data Analysis, Vol. 30, No. 4, Pp. 393-401, Jun 28 1999.

[26] W. C. Lee And Y. I. Chen, "Weighted Kaplan-Meier Tests For Umbrella Alternatives," (in English), Annals of The Institute Of Statistical Mathematics, Vol. 53, No. 4, Pp. 835-852, Dec 2001.

[27] R. C. Magel And L. Qin, "A Non-Parametric Test For Umbrella Alternatives Based on Ranked-Set Sampling," (İn English), Journal Of Applied Statistics, Vol. 30, No. 8, Pp. 925-937, Oct 2003.

[28] P. Singh And W. Liu, "A Test Against An Umbrella Ordered Alternative," (in English), Computational Statistics & Data Analysis, Vol. 51, No. 3, Pp. 1957-1964, Dec 1 2006.

[29] M. Alvo, "Nonparametric Tests Of Hypotheses For Umbrella Alternatives," (in English), Canadian Journal Of Statistics-Revue Canadienne De Statistique, Vol. 36, No. 1, Pp. 143-156, March 2008.

[30] S. V. Bhat And A. B. Patil, "A Class of K-Sample Distribution-Free Tests For The Umbrella Alternatives," (in English), International Journal Of Agricultural And Statistical Sciences, Vol. 4, No. 1, Pp. 53-61, Jun 2008.

[31] D. Basso And L. Salmaso, "A Permutation Test For Umbrella Alternatives," (in English), Statistics And Computing, Vol. 21, No. 1, Pp. 45-54, Jan 2011.

[32] B. Altunkaynak And H. Gamgam, "Comparing The Performance of Nonparametric Tests For Equality Of Location Against Ordered Alternatives," Communications in Statistics - Simulation And Computation, Pp. 1-22, 2019.

[33] J. V. Bradley, "Robustness?, " British Journal of Mathematical And Statistical Psychology, Vol. 31, No. 2, Pp. 144-152, 1978.

Smart Agriculture Applications with IoT

Mervenur SAĞLAM, Turgut ÖZSEVEN

Department of Mechatronic Engineering, Department of Computer Engineering

Tokat Gaziosmanpasa University, TURKEY

Abstract: Smart agriculture is the correct and economical use of resources, product control, and production to be more ergonomic in order to increase production efficiency in agriculture. The concept of internet of things that are widely used in conjunction with Industry 4.0 has started to be applied in agriculture. The main objective is to apply the automation systems that provide communication among themselves to agriculture in order to increase production efficiency. In this study, the current state of the smart agricultural systems with IoT has been investigated. As a result, the risks that may occur in production can be predicted in a short time with smart agriculture. Thus, proper and wasteful use of the resources required for agriculture is foreseen.

Keywords: Smart Agriculture, Industry 4.0, Internet of Things

Introduction

The Internet of Things (IoT) is the communication of the interconnected objects over the network without the need for people through mechanical and digital machines. In 1991, a study was conducted at Cambridge University to track how much coffee was left in the coffee machine without going to the machine [1]. In this study, the photo of the coffee pod was automatically taken three times per minute and these photographs were transferred to the computer. Thus, the amount of coffee used and the amount of coffee that was decreasing was followed in real time. This study can be considered as the first steps of IoT [1].

Kevin Ashton, co-founder of MIT Auto-ID Center, first spoke about IoT in 1999 in Procher & Gamble (P & G). The first industrial revolution included mechanical production facilities between the 18th and 19th centuries. After 1870, the mass production of electrical energy was called the second industrial revolution, while the third industrial revolution was called the digital

revolution in 1970. The basic structure of the 4th industrial revolution has created a new concept with the existence of the internet concept [2].

The Industrial 4.0 or 4th Industrial Revolution was first used in 2011 at the Hannover Fair in Germany. Industry 4.0, by improving living standards, technology, science, industrial automation to bring together all the advanced applications to provide communication with sensors [3]. The historical development of industry is given in Figure 1.

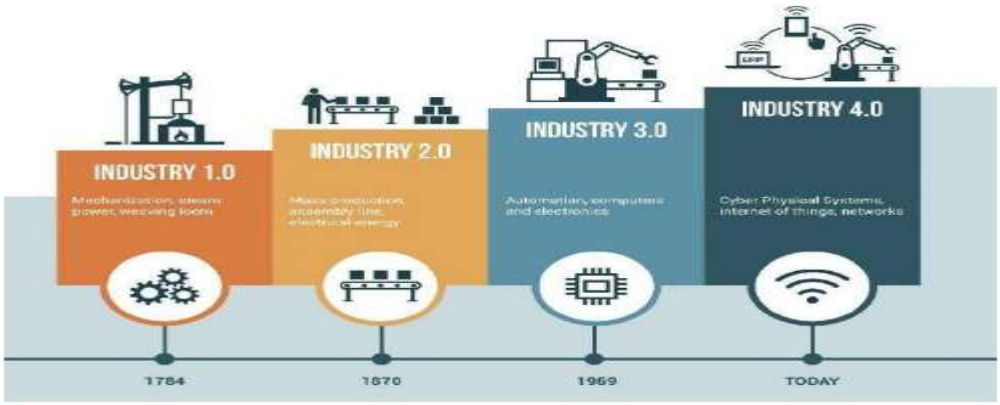


Fig. 1. Historical Development of Industry [4]

The technological developments in Industry 4.0 are given in Figure 2.

Fig. 2. Industry [5]

When Figure 2 is examined, Industry 4.0 includes many current technologies from autonomous systems to cyber security, from virtual reality to simulation. The interaction of objects through the Internet contributes to the user's ability to intervene in the device, the system to be transformed into a smart state, to reduce the cost, to minimize the amount of work to be done and thus to increase the profit.

In IoT, sensors, observation devices, chips, remote control systems are defined as objects. That is, if a device is an object, it is also assumed to be intelligent. Thus, the device has a unique ID. The sensor data is analyzed and transmitted via IDs to the device via the network. Nowadays, IoT is used in many fields such as smart home automation, smart cities, industrial controls, energy efficiency, health services, military applications. IoT sectoral distribution is given in Figure 3.

Fig. 3. IoT Sectoral Distribution [6].

Smart Agriculture

With the development of technology; terms such as intelligent agriculture, precision agriculture, digital agriculture, farm management software are used. The most well-known smart agriculture is the system of harvesting information. Smart agriculture and IoT provide remote access to data such as remote control of the system, monitoring of moisture content, harvest time.

Innovations in the 4th industrial revolution have become a turning point in technology. Cloud technology in smart agriculture systems, unmanned aerial vehicles, humidity-temperature-pressure sensors that can record all the land allow remote control of the system. Thus, it is aimed

to profit from time and product efficiency. The negativity of traditional agriculture is eliminated by this transformation in agricultural production [4]. If the objectives of smart agriculture are examined; to reduce the consumption of chemical materials, to minimize the damage to the environment, to be able to easily archive the registration information of the produced product, to obtain a higher quality of the produced product and to obtain a high quantity of product, to determine the amount of production and to make a regular production.

When we summarize a general definition of smart agriculture; It includes the intervention that can be done by taking into consideration the criteria that will be necessary for the needs that differ in terms of location and time in the field where the agricultural operator makes an agricultural production. It is difficult for farmers to evaluate the changes in heterogeneous agricultural lands and to take precautions beforehand, and in some cases it is impossible. However, with the development of technology, it can provide quantitative solutions to such problems. In this way, it provides the possibility to make agricultural areas more efficient with the exception of traditional methods [7, 8]. With the internet of things, the climate conditions in smart greenhouses can be controlled by sensors and can be recorded. Thus, it can be interfering with remote control of data instead of manually interfering.

Smart Agriculture With Iot in Turkey

Agriculture is one of the key areas that contribute to economic growth for Turkey. It is necessary to maintain this growth by increasing production, quality and yield in agriculture. The proper use of technology in the right place reduces the risks of agricultural products for operators. Correct use of technology brings together the idea of agriculture. By using the IoT and machines and intelligent systems, the least energy efficiency, water, seed, fertilizer and other agricultural products are minimized.

Even if the producer knows little about the soil structure of the field, it can produce different amounts of product on the same soil. The requirements of the cultivated product are applied in the same amount on all sides of the land. For this reason, with the changing and developing technology, timely irrigation, fertilization, spraying system will facilitate the production process and prevent environmental pollution. After mechanization era in Turkey, devices can communicate with each other, the objects that can operate synchronously is set right

timing by reducing the workload. In the field of agriculture, IoT systems can be used to manage the process from production, storage and distribution of the plant.

One of the most important factors in agriculture is meteorological results. In the air station systems developed with IoT; wind speed control, wind direction, temperature, pressure and humidity values are obtained through the sensors to obtain early intervention in situations where the risk is reduced by reducing the risk. An exemplary model of the processes that may be involved in smart agriculture is given in Figure 4.

Fig. 4. Smart Agriculture [10]

The emphasis on smart agriculture in Turkey is increasing every day. State aids and platforms have been established for this purpose (Smart Agricultural Platform). In this way, it is aimed to easily detect and analyze the developments in the field of intelligent agriculture. In this respect, early detection of potential future problems can be obtained [7].

Turkey intelligent agricultural areas, planning and pre-investment period. Domestic technology manufacturers need to be supported and supported with innovation systems. Together with smart agriculture, the young population will be able to obtain products with easier methods and thus to agriculture. The example of a remote-controlled smart greenhouse is given in Figure 5.

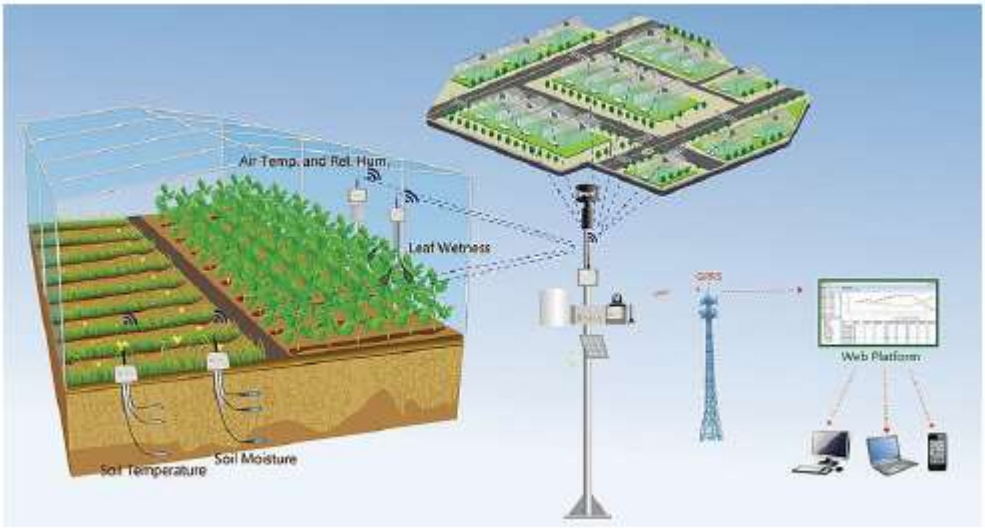


Fig. 5. Smart Greenhouse Remote Monitoring Systems [10]

In Turkey, private firms, state-supported institutions, and smart agriculture is continuing R & D efforts by individual users [4].

Nowadays, agricultural machinery has begun to integrate with intelligent technologies. From the harvesting of the harvesting tractor, everything from the harvesting to the product output in the field is recorded and big data are obtained. These data can be analyzed, and prospective estimations can be made. With the modernization of agricultural machinery, data such as GPS, wireless communication and snapshot can be obtained easily. An example of a greenhouse imaging system is given in Figure 6.

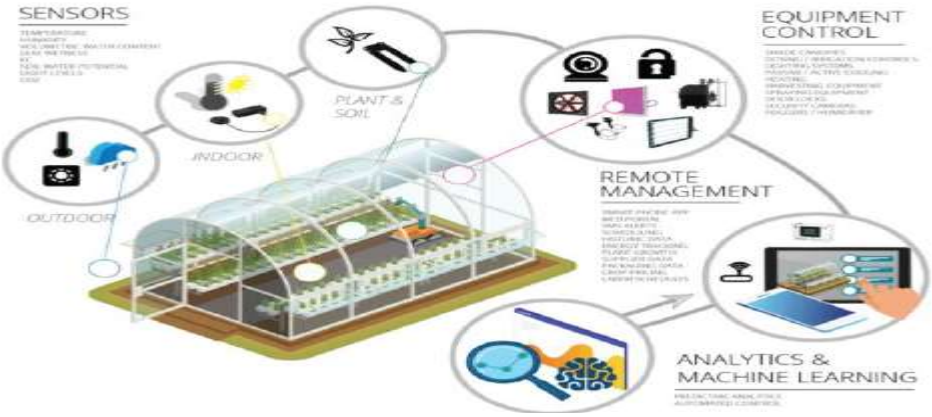


Fig. 6. Greenhouse Monitoring [11]

As can be seen from Figure 6, all information in the greenhouse is taken wirelessly and can be followed instantly from different platforms.

Results

The agriculture in Turkey is one of the biggest economic livelihoods. The importance given to agriculture in Turkey with Industry 4.0 is increasing every day.

Intelligence of everything makes the use of sensors attractive in agricultural areas. In order to obtain efficient product in agriculture, ambient values are measured and measured with sensors. These data reduce the risks of agricultural products.

Considering that environmental pollution and organic food production are important and that water resources are still inadequate, sensitive agriculture production should be transformed into sustainable agriculture. Turkey is also satisfying the necessary infrastructure in this area, manufacturers work should be done to make them conscious. Conscious producers bring conscious consumer understanding.

REFERENCES

[1] https://www.karel.com.tr/blog/nesnelerin-interneti-nasil-dogdu, 27.03.2019.

[2] Drath, R., & Horch, A. (2014). Industrie 4.0: Hit or hype? [Industry forum]. IEEE industrial electronics magazine, 8(2), 56-58.

[3] https://iiot-world.com/connected-industry/ industrie-4-0-is-all-about- speed, 25.03.2019.

[4]https://www.researchgate.net/publication/326550785_ Endustri_ 40_Ile_ Akilli_ Tarima_Gecis, 25.03.2019.

[5] Ünal, İ., & Topakcı, M. (2013). Tarımsal Üretim Uygulamalarında Bulut Hesaplama (Cloud Computing) Teknolojisi. Akademik Bilişim Konferansı-AB, 23-25.

[6] https://growthenabler.com/flipbook/pdf/IOT%20Report.pdf, 11.04.2019.

[7] http://www.akillitarim.org/tr/-atprapor - 24.03.2019.

[8] Topakcı, M. ve Ünal, İ. "Hassas Tarımda Değişken Oranlı Uygulamalar", Tarımsal Mekanizasyon 26. Ulusal Kongresi, 22-23 Eylül, Hatay, Türkiye, (2010).

[9] https://semiengineering.com/farming-goes-high-tech/, 01/04/2019

[10] https://www.postscapes.com/smart-greenhouses/, 02/04/2019

[11] https://www.caipos.com/news/irriwave/, 02/04/2019

Teknoloji Kabul Modeli Kullanarak Netflix Platformu Kullanma Maksadının Belirleyicileri

Ufuk CEBECİ, Oğuzhan İNCE

Department of Industrial Engineering, Istanbul Technical University, TURKEY

Abstract: Amazon Prime, Hulu, Apple TV, Puhutv, BluTV, Turkcell TV Plus vb. gibi digital platformlar insanlar üzerindeki etkisini artırmaya devam ediyor ve Netflix bunlardan birisidir. Netflix insanların seyretmek istediği film ve dizileri belirleyip onlara en kısa yoldan ulaştırmayı hedefleyen bir platformdur. Bu platform günümüzde çok popüler hale gelmiştir ve sinema sektörüyle yarışır seviyede bulunmaktadır. Netflix, veri madenciliğini etkili bir şekilde kullanarak insanların neyi izlemeyi sevdiğini bilmektedir. Ayrıca, Netflix aşırı ve gereksiz verilerden kurtularak büyük verilerin değerli bilgilere dönüştürülürken daha net olmasını sağladı. Bu bilgiler ışığında ülkeler için ayrı dizi ve filmlerin çekilmesine öncülük etmiş, izleyicilerin sevdiği aktör ve yönetmenleri bir araya getirerek, çektiği filmlerin izlenilirliğini arttırmış, böylece de Netflix markasını dünyaya tanıtarak izleyicileri kendine bağlamayı başarmıştır. Bizim bu çalışmadaki amacımız Netflix kullanıcılarının davranışlarını incelemek ve açıklamaktır. Bunu sağlamak için anket çalışması yapılmış olup, Netflix platformunun kullanma maksadının belirleyicilerini araştırmak için teknoloji kabul modeli kullanılmıştır.

Keywords: Veri Madenciliği Uygulamaları, Teknoloji Kabul Modeli, Netflix Kullanıcılarının Davranışları

REFERENCES

[1] Amin, M. S., Chiam, Y. K., & Varathan, K. D. (2019). Identification of Significant Features And Data Mining Techniques in Predicting Heart Disease. Telematics And Informatics, 36, 82-93.

[2] Bailey, A. A., Pentina, I., Mishra, A. S., & Ben Mimoun, M. S. (2017). Mobile Payments Adoption By US Consumers: An Extended TAM. International Journal Of Retail & Distribution Management, 45(6), 626-640.

[3] Boonsiritomachai, W., & Pitchayadejanant, K. (2017). Determinants Affecting Mobile Banking Adoption By Generation Y Based on The Unified Theory Of Acceptance And Use of Technology Model Modified By The Technology Acceptance Model Concept. Kasetsart Journal of Social Sciences.

[4] ÇELİK, H., & İPÇİOĞLU, İ. (2006). Gönüllü Teknoloji Kabulü: İnternet Kullanimini Benimseme Davranişi Üzerine Bir Araştırma. Hacettepe Üniversitesi İktisadi ve İdari Bilimler Fakültesi Dergisi, 24(1), 111-159.

[5] Davis, F. D. (1989). Perceived Usefulness, Perceived Ease of Use, and User Acceptance of Information Technology. MIS Quarterly, 319-340.

[6] Dutot, V., Bhatiasevi, V., & Bellallahom, N. (2019). Applying The Technology Acceptance Model İn A Three-Countries Study of Smartwatch Adoption. The Journal Of High Technology Management Research.

[7] Estriegana, R., Medina-Merodio, J. A., & Barchino, R. (2019). Student Acceptance of Virtual Laboratory And Practical Work: An Extension of The Technology Acceptance Model. Computers & Education, 135, 1-14.

[8] Kim, C., Mirusmonov, M., & Lee, I. (2010). An Empirical Examination of Factors Influencing the Intention To Use Mobile Payment. Computers in Human Behavior, 26(3), 310-322.

[9] Netflix 2019. Https://Media.Netflix.Com/En/About-Netflix.

[10] Ozturk, A. B., Bilgihan, A., Nusair, K., & Okumus, F. (2016). What Keeps The Mobile Hotel Booking Users Loyal? Investigating The Roles Of Self-Efficacy, Compatibility, Perceived Ease of Use, And Perceived Convenience. International Journal of Information Management, 36(6), 1350-1359.

[11] Pavlou, P. A. (2003). Consumer Acceptance Of Electronic Commerce: Integrating Trust And Risk With the Technology Acceptance Model. International Journal of Electronic Commerce, 7(3), 101-134.

[12] Scherer, R., Siddiq, F., & Tondeur, J. (2019). The Technology Acceptance Model (Tam): A Meta-Analytic Structural Equation.

[13] Subramaniam, M., Iyer, B., & Venkatraman, V. (2019). Competing in Digital Ecosystems. Business Horizons, 62(1), 83-94.

[14] Verma, P., & Sinha, N. (2018). Integrating Perceived Economic Wellbeing to Technology Acceptance Model: The Case of Mobile Based Agricultural Extension Service. Technological Forecasting and Social Change, 126, 207-216.

[15] Verma, S., Bhattacharyya, S. S., & Kumar, S. (2018). An Extension of The Technology Acceptance Model in The Big Data Analytics System Implementation Environment. Information Processing & Management, 54(5), 791-806.

[16] Wallace, L. G., & Sheetz, S. D. (2014). The Adoption Of Software Measures: A Technology Acceptance Model (TAM) Perspective. Information & Management, 51(2), 249-259.

[17] Wu, J. H., & Wang, S. C. (2005). What Drives Mobile Commerce? An Empirical Evaluation of The Revised Technology Acceptance Model. Information & Management, 42(5), 719-729.

[18] Yang, Y., & Wang, X. (2019). Modeling The İntention To Use Machine Translation For Student Translators: An Extension of Technology Acceptance Model. Computers & Education, 133, 116-126.

[19] Yoo, Y., Henfridsson, O., & Lyytinen, K. (2010). Research Commentary—The New Organizing Logic of Digital İnnovation: An Agenda For Information Systems Research. Information Systems Research, 21(4), 724-735.

[20] Zainab, B., Awais Bhatti, M., & Alshagawi, M. (2017). Factors Affecting E-Training Adoption: An Examination Of Perceived Cost, Computer Self-Efficacy and The Technology Acceptance Model. Behaviour & Information Technology, 36(12), 1261-1273.

Meta-Heuristic Methods Used in Optimization of SVM Learning Parameters

Zübeyir ÖZKORUCU, Turgut ÖZSEVEN

Department of Mechatronic Engineering, Department of Computer Engineering

Tokat Gaziosmanpasa University, TURKEY

Abstract: Support vector machine is an effective machine learning method based on statistical learning theory and used for classification problems. The optimization of the parameter is very important in order to increase the classification accuracy. Meta-heuristic methods are one of the main optimization approaches that can be applied in this context and have been used frequently for parameter optimization in recent years. These methods are generally particle swarm optimization, genetic algorithm, grid search method, differential evolution algorithm, ant colony optimization. In this study, support vector machine parameter optimization studies between 2010-2019 were investigated. According to the results of these studies, it was observed that parameter optimization through meta-heuristic methods significantly increased the rate of classification accuracy of classifier and significantly reduced the workload.

Keywords: Support Vector Machines, Meta-Heuristic Methods, Parameter Optimization, Classification Accuracy

Introduction

The classification process plays an important role in machine learning and data mining. One of the machine learning algorithms is Support Vector Machine (SVM). Based on statistical learning theory and structural risk minimization, SVM is an effective method for solving pattern recognition and classification problems. The basic idea in SVM is to determine a hyperplane that separates the data appropriately from each other by maximizing the spacing of the nearest vectors in the data set (Fig. 1-left) [1]. Although the number of hyperplanes may be more than one for the data set which can be classified as linear (Fig. 1-right), only one of them makes the maximum distance between the two classes, which is the hyperplane that the SVM must find.

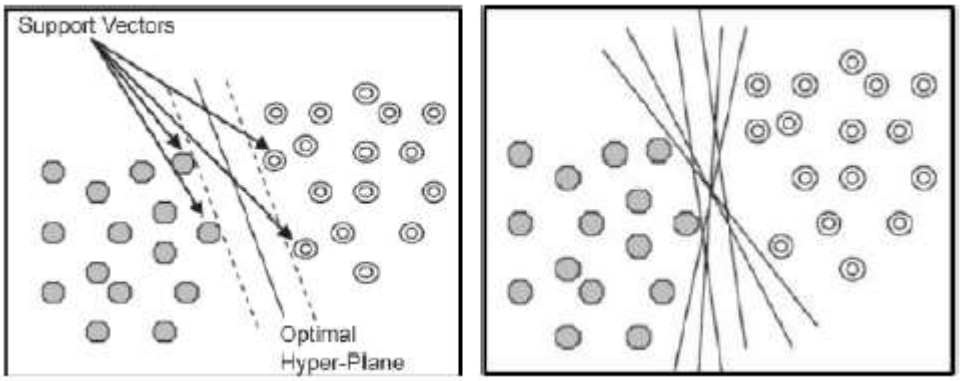


Fig. 1. Optimal Hyperplane (Left) and Possible Hyperplanes (right) [2]

The studies to get the most appropriate result from a system under certain constraints are called optimization. In other words, it is the process of determining the most appropriate values of the variables of the problem in order to minimize or maximize the target function. The first thing to do in the optimization process is to determine the decision parameters. Then the limiting functions that define the values that the parameters cannot take, a cost function to be minimized depending on the parameters or a profit function to be maximized are defined [3].

Meta-heuristic methods are algorithms that are proposed by inspiration from nature that produces effective and appropriate solutions when traditional optimization methods produce unacceptable solutions. These methods examine the search area effectively and efficiently and conduct a solution research in a stochastic way. They set out from the set of solutions created in each iteration and produce new solutions. However, they don't guarantee the global optimum solution [4-6].

In this study, the studies conducted between 2010-2019 for the optimization of SVM parameters were examined and analyzed. Analysis results are described in Section 3.

Used Optimization Methods For SVM

Meta-heuristic algorithms have been rapidly progressing in the literature in the last 20 years because of their successful results [7-8]. Genetic Algorithm (GA) [9], Particle Swarm Optimization (PSO) [10], Bat Algorithm (BA) [11], Differential Evolution Algorithm (DEA) [12], Simulated Annealing (SA) [13] and Ant Colony Algorithm [14] are the most popular meta-heuristic algorithms. These methods have advantages and disadvantages compared to each other in terms of calculation complexity, classification accuracy rate and working time. The distribution

of meta-heuristic methods used in the studies examined and analyzed in this study is given in Fig. 2.

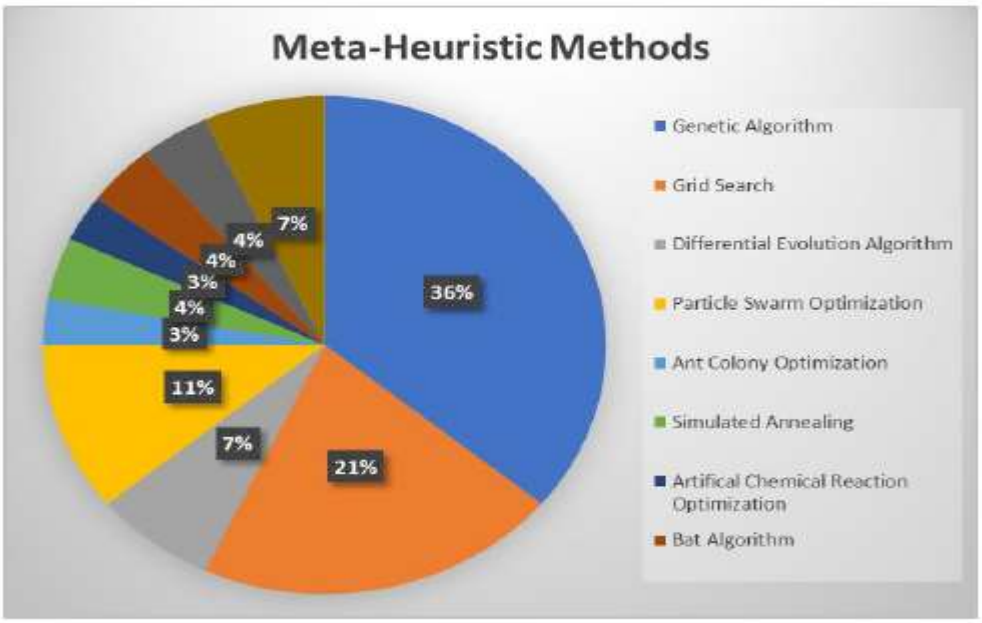


Fig. 2. Meta-heuristic Methods Used in Studies Examined

GA, which was proposed by John Holland (1975), is one of the evolutionary approaches used in problems that are difficult to solve with traditional methods [9]. This algorithm is based on the principle of protecting the lives of good generations and the destruction of the bad generations.

PSO, which was proposed R.C. Eberhart and J.Kennedy in 1995, is an optimization algorithm based on the behavior of bird flocks trying to find food [16]. Each individual in the swarm is called a particle and particles form population. Each of these individuals is the candidate solution. In this method, each particle adjusts its position to the best position by taking advantage of its previous experience and individual with best position in swarm. Particles usually have a better position than their previous position in their next position.

PSO has a fast convergence rate, but it is very easy to drop to the local optimum and has premature convergence problems. Therefore, it is difficult to find the best solution in the solution space. To overcome this problem, a mutation process was added to the PSO. Particles can be reinitialized and updated to set the population search area after each iteration thanks to mutation operation. In addition, this process prevents premature convergence by changing the position and velocity of the particles and provides population diversity [17].

DEA, which was proposed by Price and Storn in 1995, is a population-based algorithm [18]. This method gives effective results especially in problems with continuous data and its operators and operation are similar to GA [12][18][19][20]. However, it is tried to improve the performance of the solution by changing the usage of operators in DEA. The crossover, selection and mutation operators used in GA are also used in this method. However, unlike GA, every operator is not sequentially applied to the whole population. Moreover, chromosomes are taken into consideration in this method and a new individual is obtained by using three other randomly selected chromosomes. The fitness values between the existing chromosome and the new chromosome obtained is compared and better one is transferred to the next population.

Ant Colony Optimization (ACO), which is created by Marco Dorigo, is a method inspired by the methods of finding the shortest path between nests and food sources of ant colonies [21][22]. Dorigo has applied this method, which is a population based approach, to the Travelling Salesman Problem and observed that ACO is highly effective in solving this problem. This method is widely used in the solution of today's optimization problems [14][21].

SA is a probability-based method proposed by Kirkpatrick et al. This method is called simulated annealing because it is based on annealing process of solids. The annealing is to cool the material slowly until it is crystallized after being heated to a certain maximum degree. If the cooling is done appropriately, the crystal structure is very regular and a super cage structure is obtained. If the cooling process is carried out very quickly, irregularities and disturbances occur in the crystal structure. Therefore, the cooling process is very important. This method is especially used in optimization of combinatorial problems which cannot be shown with mathematical models [23].

BA is a population-based optimization algorithm based on the direction and distance determining behavior of the object by utilizing the echoing of sound called echolocation [24]. Echolocation is a type of sonar, which many bat species use to communicate, to move without hitting obstacles by perceiving objects around them, to determine the distance between itself and its hunts [24][25][26]. All living things emit signals at a given frequency [26]. The bats listen to the echo that occurs after they emit the signal and they determine the position of the objects and prey around them by analyzing these echoes [26][27].

Artificial Chemical Reaction Optimization Algorithm (ACROA) is a probabilistic optimization algorithm based on chemical reactions. During the chemical reaction, while bond rupturing between some substances occurs, new bonds are formed between certain substances and

chemical changes occur. Thus, the energy and structure of reactants change. Furthermore, new molecules formed as a result of the reaction can be used as reactants in another reaction. The new molecules formed can be converted into their first reactants in two-way reactions. ACROA is a method which is developed according to the types of chemical reactions and requires less parameters [28][29][30]. Entropy for a maximization problem and enthalpy for a minimization problem can be used as objective function in this method [28][30].

Elephant Herd Optimization (EHO) is one of the swarm intelligence algorithms proposed by Wang, Den, Gao and Coelho in 2016. This method is based on the basic herd behaviors of elephant groups. The elephant population consists of clans and each clan has a fixed number of elephants. The elephants in each clan continue to live under the leadership of a matriarch. In every generation, male elephants leave their clans and live away from them [31][32]. When we associate these behaviors of the elephant herd with the optimization method, each elephant represents a candidate solution. The population is divided into subgroups and forms the clans and the elephant with the best fitness value in each clan is called the matriarch. After each iteration in the algorithm, the worst elephant in the clan leaves from the herd and identifies a new location for itself. Each elephant moves according to the clan's matriarch [32].

Grid search (GS) is a traditional method for parameter optimization. This method tries to find all combinations with brute force and makes a comprehensive search. GS requires creating two sets called learning rate and number of layers. The method trains the algorithm for all combinations using these two sets and uses the Cross Validation (CV) method to measure performance. Although GS is a simpler algorithm compared to other optimization algorithms, the runtime is quite long, especially when the data set is too large [33].

Results of Existing Studies For SVM Optimization

According to the results of the studies examined within this study, the optimization of the parameters in SVM has different effects on the results. This difference varies depending on the optimized parameter and the method used.

Each kernel function has different kernel parameters in SVM. The optimization of learning parameters in SVM is very important. Kernel parameters are learning parameters, too. One of learning parameters is the penalty coefficient (C) that determines the trade-off between maximization between class distances and the minimization of fault tolerance. When the value C is greater, SVM results in fewer training errors and narrower distances, whereas it leads to more

distance and more educational errors in the small values [34]. However, the education data is also very important in the high or low education error rate. This problem cannot be solved by only the large value of C. Therefore, it is important to find an appropriate C value for the sample data in the problem. Thus, the training model can establish a balance between sample data and educational error. Another parameter is the kernel parameter that determines the complexity of the sample data [35]. This parameter can both reduce complexity and guarantee classification accuracy. In short, learning parameters have a great effect on the generalization ability and efficiency of SVM [36]. The optimization methods used in the studies examined in this article optimize these learning parameters.

As a result of the studies examined, GA is a more suitable method than GS in optimization of SVM parameters [37-39]. GS method results in a higher classification accuracy than original SVM model [40], but when GA and GS are compared, GA increases the classification accuracy more and the running time is shorter [37-39]. Although GA is a more favorable method than GS, ACO [41], DE [42], ACROA [43], BA [44], EHO [37], M-PSO [45], PSO [46], SA [47] methods are more successful than the GA method. DE [42], BA [44], M-PSO [45], ACROA [43] methods are better than PSO in the studies examined. In the study, which has been used to optimize the SVM parameters by using M-PSO [17], M-PSO method prevented falling to local optimum and premature convergence problems of PSO.

Conclusion and Discussion

SVM is a powerful machine learning method used in classification problems. SVM parameter optimization research has increased in recent years because of the heavy computational complexity of SVM in training stage. Incorrect parameter selection in the SVM can affect the classification performance of the method and increase the computational complexity [8]. Therefore, researchers are looking for appropriate optimization algorithms and there are many studies in literature about this topic [36-43]. In this article, the studies of SVM parameter optimization between 2010-2019 was analyzed. According to the results of the analysis,

- Meta-heuristic methods used significantly increased the success of the classification accuracy of SVM [17,35,37-40,42,44-46,48-57].
- Most commonly used methods are GA and GS. However, ACO, SA, BA and PSO have started to become widespread in recent years. Because these methods compared with GA

and GS in the studied examined, the rate of classification accuracy of these methods was higher GA and GS [37,44].

- Although meta-heuristic methods increase the accuracy of classification, these methods take a long time to optimize the parameters in big data sets [58].
- The adjustment of the kernel parameter and penalty coefficient in the SVM has vital importance in increasing the classification accuracy.
- The selection of the kernel function in increasing the classification success of SVM is also very important.

New optimization algorithms are constantly introduced into the literature and these algorithms are used in parameter optimization of SVM. In future studies, it is aimed to gain a new DVM parameter optimization model by considering the success rates of the optimization algorithms used in the studies examined.

REFERENCES

[1] V.N. Vapnik, The Nature of Statistical Learning Theory, 2nd ed., 2000, Springer-Verlag: New York.

[2] T. Özseven, M. Ertürkler, M. Nurmuhammed, M. Gül and M. Harputluoğlu, "Quantifying the Necrotic Areas on Liver Tissues Using Support Vector Machine (SVM) Algorithm and Gabor Filters," International Symposium on Innovations in Intelligent Systems and Applications Innovations in Intelligent Systems and Applications, pp. 1-5, 2012.

[3] D. Karaboğa, Yapay Zekâ Optimizasyon Algoritmaları, 2nd ed., 2011, pp. 1-5.

[4] A. Onan, "Metaheuristic Methods and Their Application Areas," Journal of Çukurova University Faculty of Economics and Administrative Sciences, vol. 17,2013, pp.113-128.

[5] M. Yetkin, "Metasezgisel Algortimaların Jeodezi'de Kullanımı," Journal Of Geomatic Engineering Research, Vol. 1, 2016, Pp. 8-13.

[6] M. Yetkin, M. Berber, "Implementation of Robust Estimation İn GPS Networks Using The Artificial Bee Colony Algorithm," Earth Science Informatics, Vol. 7,2014, Pp. 39-46.

[7] X.-S. Yang, And I. Wiley, Engineering Optimization: An Introduction With Metaheuristic Applications, 2010.

[8] B. Tezcan, "Determination Of Support Vector Machine Parameters By Meta Heuristic Optimization Algorithms (MS Thesis)", The Graduate School Of Natural And Applied Science Of Selçuk University The Degree Of Master Of Science In Computer Engineering, 2018, Konya.

[9] J.H. Holland, Adaptation İn Natural And Artificial Systems. MIT Press, 1992.

[10] J. Kennedy And R. Eberhart, "Particle Swarm Optimization," Proceedings Of ICNN'95 - International Conference On Neural Networks, Vol. 1944, Pp. 1942-1948, Dec 1995.

[11] X.-S. Yang, "A New Metaheuristic Bat-Inspired Algorithm," Nicso 2010: Nature Inspired Cooperative Strategies For Optimization, Vol.284, 2010b, Pp.65-74.

[12] R. Storn, K. Price, "Differential Evolution - A Simple And Efficient Adaptive Scheme For Global Optimization Over Continuous Spaces," Journal Of Global Optimization, Vol. 23, 1995.

[13] S. Kirkpatrick, C.D Gelatt, M.P. Vecchi, "Optimization By Simulated Annealing," Science, Vol. 220, May 1983, Pp. 671-680.

[14] M. Dorigo, G.A. Di Caro, L.M. Gambardella, "Ant Algorithms For Discrete Optimization," Artificial Life, Vol. 5,1999, Pp. 137-172.

[15] M.N. Ab Wahab, S. Nefti-Meziani And A. Atyabi, "A Compherensive Review Of Swarm Optimization Algorithms," PLOS One, Vol. 23, 2015, Pp. 141-161.

[16] J. Kennedy And R. Eberhart, "Particle Swarm Optimization," Proceedings Of ICNN'95 - International Conference On Neural Networks, Vol. 1944, Pp. 1942-1948, Dec 1995.

[17] Z. He, S. Ding, B. Li, M. Yin And X. Zhang, "An İmproved Particle Swarm Optimization of Support Vector Machine Parameters For Hyperspectral İmage Classification", 12th IEEE Conference On Industrial Electronics And Applications (ICIEA), Pp. 499-503, 2017.

[18] R. Storn, Http://Www1.İcsi.Berkeley.Edu/~Storn/Code.Html, [20.04.2019]

[19] D.G. Mayer, B.P. Kinghorn, A.A Archer, "Differential Evolution – An Easy And Efficient Evolutionary Algorithm For Model Optimisation," Agricultural Systems, Vol. 83, March 2005, Pp. 315-328.

[20] T. Keskintürk, "Differential Evolution Algorithm," Istanbul Commerce University Journal Of Science, Vol. 5, 2006, Pp. 85-99.

[21] A.E. Şeneras, Ş. İnanç, "Ant Colony Optimization For Solving Vehicle Routing Problems, Papers on Social Science, 2017, Pp. 58- 67.

[22] M. Demircioğlu, "A Heuristic Approach To Vehicle Routing Problem And An Application (Doctoral Thesis)," Çukurova University Institute Business Administration Department, 2009, P. 163.

[23] S. Kirkpatrick, C.D Gelatt, M.P Vecchi, "Optimization By Simulated Annealing," Science, Vol. 220, May 1983 Pp. 671-680.

[24] L.L.K. Kamal, "Training Artificial Neural Networks Using Bat Optimization Algorithm (MS Thesis)," The Graduate School Of Natural And Applied Science Of Selçuk University The Degree Of Master Of Science In Computer Engineering, 2018, P. 62.

[25] M.B Fenton, "Bat Natural History And Echolocation," Bat Echolocation Research: Tools, Techniques And Analysis,2004, Pp. 2-6.

[26] S. Yılmaz, "Modification of Bat Algorithm on Unimodal, Multimodal and Shifted Numeric Optimization Problems (MSc Thesis), "Süleyman Demirel University Graduate School of Applied And Natural Sciences Department of Computer Engineering, 2014, P. 111.

[27] H.-U. Schnitzler, E.K. Kalko, "Echolocation By Insect-Eating Bats: We Define Four Distinct Functional Groups Of Bats And Find Differences in Signal Structure That Correlate With The Typical Echolocation Tasks Faced By Each Group," Bioscience, Vol. 51, July 2001, Pp. 557-569.

[28] B. Alatas, A.B. Özer, "Mining Of Generalized Interesting Classification Rules With Artificial Chemical Reaction Optimization Algorithm" Journal of The Faculty of Engineering And Architecture of Gazi University, Vol. 32, 2017, Pp. 101-118.

[29] B. Alatas, "ACROA: Artificial Chemical Reaction Optimization Algorithm For Global Optimization," Expert Systems With Applications, Vol. 38, September 2011, Pp. 13170-13180.

[30] B. Alatas, "A Novel Chemistry Based Metaheuristic Optimization Method For Mining of Classification Rules," Expert Systems With Applications, Vol. 39, September 2012, Pp. 11080-11088.

[31] G.-G. Wang, S. Deb, X.-Z. Gao, and L. D. S. Coelho, "A New Metaheuristic Optimisation Algorithm Motivated By Elephant Herding Behaviour," International Journal of Bio-Inspired Computation, vol. 8, 2016, pp. 394–409.

[32] H. Haklı, "Elephant Herding Optimization Using Multi-Search Strategy for Continuous Optimization Problems," Academic Platform Journal of Engineering and Science, vol. 7, 2019, pp. 261-268.

[33] Prabhu, https://towardsdatascience.com/understandinghyperparameters- and-its-optimisation -techniques-f0debba07568, 2018. [21.04.2019]

[34] T. Eitrich and B. Lang, "Efficient Optimization of Support Vector Machine Learning Parameters For Unbalanced Datasets," Journal of Computational and Applied Mathematics, vol. 196, November 2006, pp. 425-436.

[35] J. Han, W. Jiang, C. Dai and H. Ma, "The Design of Diabetic Retinopathy Classifier Based on Parameter Optimization SVM," 2018 International Conference on Intelligent Informatics and Biomedical Sciences (ICIIBMS), vol. 3, 2018, pp. 52-58.

[36] X. Zhang, X. Chen and Z. He, "An ACO-Based Algorithm For Parameter Optimization of Support Vector Machines," Expert Systems With Applications, vol.37, September 2010, pp. 6618-6628.

[37] E. Tuba, Z. Stanimirovic, "Elephant Herding Optimization Algorithm for Support Vector Machine Parameters Tuning," 9th International Conference on Electronics, Computers and Artificial Intelligence (ECAI), pp. 1-4, 2017.

[38] I. Syarif, A. Prugel-Bennett, G. Wills, "SVM Parameter Optimization Using Grid Search and Genetic Algorithm to Improve Classification Performance," Telkomnika, vol. 14, 2016, pp. 1502- 1509.

[39] A.V. Phan, M.L. Nguyen and L.T. Bui, "Feature Weighting And SVM Parameters Optimization Based on Genetic Algorithms For Classification Problems," Applied Intelligence, vol. 46, 2017, pp. 455-469.

[40] L. Grama, L. Tuns and C. Rusu, "On the Optimization of SVM Kernel Parameters for Improving Audio Classification Accuracy," 14th International Conference on Engineering of Modern Electric Systems (EMES), pp. 224-227, 2017.

[41] X. Zhang, X. Chen And Z. He, "An Aco-Based Algorithm For Parameter Optimization of Support Vector Machines," Expert Systems With Applications, Vol.37, September 2010, Pp. 6618-6628.

[42] L. Yang And Z. Xu, "Feature Extraction By Pca And Diagnosis of Breast Tumors Using Svm With De-Based Parameter Tuning," International Journal of Machine Learning And Cybernetics, Vol. 10, 2019, Pp. 591-601.

[43] H. Ao, J. Cheng, Y. Yang And T.K. Truong, "The Support Vector Machine Parameter Optimization Method Based On Artificial Chemical Reaction Optimization Algorithm And İts Application To Roller Bearing Fault Diagnosis," Journal Of Vibration And Control, Vol. 21, 2015, Pp. 2434-2445.

[44] A. Tharwat, A.E. Hassanien And B.E. Elnaghi, "A Ba-Based Algorithm For Parameter Optimization of Support Vector Machine," Pattern Recognition Letters, Vol. 93, 2017, Pp. 13-22.

[45] C.-N. Chen, T.T. Hoang And M.-Y. Cho, "Parameter Optimisation Of Support Vector Machine Using Mutant Particle Swarm Optimisation For Diagnosis of Metal-Oxide Surge Arrester Conditions" Journal Of Experimental & Theoretical Artifical Intelligence, Vol. 31, 2019, Pp. 163-175.

[46] M. Hric, M. Chmulík And R. Jarina, "Model Parameters Selection For Svm Classification Using Particle Swarm Optimization," Proceedings of 21st International Conference Radioelektronika 2011, Pp. 1-4, 2011.

[47] S. Wang, "Support Vector Machines Classification For Highdimentional Dataset," Fourth International Conference On Multimedia Information Networking And Security, Pp. 315-318, 2012.

[48] A. Özbeyaz, M.İ. Gürsoy, R. Çoban, "Regularization And Kernel Parameters Optimization Based on Pso Algorithm in Eeg Signals Classification With SVM," Ieee 19th Signal Processing And Communications Applications Conference, Pp. 399-402, April 2011.

[49] L. Wang, G. Xu, J. Wang, S. Yang, L. Guo And W. Yan, "Ga-Svm Based Feature Selection And Parameters Optimization For BCI Research," Seventh International Conference on Natural Seventh International Conference on Natural Computation, Vol. 1, Pp. 580- 583, 2011.

[50] L. Jun, D. Lixin And X. Ying, "Differential Evolution Based Parameters Selection For Support Vector Machine," Ninth International Conference on Computational Intelligence And Security, Pp. 284-288, 2013.

[51] X.Z. Li And J.M. Kong, "Application Of Ga–Svm Method With Parameter Optimization For Landslide Development Prediction," Natural Hazards And Earth System Sciences Discussions, Vol. 14, 2014, Pp. 5295-5322.

[52] B.M. Sherin, And M.H. Supriya, "Ga Based Selection And Parameter Optimization For An Svm Based Underwater Target Classifier," International Symposium On Ocean Electronics (Sympol), Pp. 1-7, 2015.

[53] M.Yang, W. Li,H. Zhang And H. Wang, "Parameters Optimization Improvement of Svm On Load Forecasting," 8th International Conference On Intelligent Human-Machine Systems And Cybernetics (Ihmsc), Vol. 02, Pp. 257-260, 2016.

[54] J. Manurung, H. Mawengkang And E. Zamzami, "Optimizing Support Vector Machine Parameters With Genetic Algorithm For Credit Risk Assessment," Journal Of Physics: Conference Series, Vol. 930, 2017.

[55] M. -Y. Cho And T.T. Hoang, "Feature Selection And Parameters Optimization of Svm Using Particle Swarm Optimization For Fault Classification İn Power Distribution Systems," Computational Intelligence And Neuroscience, Vol. 2017, 2017, P. 9.

[56] P. Tao, Z. Sun And Z.Sun, "An Improved Intrusion Detection Algorithm Based on Ga And Svm," Ieee Access, Vol. 6, 2018, Pp. 13624-13631.

[57] Z. Tao, L. Huiling, W. Wenwen And Y. Xia, "Ga-Svm Based Feature Selection and Parameter Optimization İn Hospitalization Expense Modeling," Applied Soft Computing, Vol. 75, February 2019, Pp. 323-332.

[58] Y. Liu, J. Du, "Parameter Optimization Of The Svm For Big Data," 8th International Symposium On Computational Intelligence And Design (Iscıd), Vol. 2, Pp. 341-344, 2015.

Determination Similarities of Basic Financial Indicators of Enterprises Included in the NASDAQ Index Using by Hierarchical Clustering Distance Methods

Ebru Şeyma KARAKOYUN

Industrial Engineering Department, Karabük University, TURKEY

Abstract: Many indicators are used in the analysis of the financial status of companies. Indicators are usually presented quarterly. However, some of the financial indicators of some companies are not given data. In this study, it is examined whether there will be other indicators that Show the same character instead of the data which is not present when the companies are evaluated as a whole. A hierarchical clustering approach was used to evaluate whether the financial indicators of 10 randomly selected companies traded on the NASDAQ stock exchange to have similar characteristics using distance methods. According to the results obtained, indicators with similar characteristics are listed.

REFERENCES

[1] Rai, P., Singh, S. (2010). A Survey of Clustering Techniques, Int. J. Com. Appl. 7 (12) 1–5.

[2] Ando, T. And Bai J. (2016). Clustering Huge Number of Financial Time Series: A Panel Data Approach With High-Dimensional Predictor And Factor Structures, Journal of The American Statistical Association, İn Press.

[3] Kakizawa, Y., Shumway, R.H., Taniguchi, M. (1998). Discrimination And Clustering For Multivariate Time Series, Journal of The American Statistical Association, 93:441, 328-340.

[4] Wang, X., Smith, K., Hyndman, R. (2006). Characteristic-Based Clustering For Time Series Data, Data Min. Knowl. Discov. 13 (3) 335–364.

[5] Pavlidis, N., Plagianakos, V.P., Tasoulis, D.K., Vrahatis, M.N. (2006). Financial Forecasting Through Unsupervised Clustering and Neural Networks, Oper. Res. 6 (2), 103–127.

[6] Scotto, M.G., Alonso, A.M., Barbosa, S.M. (2010). Clustering Time Series of Sea Levels: Extreme Value Approach, J. Waterway Port, Coastal, Ocean Eng. 136 (4), 215–225.

[7] Gorji Sefidmazgi, M., Sayemuzzaman, M., Homaifar, A., Jha, M.K., Liess, S. (2014). Trend Analysis Using Non-Stationary Time Series Clustering Based on The Finite Element Method, Nonlinear Process. Geophys. 21 (3), 605–615.

[8] Ramoni, M., Sebastiani, P., Cohen, P. (2000). Multivariate Clustering By Dynamics, in: Proceedings Of The National Conference On Artificial Intelligence, Pp. 633–638.

[9] Guo, C., Jia, H., Zhang, N. (2008). Time Series Clustering Based On ICA For Stock Data Analysis, in: Proceedings Of 4th International Conference On Wireless Communications, Networking And Mobile Computing, Wicom '08, 2008, Pp. 1–4.

[10] Guan, H., Jiang, Q. (2007). Cluster Financial Time Series For Portfolio, in: ProceedingsoOf The International Conference on Wavelet Analysis And Pattern Recognition, Pp. 851–856.

[11] Wang, X., Mueen, A., Ding, H., Trajcevski, G., Scheuermann, P., Keogh, E. (2012). Experimental Comparison Of Representation Methods And Distance Measures For Time Series Data, Data Min. Knowl. Discov., P. Springer Netherlands.

[12] Lkhagva, B., Suzuki, Y.U., Kawagoe, K. (2006). New Time Series Data Representation ESAX For Financial Applications, in: Proceedings of 22nd International Conference on Data Engineering Workshops, Pp. 17–22.

[13] Petitjean, F., Ketterlin, A., Gançarski, P. (2011). A Global Averaging Method For Dynamic Time Warping, With Applications To Clustering, Pattern Recognit 44 (3), 678–693.

[14] Kakizawa, Y., Shumway, R.H., Taniguchi, M. (1998). Discrimination And Clustering For Multivariate Time Series, J. Am. Stat. Assoc 93 (441), 328–340.

[15] Focardi, S. M. (2001). Clustering Economic And Financial Time Series: Exploring The Existence of Stable Correlation Conditions. Technical Report 2001-04, The Intertek Group.

[16] Bao, D. (2007). A Generalized Model For Financial Time Series Representation and Prediction, Appl. Intell. 29 (1), 1–11.

[17] Fu, T.C., Chung, F.L., Luk, R., Ng, C.M. (2010). Financial Time Series Indexing Based on Low Resolution Clustering, in: Proceedings of The 4th IEEE International Conference on Data Mining (ICDM-2004), Pp. 5–14.

[18] Ghassempour, S., Girosi, F., Maeder, A. (2014). Clustering Multivariate Time Series Using Hidden Markov Models, Int. J. Environ. Res. Public Health 11 (3), 2741–2763.

[19] Spiegel S., Gaebler J., Lommatzsch A., De Luca E., Albayrak S., (2011). Pattern Recognition And Classification For Multivariate Time Series, Sensorkdd '11.

[20] An S., Gao X., Jiang M., Sun X., (2018). Multivariate Financial Time Series in The Light Of Complex Network Analysis, Physica A 503, 1241–1255, 2018.

[21] Agrawal, R., Faloutsos, C., Swami, A. (1993). Efficient Similarity Search in Sequence Databases, Found. Data Organ. Algorithms 46, 69–84.

[22] Keogh, E. (2002). Exact İndexing Of Dynamic Time Warping. In 28th International Conference on Very Large Databases, Hong Kong. Pp. 406-417.

[23] Berndt, D. & Clifford, J. (1994). Using Dynamic Time Warping To Find Patterns in Time Series. AAAI-94 Workshop On Knowledge Discovery İn Databases.

[24] Eiter, T., & Mannila, H. (1994). Computing Discrete Fréchet Distance. Tech. Report CD-TR 94/64, Information Systems Department, Technical University Of Vienna.

[25] Drimel A., Krivošija A., Sohler C., (2015). Clustering Time Series Under The Frèchet Distance, Arxiv.Org.

[26] Casado De Lucas D (2010). Classification Techniques For Time Series and Functional Data. Ph.D. Thesis, Universidad Carlos III De Madrid.

[27] Keogh, E. And Kasetty, S., (2002). On The Need For Time Series Data Mining Benchmarks: A Survey And Empirical Demonstration. In Proceedings Of The 8th ACM SIGKDD Int'l Conference On Knowledge Discovery And Data Mining. Edmonton, Alberta, Canada, July 23-26. Pp. 102-111.

RFM Model for Segmentation in Retail Analytics: A Case Study

İnanç KABASAKAL

Faculty of Economics and Administrative Sciences, Ege University, TURKEY

Abstract: Over the last decades, marketing scholars have often drawn attention to the value of customers for businesses that aim to endure in a harsh competitive environment. Customer Relationship Management (CRM) has been a prominent approach in marketing management that aims to improve relationships with customers. A practical implication of the CRM approach is the analysis of customer data to extract value for businesses, as well as customers. Segmentation has been a useful task that helps to group customers with similar attributes and designate better-tailored marketing strategies for customer groups. Among a variety of approaches for customer segmentation, Recency Frequency Monetary (RFM) Model stands out as an easy-to-adopt and effective technique. In this study, segmentation with RFM approach will be conducted over the purchase records obtained from an e-retailer. The segments and relevant marketing strategies will be presented in the findings. Moreover, a software implementation for the RFM model will be introduced along with a case study.

Keywords: Business Analytics, Analytical CRM, Customer Segmentation, RFM Model

REFERENCES

[1] E. Chablo, "The İmportance Of Marketing Data İntelligence in Delivering Successful CRM", in Customer Relationship Management, Vieweg+ Teubner Verlag, Wiesbaden, 2000, Pp. 57–70.

[2] S. Gupta, D. R. Lehmann. "Customers As Assets". Journal Of Interactive Marketing, 2003, 17(1), 9–24.

[3] P. E. Pfeifer. "The Optimal Ratio Of Acquisition And Retention Costs", Journal of Targeting, Measurement And Analysis For Marketing, 2005, 13(2), 179–188.

[4] Han, J. and Kamber, M. (2006) Data Mining Concepts and Techniques. 2nd Edition, Morgan Kaufmann Publishers, San Francisco.

[5] C. H. Cheng, And Y.S. Chen. "Classifying The Segmentation Of Customer Value Via Rfm Model And Rs Theory", Expert Systems With Applications, 2009, 36(3), Pp. 4176–4184.

[6] K. Elliott, R. Scionti, And M. Page. "The Confluence Of Data Mining And Market Research For Smarter Crm", Spss White Paper, Chicago, 2003, Pp. 1–11.

[7] M. Xu, And J. Walton. "Gaining Customer Knowledge Through Analytical CRM", Industrial Management & Data Systems, 2005, 105(7), Pp. 955–971.

[8] S.A.C. Madeira. "Comparison Of Target Selection Methods in Direct Marketing", 2002, Unpublished Msc. Thesis.

[9] J. T. Wei, S. Y. Lin, And H. H. Wu. "A Review of The Application Of RFM Model", African Journal of Business Management, 2010, 4(19), 4199–4206.

[10] V. Aggelis, And D. Christodoulakis. "Customer Clustering Using RFM Analysis", in Proceedings of The 9th World Scientific and Engineering Academy And Society International Conference on Computers, July 2005, Pp. 2–7.

[11] Kumar, V. "Clv: The Databased Approach", Journal of Relationship Marketing, 2006, 5(2-3), 7–35.

[12] M. Khajvand, K. Zolfaghar, S. Ashoori, And S. Alizadeh. "Estimating Customer Lifetime Value Based on Rfm Analysis of Customer Purchase Behavior: Case Study", Procedia Computer Science, 2011, 3, 57–63.

[13] B. Sohrabi, A. Khanlari. "Customer Lifetime Value (Clv) Measurement Based on Rfm Model", Iranian Accounting & Auditing Review, 2007, 14(47), Pp. 7–20.

[14] Sqlite Official Website. "About Sqlite", Url: Https://Www.Sqlite.Org/Index.Html, Date Accessed: 25.04.2019.

[15] Sqlite Official Website. "Sqlite: Most Widely Deployed and Used Database Engine", Url: Https://Sqlite.Org/Mostdeployed.Html, Date Accessed: 25.04.2019.

[16] M. Fuchs, W. Höpken, And M. Lexhagen. "Big Data Analytics For Knowledge Generation in Tourism Destinations – A Case From Sweden", Journal Of Destination Marketing & Management, 2014, 3(4), Pp.198–209.

[17] S. Sayad. "An Introduction to Data Science: Data Preparation", Https://Www.Saedsayad.Com/Data_Preparation.Htm, Date Accessed: 20.04.2019.

[18] K. K. Tsiptsis, And A. Chorianopoulos. "Data Mining Techniques in Crm: Inside Customer Segmentation", John Wiley & Sons, 2009.

[19] Putler Blog. "Rfm Analysis For Successful Customer Segmentation", Url: Https://Www.Putler.Com/Rfm-Analysis, Access Date: 26.04.2019.

Siber Tehdit İstihbaratı Alanında Makine Öğrenmesi Algoritmalarının Kullanılması

Cemile SARICAOĞLU, Mehmet DEMİRCİ

Department of Computer Engineering, Gazi University, TURKEY

Abstract: Nowadays, with the developing technology, the amount of data that is owned and processed is increasing day by day. It is very important to ensure the security of data, which is one of the biggest assets for institutions and organizations. With traditional security methods, attacks can be detected and prevented, but cybercriminals spend a lot of time and resources on advanced and targeted attacks that can bypass these methods. The present methods are reactive because they are generally updated with the information obtained from the analyzes performed after a successful attack. More proactive approaches are needed to improve safety. Cyber threat intelligence represents such a proactive approach and involves collecting and analyzing information for potential threats from a wide variety of data sources. The purpose of cyber-threat intelligence is to proactively adapt security controls to understand the methodology used by different attackers and to detect and prevent such activities. In the world of technology, the defense against attacks is one of the most important issues. Today, different approaches and effective methods have been used to obtain intelligence. These include vital information about security threats, which are used by hacker forums and other platforms as a means of communication between hackers. The amount of data on such platforms is very large. The manual analysis of these data is time-consuming, ineffective and requires a considerable amount of resources. In this sense, machine learning has become one of the popular approaches used in the field of cyber-threat intelligence in terms of its suitability to the subject, producing beneficial and effective results. In this study, information is given about cyber threat intelligence and in the world of hackers, how to obtain intelligence by using machine learning techniques is examined and evaluated in detail by supporting the studies conducted in the literature.

Keywords: Cyber Security, Threat Intelligence, Machine Learning, Hacker Forums

REFERENCES

[1] Webroot Threat Intelligence: What is it, How Can it Protect You From Today's Advanced Cyber-Attacks?URL:Https://Www.Gartner.Com/İmagesrv/Mediaproducts/Pdf/Webroot/İssue1_ Webroot. Pdf

[2] Shackleford, D. The SANS State Of Cyber Threat Intelligence Survey: CTI Important And Maturing. Technical Report, SANS Institute, 2016. URL: Https://Www.Sans.Org/Reading-Room/Whitepapers/Bestprac/Statecyber-Threat-İntelligence-Survey-Cti-İmportant-Maturing-37177.

[3] Türkmenoğlu, C. (2016). Türkçe Metinlerde Duygu Analizi (Doctoral Dissertation, Fen Bilimleri Enstitüsü).

[4] Alpaydin, E. (2004). Introduction to Machine Learning (Adaptive Computation And Machine Learning Series). The MIT Press Cambridge.

[5] Öztemel E., "Yapay Sinir Ağları", Papatya Yayıncılık, İstanbul (2003).

[6] Koyuncugil, A., & Özgülbaş, N. (2009). Veri Madenciliği: Tıp ve Sağlık Hizmetlerinde Kullanımı ve Uygulamaları. Bilişim Teknolojileri Dergisi, 2(2).

[7] Çoban, O. (2016). Metin Sınıflandırma Teknikleri ile Türkçe Twitter Duygu Analizi. Yüksek Lisans Tezi, Atatürk Üniversitesi, Fen Bilimleri Enstitüs.

[8] Ardıl, E. (2009). Esnek Hesaplama Yaklaşımı ile Yazılım Hata Kestrimi (Master's Thesis).

[9] KARTAL, E., Programı, E., & Balaban, M. E. Sınıflandırmaya Dayalı Makine Öğrenmesi Teknikleri ve Kardiyolojik Risk Değerlendirmesine İlişkin Bir Uygulama.

[10] Shearer, C. (2000). The CRISP-DM Model: The New Blueprint For Data Mining. Journal of Data Warehousing, 5(4), 13-22.

[11] Deliu, I. (2017). Extracting Cyber Threat İntelligence From Hacker Forums (Master's Thesis, NTNU).

[12] Nunes, E., Diab, A., Gunn, A., Marin, E., Mishra, V., Paliath, V& Shakarian, P. (2016, September). Darknet and Deepnet Mining For Proactive Cybersecurity Threat İntelligence. In 2016 IEEE Conference on Intelligence And Security Informatics (ISI) (Pp. 7-12).

[13] Marin, E., Diab, A., & Shakarian, P. (2016, September). Product Offerings in Malicious Hacker Markets. In 2016 IEEE Conference on İntelligence And Security İnformatics (ISI) (Pp. 187-189). IEEE.

[14] Samtani, S., Chinn, R., & Chen, H. (2015, May). Exploring Hacker Assets in Underground Forums. In 2015 IEEE International Conference on Intelligence and Security Informatics (ISI) (Pp. 31-36). IEEE.

[15] Samtani, S., Chinn, K., Larson, C., & Chen, H. (2016, September). Azsecure Hacker Assets Portal: Cyber Threat İntelligence And Malware Analysis. In 2016 IEEE Conference on Intelligence And Security Informatics (ISI) (Pp. 19-24). Ieee.

[16] Fang, Z., Zhao, X., Wei, Q., Chen, G., Zhang, Y., Xing, C.& Chen, H. (2016, September). Exploring Key Hackers And Cybersecurity Threats İn Chinese Hacker Communities. In 2016 IEEE Conference On Intelligence And Security Informatics (ISI)(Pp. 13-18). IEEE.

[17] Benjamin, V., Li, W., Holt, T., & Chen, H. (2015, May). Exploring Threats And Vulnerabilities in Hacker Web: Forums, IRC And Carding Shops. In 2015 IEEE International Conference On Intelligence And Security Informatics (ISI) (Pp. 85-90). IEEE.

[18] Macdonald, M., Frank, R., Mei, J., & Monk, B. (2015, August). Identifying Digital Threats in A Hacker Web Forum. In 2015 IEEE/ACM International Conference On Advances İn Social Networks Analysis And Mining (ASONAM) (Pp. 926-933). IEEE.

Analysis of Non-Risked Provinces; Unemployment and Traffic Accidents

Taner ERSÖZ, Betül CEBESOY

Department of Actuary and Risk Management, Department of Industrial Engineering
Karabük University, TURKEY

Abstract: Risk is any possibility that affects the achievement of the intended objectives. In this study, it was investigated which provinces did not contain any risks. In the calculation of risk values of provinces; unemployment and traffic accidents were taken into consideration. Turkey Statistical Institute (TURKSTAT), the General Directorate of Security, the Social Security Institution (SSI) and Turkey Business Association (TBA) risk values obtained from benefiting from the data of the years 2013 to 2017 were calculated. Risk values are evaluated between 1 and 5. The green-colored provinces are risk-free, while the red-colored region is cons idered to be of high risk. Fine-Kinney method was used for risk analysis.

Keywords: Free Risk Zone, Risk Analysis, Fine-Kinney Method, Risk Map

Use of Grid Search in Hyper-Parameter Selection for Time Series Analysis: A Case Study with Ad Mediation Software

Görkem GİRAY[1], Murat Osman ÜNALIR[2], Şeyma TAHMAZ[1]

[1]Kokteyl A.Ş.

[2]Computer Engineering Department, Ege University, TURKEY

Abstract: The success of an Ad Mediation Software's decision on from which ad network to request ads and in which order depends on the ability to estimate eCPM (effective Cost Per Mille) value used to measure ad revenue. This value varies for different applications depending on different external factors. It is not possible for domain experts to make successful predictions by analyzing different sets of external factors for a large number of applications and to keep these estimates constantly up to date. Therefore, eCPM values were automatically estimated separately for each application based on different advertising spaces and different countries using time series analysis. The ARIMA model was used to estimate and the hyper-parameters of the model were optimized by using grid search method. For most of the values obtained, it was found that the values obtained with the intuition of domain experts were closer to the actual values.

Keywords: Time Series Analysis, ARIMA, Hyper-parameter Optimization, Grid Search Method, Ad Mediation Software

Giriş

Mobil cihazların ve bununla paralel olarak mobil uygulamaların kullanımı son yıllarda oldukça yaygınlaşmıştır. Bu uygulamaların gelirlerinin önemli bir bölümü reklam gösteriminden gelmektedir. Bunun sonucu olarak mobil uygulamalara reklam sağlayan çok sayıda reklam ağı ortaya çıkmıştır. Reklam ağları, reklam verenlerden reklam alarak bu reklamları uygun yayıncılara (bu bildiri kapsamında mobil uygulamalara) sağlamaktadır. Çok sayıda reklam ağının ortaya çıkmasıyla birlikte bir yayıncının kendisine en çok kazanç sağlayacak reklamı yayımlayabilmesi için çok sayıda reklam ağı ile çalışması gerekliliği ortaya çıkmıştır. Bu karmaşıklığı yayıncıların

yerine yönetmek için de reklam ağları ile reklam yayıncıları arasında reklam aracıları konumlanmıştır.

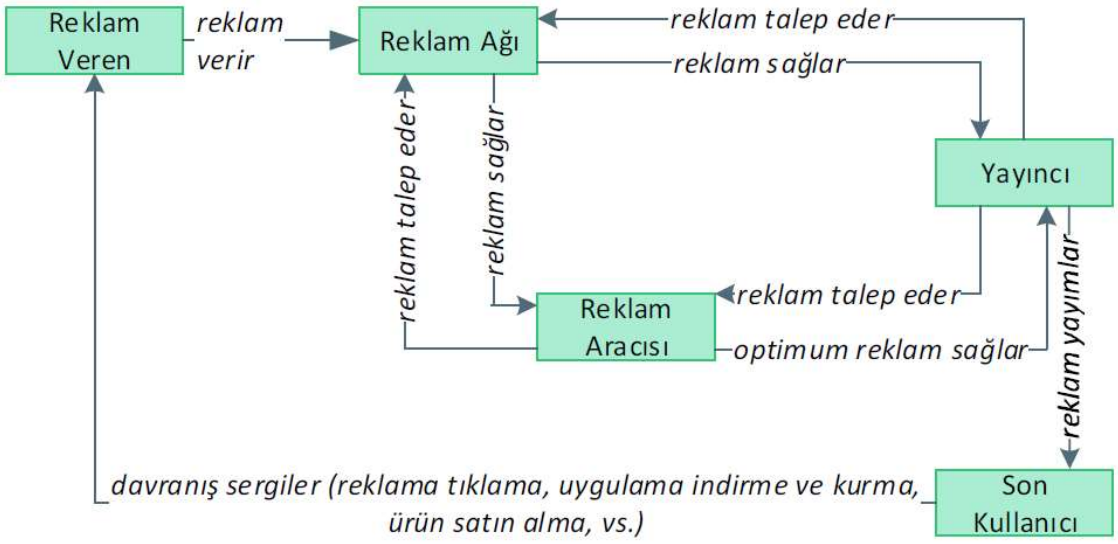
Bir reklam aracısı, bir uygulamaya, yani yayıncıya, reklam sağlarken, hangi reklam ağlarından hangi sırada reklam talep edeceğine dair bir karar verir. Bu kararı verirken uygulamanın reklam gelirini (aynı zamanda da kendi reklam gelirini) azami seviyeye çıkarmayı hedefler. Bunun için de reklam gösteriminden en fazla gelir sağlanabilecek reklam ağından başlayarak reklam talebinde bulunur. Mobil reklam sektöründe reklam gösterimi başına geliri ifade etmek için eBGBM (Etkin Bin Gösterim Başına Maliyet; İngilizcesi: eCPM – Effective Cost per Mille) kavramı kullanılmaktadır. Mevcut durumda Kokteyl şirketinin Reklam Aracısı Yazılımında (RAY) gerçekleşecek eBGBM değerleri alan uzmanları tarafından sezgiye dayalı olarak tahmin edilmektedir. eBGBM değerini etkileyebilecek değişkenler tespit edilerek, bu değişkenlerle eBGBM arasındaki örüntülerin makine öğrenmesi algoritmalarıyla tespit edilmesi mümkün görünmektedir [1 - 4]. Bu çalışma kapsamında mevcut durumda sezgisel olarak yapılan tahminlerin başarısı istatistiksel yöntemler kullanılarak incelenmiştir. Elde edilen sonuçlara göre başka yöntemler kullanılarak daha iyi tahmin yapılabileceği görülmüştür.

Bildirinin geri kalan bölümleri şu şekilde düzenlenmiştir: İkinci bölüm mobil reklam sektörü ve RAY hakkında genel bir bilgi vermektedir. Üçüncü bölümde problem tanımı, beşinci bölümde ise bu problem için geliştirilen çözüm anlatılmaktadır. Son bölümde sonuçlar ve gelecek çalışmalar sunulmuştur.

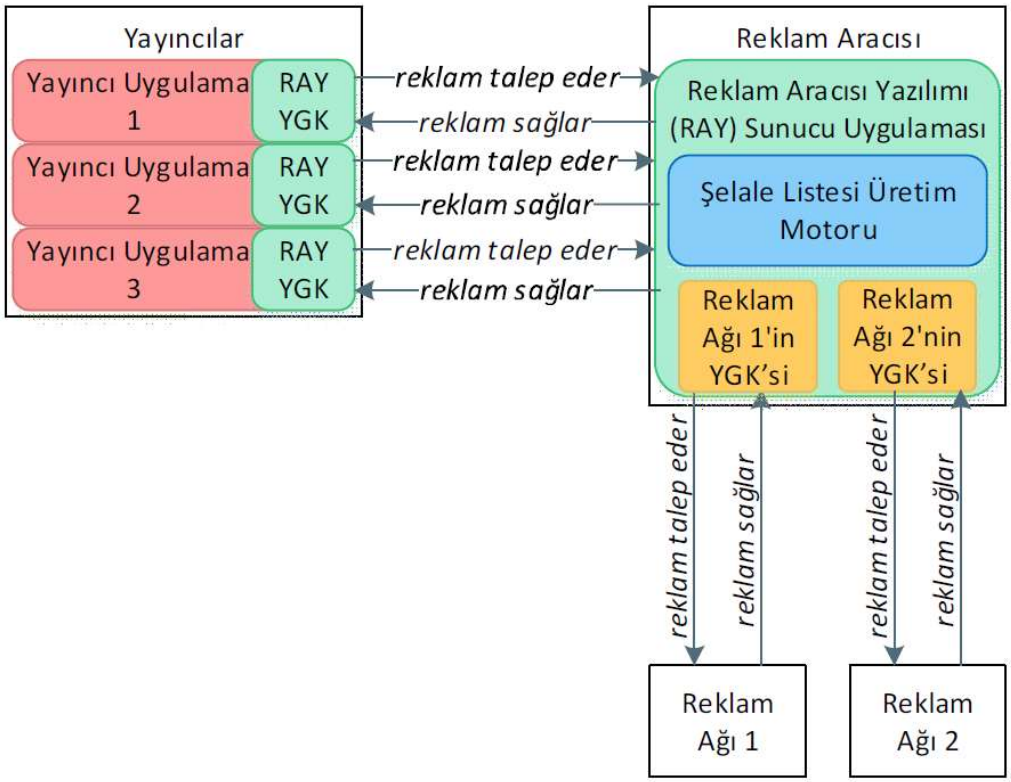
Arka Plan

Şekil 1 mobil reklam sektöründeki ana paydaşları göstermektedir. Reklam verenler, reklam ağlarına ürünleri ve/veya hizmetleri hakkında reklam vermektedir. Yayıncılar, reklam ağlarından reklam talep etmekte ve reklam ağları yayıncıya o andaki en uygun reklamı sağlamaktadır. Uygun reklamın seçimi konusundaki karar birçok değişkene (yayıncının hangi uygulama olduğu, son kullanıcı hakkındaki bilgiler, uygulamanın çalıştığı cihaz gibi) bağlı olarak değişmektedir. Yayıncıların reklam ağlarından doğrudan reklam alması durumunda bazı zorluklar ortaya çıkmaktadır. Uygulamanın geliştirilmesi açısından ortaya çıkan bir zorluk, yayıncı uygulamanın reklam talep edeceği tüm reklam ağlarıyla ayrı ayrı entegre olma gereksinimidir. İş amaçları açısından bakıldığında bir yayıncının birçok reklam ağından en uygun reklamı (en fazla geliri getirecek reklam) alamaması durumunda gelir kaybı yaşanacaktır. Dolayısıyla yayıncının en

uygun reklamı seçecek karar algoritmasını uygulamanın bir parçası olarak geliştirmesi gerekmektedir. Reklam aracıları birçok reklam ağı ile entegre olarak en uygun reklamı sağlama konusunda uzmanlaşmaktadır. Uygulama geliştirme açısından yayıncı uygulamanın sadece bir reklam aracısı ile entegre olarak birçok reklam ağından reklam alabilmektedir ve çok sayıda reklam ağı ile entegre olmanın getirdiği karmaşıklık reklam aracısı tarafından yönetilmektedir. Bunun yanında en uygun reklamı seçme konusunda uzmanlaşan reklam aracıları bu işi de yayıncı uygulamaların yerine yapmaktadır. Böylece yayıncıların reklam gelirlerini arttırmasına yardımcı olmaktadır. Son kullanıcılar ise yayıncı uygulamaları kullanarak reklamları görüntülemekte ve bir davranış (reklama tıklama, uygulama indirme ve kurma, ürün satın alma gibi) sergileyebilmektedir. Bu davranışlar yayıncılara gelir sağlamaktadır.



Şekil 1. Mobil Reklam Sektöründeki Ana Paydaşlar.



Şekil 2. Reklam Aracısı Yazılımının Yayıncılar Ve Reklam Ağları İle İlişkisi.

Şekil 2'de RAY'nin (bu yazılımın mimarisi [5] numaralı kaynakta anlatılmaktadır) yayıncılarla ve reklam ağlarıyla olan ilişkisi gösterilmektedir. RAY'yi oluşturan iki bileşen yeşil renk ile gösterilmiştir. RAY sunucu uygulaması, reklam ağlarının yazılım geliştirme kitlerini (YGK) içermektedir. Bu YGK'ler (Şekil 2'deki Reklam Ağı 1'in ve 2'nin YGK'si) aracılığıyla sunucu uygulaması reklam ağlarından reklam talep etmektedir ve reklam almaktadır. Yayıncı uygulamalar ise RAY'nin YGK'sini uygulamalarının içine yerleştirerek RAY'den reklam talep etmektedir ve almaktadır. Daha önce anlatıldığı gibi, teknik açıdan RAY birçok reklam ağı ile entegre olmanın getirdiği karmaşıklığı yayıncı uygulamalara saydam hale getirmektedir. İş amaçları açısından ise RAY birçok reklam ağından sağladığı reklamlar arasından en uygun reklamı yayıncı uygulamaya göndererek reklam gelirlerinin arttırılmasını sağlamaktadır.

Problem Tanımı

Şekil 2'de gösterilen RAY'nin içerdiği bir bileşen olan şelale listesi üretim motorunun ürettiği şelale listesi, yayıncının hangi reklam ağlarından hangi sırada reklam talebinde bulunması

gerektiğini belirleyen bir öncelik listesidir (Bkz. Şekil 3). Şelale listesi, her bir yayıncı için belirli aralıklarla (örneğin günlük) çeşitli değişkenler (reklam formatı, reklam alanı, geçmiş eBGBM değerleri gibi) göz önüne alınarak şelale listesi üretim motoru tarafından üretilir. Her yayıncı uygulamanın RAY YGK'si bu şelale listesini belirli aralıklarla sunucu uygulamasından indirerek reklam taleplerini bu güncellenmiş listeye göre yapar. Şelale listesindeki sıralama gelecekte gerçekleşecek eBGBM değerlerini ne kadar iyi yansıtırsa reklam gelirleri o kadar artacaktır.

```
1. sıradaki reklam ağından reklam talep et
Eğer reklam alındıysa
    Reklamı göster
    Çık
2. sıradaki reklam ağından reklam talep et
Eğer reklam alındıysa
    Reklamı göster
    Çık
3. sıradaki reklam ağından reklam talep et
Eğer reklam alındıysa
    Reklamı göster
    Çık
…
```

Şekil 3. Örnek Bir Şelale Listesi.

eBGBM reklam gösterimi başına geliri ifade etmektedir ve "(Gelir / Gösterim Sayısı) x 1000" formülü ile hesaplanmaktadır. Mevcut durumda şelale listesi oluşturmak için kullanılan girdiler (eBGBM tahminleri gibi) alan uzmanları tarafından yapılan gözlemler sonucunda sezgisel olarak belirlenmektedir ve güncellenmektedir.

Gerçekleşen eBGBM değerleri ve bağlantılı veriler RAY'de Tablo 1'de gösterilen alanlarda saklanmaktadır. Her yayıncı uygulamanın her bir reklam alanı için ülke bazında günlük olarak gerçekleşen ve daha önce tahmin edilen eBGBM değerleri RAY tarafından kaydedilmektedir. Her ülkede mobil uygulamaların kullanım dinamiklerinin farklı olması nedeniyle eBGBM değerleri ülke bazında ayrı ayrı tahmin edilmektedir. Bunun yanında reklam ağları da her yayıncı uygulama için ülke bazında gerçekleşen eBGBM değerlerini reklam aracılarıyla paylaşmaktadır.

Örnek olarak bir yayıncı uygulama için (uygulamanın adı gizlilik nedeniyle verilmemiştir) üç günde, dört ülkede gerçekleşen ve daha önce tahmin edilen eBGBM değerleri Tablo 2'de gösterilmektedir.

Tablo 1. eBGBM Değerlerinin ve Bağlantılı Verilerin Saklandığı Alanlar Ve Açıklamaları

Alan adı (Türkçe)	**Alan adı (İngilizce)**	**Açıklama**
Tarih	date	tarih
Reklam alanı no	placement_id	yayıncı uygulamanın reklam gösterilebilecek bir konumunun RAY'deki kimlik numarası
Yayıncı uygulamanın adı	app_name	yayıncı uygulamanın adı
Ülke	country	yayıncı uygulamanın kullanıldığı ülke
Gerçekleşen eBGBM	real_ecpm	gerçekleşen eBGBM değeri
Tahmin edilen eBGBM	weighted_ecpm	alan uzmanları tarafından tahmin edilen eBGBM değeri

Tablo 2. Örnek eBGBM Değerleri ve Bağlantılı Veriler

Tarih	**Reklam alanı no**	**Uygulamanın adı**	**Gerçekleşen eBGBM**	**Tahmin edilen eBGBM**	**Ülke**
08-09-18	00a78a52-7572-4ab6-863c-df5eb77ebe99	yayıncı uygulama	1.280701754	0.012145187	BR
09-09-18	00a78a52-7572-4ab6-863c-df5eb77ebe99	yayıncı uygulama	0.482967144	0.240236608	BR
10-09-18	00a78a52-7572-4ab6-863c-df5eb77ebe99	yayıncı uygulama	0.769817073	0.323807174	BR
08-09-18	00a78a52-7572-4ab6-863c-df5eb77ebe99	yayıncı uygulama	5.929345644	0.0260193	DE
09-09-18	00a78a52-7572-4ab6-863c-df5eb77ebe99	yayıncı uygulama	1.026119403	0.429765486	DE
10-09-18	00a78a52-7572-4ab6-863c-df5eb77ebe99	yayıncı uygulama	1.94153337	0.460431981	DE
08-09-18	00a78a52-7572-4ab6-863c-df5eb77ebe99	yayıncı uygulama	6.718092567	0.021978022	FR
09-09-18	00a78a52-7572-4ab6-863c-df5eb77ebe99	yayıncı uygulama	0.885636856	0.549004147	FR
10-09-18	00a78a52-7572-4ab6-863c-df5eb77ebe99	yayıncı uygulama	2.264134582	0.626127083	FR
08-09-18	00a78a52-7572-4ab6-863c-df5eb77ebe99	yayıncı uygulama	1.640059765	0.005023753	TR
09-09-18	00a78a52-7572-4ab6-863c-df5eb77ebe99	yayıncı uygulama	0.515668347	0.561521377	TR
10-09-18	00a78a52-7572-4ab6-863c-df5eb77ebe99	yayıncı uygulama	0.684581689	0.535929221	TR

Çözüm Önerisi

Alan uzmanlarının sezgileriyle yaptıkları tahminlerden daha iyi tahmin yapabilen modellerin oluşturulması için öncelikle eldeki veri ve çözülmeye çalışılan problem (eBGBM değerinin tahmin edilmesi) analiz edilmiştir. Öncelikle veri kümesinden örnek alt kümeler seçilmiştir. Üç reklam

alanı ve beş ülke için seçilen veri kümelerindeki gözlem sayısı (gerçekleşen eBGBM değerleri) Tablo 3'te gösterilmektedir.

Tablo 3. Analiz için Kullanılan Veri Kümelerindeki Gözlem (eBGBM değeri) Sayıları

	Ülke				
	BR	**DE**	**FR**	**TR**	**US**
Reklam alanı 1	144	144	144	144	139
Reklam alanı 2	250	261	261	259	262
Reklam alanı 3	275	280	280	283	280

Tablo 3'te gösterilen her bir veri kümesinin durağan olup olmadığını anlamak için Augmented Dickey-Fuller testi yapılmıştır. Durağan bir zaman serisinde gözlemler (bu çalışma kapsamında gerçekleşen eBGBM değerleri) zamana bağlı değildir. Bir zaman serisi trend ya da mevsimsel etki içermiyor ise durağan olarak nitelendirilmektedir. Durağan olmayan zaman serilerinde ise mevsimsel etkiler, trendler ve zamana bağlı başka örüntüler gözlemlenmektedir. Tablo 4'te gösterildiği gibi az sayıda durağan veri kümesi (Tabloda "D" ile ifade edilmiştir) bulunurken veri kümelerinin çoğu durağan değildir (Tabloda "DD" ile ifade edilmiştir).

Tablo 4. Veri Kümelerinin Durağanlık Durumu (D: Durağan; Dd: Durağan Değil)

	Ülke				
	BR	**DE**	**FR**	**TR**	**US**
Reklam alanı 1	D	DD	DD	D	D
Reklam alanı 2	DD	DD	DD	DD	DD
Reklam alanı 3	DD	DD	DD	DD	DD

Bu analiz ve veri üzerinde alan uzmanları ile yapılan çalışmalar sonucunda ARİMA modeli kullanılarak zaman serisi analizi yapılmasına karar verilmiştir. ARİMA modeli, zaman serisi verilerini analiz etmek ve tahmin etmek için kullanılmaktadır. ARİMA, "AutoRegressive Integrated Moving Average" kavramını temsil etmektedir. ARİMA modelinde üç hiper-parametre bulunmaktadır [6]:

- p: Model içerisinde yer alan gecikmeli gözlemlerin sayısını belirtir; gecikme sırası olarak adlandırılır.
- d: Ham gözlemlerin sayısını belirtir; ayırt edilme derecesi olarak adlandırılır.
- q: Hareketli ortalamanın penceresinin boyutudur; hareketli ortalama sıralaması olarak adlandırılır.

Bu üç hiper-parametrenin hangi kombinasyonunun en iyi tahmin değerlerini verdiğini belirlemek için grid arama yöntemi kullanılmıştır. En iyi kombinasyonu bulmak için her parametre

için aşağıdaki değerler kullanılmıştır. Tablo 4'te gösterildiği gibi hem durağan hem de durağan olmayan veri kümeleri olduğu için d parametresi hem 0 hem de 1 değeri alacak şekilde düzenlenmiştir.

- p = 0, 1, 2, 3, 4, 5, 6, 7
- d = 0, 1
- q = 0,1, 2

En iyi kombinasyon, bir yayıncı uygulamanın üç reklam alanının beş farkı ülkedeki eBGBM değerleri kullanılarak yapılmıştır. Her bir veri kümesinin üçte ikisi öğrenme, geri kalan üçte biri ise test için kullanılmıştır. Elde edilen hiper-parametre (p, d, q) değerleri Tablo 5'te gösterilmektedir. Bu hiper-parametre kombinasyonları kök ortalama kare hatası (KOKH) (root mean square error) en az olacak şekilde oluşturulmuştur. Hesaplama için kullanılan formül aşağıda gösterilmektedir.

$$\text{Kök ortalama kare hatası} = \sqrt{\sum_{j=1}^{n} (\text{Gerçekleşen eBGBM}_j - \text{Tahmin edilen eBGBM}_j)^2}$$

Tablo 5. Deney Sonucunda Elde Edilen Hiper-Parametre Değerleri

	Ülke	p	d	q
Reklam alanı 1	BR	0	1	1
	DE	1	0	0
	FR	0	0	1
	TR	1	0	0
	US	1	0	1
Reklam alanı 2	BR	1	1	0
	DE	1	0	1
	FR	1	0	1
	TR	1	0	1
	US	1	1	0
Reklam alanı 3	BR	0	1	1
	DE	4	0	2
	FR	1	1	0
	TR	1	0	1
	US	0	1	1

Elde edilen hiper-parametre kombinasyonları ve öğrenme için ayrılan veri kümesi kullanılarak geri kalan test veri kümesindeki gerçekleşen değerler için tahmin yapılmıştır. Tahmin yapılırken test kümesi için ayrılan gerçekleşen eBGBM değerleri kullanılmamıştır. Sonrasında tahmin edilen eBGBM değerleri ile gerçek eBGBM değerleri karşılaştırılarak kök ortalama kare

hataları hesaplanmıştır. Aynı şekilde gerçek eBGBM değerleri ile geçmişte alan uzmanları tarafından oluşturulan tahminler de karşılaştırılarak kök ortalama kare hataları hesaplanmıştır. Sonuçlar Tablo 6'da gösterilmektedir.

Tablo 6. Gerçekleşen Ebgbm Değerleri ile ARİMA Modeli ve Alan Uzmanının Tahmin Değerleri Arasındaki Kök Ortalama Kare Hataları

	Ülke	Gerçek – ARİMA tahmini için KOKH	Gerçek – Alan uzmanı tahmini için KOKH
Reklam alanı 1	BR	***0,171***	0,162
	DE	***1,895***	2,300
	FR	0,831	***0,812***
	TR	***0,154***	0,183
	US	3,941	***3,648***
Reklam alanı 2	BR	4,601	***2,402***
	DE	6,628	6,738
	FR	11,315	***10,437***
	TR	2,639	***2,109***
	US	6,211	***5,224***
Reklam alanı 3	BR	1,940	***1,599***
	DE	4,533	***3,802***
	FR	4,914	***3,369***
	TR	1,615	***1,295***
	US	6,104	***5,088***

Tablo 6'da koyu ve italik olarak gösterilen sonuçlar, gerçekleşen eBGBM değerleriyle daha az bir farkı temsil ettiği için diğerine göre daha başarılıdır. Elde edilen sonuçlara göre alan uzmanlarının kurduğu modelin yaptığı tahminler ARİMA modeli ile elde edilen tahminlerin çoğuna göre daha başarılıdır.

Sonuçlar ve Gelecek Çalışmalar

Reklam aracısı yazılımları, gerçekleşecek eBGBM değerlerini tahmin ederek yayıncı uygulamaların reklam gelirlerini arttırmayı hedeflemektedir. Bu tahmin işlemi için kurulabilecek en basit modellerden birisi geçmiş belirli bir dönemin ortalamasını tahmin olarak kullanmaktır. Diğer taraftan zaman serisi analizi gibi yüksek hacimli veri kümelerindeki gizli örüntüleri bularak daha iyi tahminler yapabilme potansiyeline sahip modeller bulunmaktadır. ARIMA bu modellerden birisidir. Bu çalışma kapsamında ARIMA modeli kullanılarak eBGBM değerleri tahmin edilmiş ve bunlar alan uzmanlarının tasarladığı modelin yaptığı sonuçlarla karşılaştırılmıştır. Yapılan deney sonucunda alan uzmanlarının tahminleri daha başarılı çıkmıştır.

Gelecek çalışmalar kapsamında, elde edilen sonucun geçerliliği daha fazla veri ile test edilecektir. Bunun yanında ARİMAX modeli kullanılarak ve tahmin başarısını arttırabilecek değişkenler bu modele eklenerek deney yapılacaktır. Ayrıca hiper-parametre optimizasyonu sırasında kombinasyon sayısı ve parametre değerleri arttıkça performans problemleri yaşanmıştır. Bu problemi gidermek için optimizasyonun MapReduce dağıtık programlama modeli ile yapılması planlanmaktadır.

TEŞEKKÜR

Bu çalışma TÜBİTAK'ın desteğiyle 3171053 numaralı proje kapsamında yapılmıştır.

KAYNAKÇA

[1] E. Llach, "System For Automatically Selling and Purchasing Highly Targeted and Dynamic Advertising Impressions Using A Mixture Of Price Metrics," U.S. Patent Application No. 10/767,050, 2004.

[2] B. O'kelley, "Method And System For Pricing Electronic Advertisements," U.S. Patent Application No. 11/006,121, 2006.

[3] A. Paunikar And M. Hochberg, "Dynamic Pricing For Content Presentations," U.S. Patent No. 8,706,547. 22, 2014.

[4] O. Chapelle, E. Manavoglu, R. Rosales, "Simple And Scalable Response Prediction For Display Advertising," Acm Transactions On Intelligent Systems And Technology (Tıst), Vol. 5(4), Article 61, 2015.

[5] Ö. Kırkağaçlıoğlu, G. Giray, A. Ersin, C. Çatıkkaş, S. Koçer, T. Şeremet, M. O. Ünalır, "Bir Reklam Aracısı Yazılımının Mimari Evrimi," 12. Ulusal Yazılım Mühendisliği Sempozyumu, İstanbul, 2018.

[6] J. Brownlee, Introduction To Time Series Forecasting With Python, Edition V1.6, 2019.

Marketing and Data Analytics; Increasing Importance of Marketing

Muhammet GİRGİN

Social Sciences Vocational School, Karabük University, TURKEY

Abstract: —At the moment we are entering a new era of knowledge, we are experiencing the first stages of a digital transformation. Businesses should be restructured in a structure suitable for the digital age when they are moving towards Industry 4.0, which is rising on cyber physical systems. In this period when the digital enterprises are rising, each business unit of the enterprise should be in the effort of using these innovations in accordance with this structure and using them in the most effective way to reach their goals. With the understanding of marketing in the modern sense, one of the most basic functions of marketing is to provide the information about the customer demands and expectations in order to be used in the product design and planning stage. For this reason, marketing is not only about the sales and marketing of the products or services produced, but also for the decision of what to produce. In this period, where the effects of the digital age are becoming more evident, the weight of digital applications in marketing is increasing and the concept of Digital Marketing is becoming more prominent. Innovations on the transformation of Data (which is called New oil) into information, give enterprises more competitive advantages and help to make more accurate decisions. In parallel with the developments in technology, the change in the market and the business environment necessitates the marketing to be prepared for these changes and to transform itself with this change. In this study, the analytical concept is discussed with marketing. In addition, the relationship between the analytical concepts of marketing is tried to be explained and the opportunities awaiting marketing are put forward. At the end of the study, it has been concluded that this digital transformation and developments all over the world increase the importance of marketing.

Keywords: Analytics, Marketing Analytics, Marketing 4.0, Big Data Analytics, Digital Marketing

REFERENCES

[1] A. C. Ramos, S. C. Fermont: Insider SEO & PPC. Jain Publishing. 2006

[2] A. Gilmore, D. Gallagher, H. Scott, E-Marketing And Smes: Operational Lessons For The Future. European Business Review, 2007, 19(3), 234-247.

[4] B. Bulunmaz, Gelişen Teknolojiyle Birlikte Değişen Pazarlama Yöntemleri ve Dijital Pazarlama TRT Akademi Cilt 1 / Sayı 2 / Temmuz 2016 Dijital Medya Sayısı.

[5] B. Manel, S. Dupuy-Chessa, L. Gzara, N. Mandran, C. Jeannet, BI4BI: A Continuous Evaluation System For Business Intelligence Systems Expert Systems With Applications, Journal Homepage: www.Elsevier.Com/Locate/Eswa, Volume 76 Pages 1-166 (15 June 2017) S:97.

[6] C. Dave, M. Patron, From Web Analytics To Digital Marketing Optimization: Increasing The Commercial Value Of Digital Analytics, Journal Of Direct, Data And Digital Marketing Practice. July 2012, Volume 14, Issue 1.

[7] C. Gurâu, F. Duquesnois, The Website As An İntegrated Marketing Tool: An Exploratory Study Of French Wine Producers. Journal Of Small Business And Entrepreneurship, 2011 24(1), 17-28.

[8] C. L. Bovee, ve J. V. Thill., Marketing, McGraw Hill Inc., USA, 1992.

[9] C. J. Shultz II. There İs More To Marketing: An Encore To Layton's Ring Cycle And The Compelling Case For Marketing Systems Theory, Research And Management, Australasian Marketing Journal 24, 2016.

[10] D. Chaffey, P.R. Smith, E-Marketing Excellence. New York: Routledge. 2013.

[11] D. Chaffey M. Patron, From Web Analyti Cs To Digital Marketing Optimization: Increasing the commercial valueof digital analytics, Journal of Direct, Data and Digital Marketing Practice (2012) 14, 30– 45.

[12] D. Q. Chen, D. S. Preston & M. Swink, How The Use Of Big Data Analytics Affects Value Creation İn Supply Chain Management Journal Of Management Information Systems, Journal Of Management Information Systems / 2015, Vol. 32, No. 4, Pp. 4–39.

[13] E. Uzunoğlu, İnternet Reklamcılığında Yeni Bir Yaklaşım, Edt. Sinem Yelgen Çakır Teknolojinin Pazarlama İletişimine Etkileri, Nobel Yay. Dağ. 2010.

[14] E. Schmidt, J. Cohen, Yeni Dijital Çağ, Çev. Ümit Şensoy, Optimist Yayınevi, 2014.

[15] F. B. Peltekoğlu, Halkla İlişkiler Nedir? İstanbul: Beta Basım Yayım Dağıtım. 2016.

[16] G. Çözen, (2011). Dijital Pazarlama Nedir? Teknikleri Nelerdir? Http://Www.Dijitalmarketing. Net/2011/03/31/Dijital-PazarlamaNedir-Teknikleri-Nelerdir/, Erişim: 07 Mart 2018.

[17] G. J. Dehkordi, S. Rezvani, M.S. Rahman, F. Fouladivanda, N. Nahid, S.F. Jouya, A Conceptual Study On E-Marketing And İts Operation On Firm's Promotion And Understanding Customer's Response. International Journal Of Business And Management, 7(19), 114-124. 2012.

[18] G. Viglia, Princing, Online Marketing Behavior, And Analitics, Palgrave Macmillian Publihed, 2014.

[19] H. J. Watson, H. W. Barbara, The Current State Of Business Intelligence, 96-97 IT Systems Perspectıves IEEE Computer Society, 2007.

[20] İ. E. Emre, Ç. S. Erol, Veri Analizinde İstatistik mi Veri Madenciliği mi? Bilişim Teknolojileri Dergisi, Cilt: 10, Sayı: 2, Nisan 2017.

[21] İ. Ertuğrul, Deniz, Gözde. 4.0 Dünyası: Pazarlama 4.0 ve Endüstri 4.0. Bitlis Eren Üniversitesi Sosyal Bilimler Enstitüsü Dergisi, 2018.

[22] J. Järvinen, H. Karjaluoto, The Use Ofweb Analytics For Digital Marketing Performance Measurement, Industrial Marketing Management, 2015.

[23] K. Nakatani, T. Chuang, (2011) "A Web Analytics Tool Selection Method: An Analytical Hierarchy Process Approach", Internet Research, Vol. 21 Issue: 2, Pp.171- 186, Https://Doi.Org/10.1108/10662241111123757.

[24] K. Suryanarayanan, S. K. Mathew, Business Analytics and Business Value: A Comparative Case Study, Information & Management, Available Online 31 January 2018.

[25] L. Fink, N. Yogev, A. Even, Business Intelligence and Organizational Learning: An Empirical İnvestigation Of Value Creation Processes, Information & Management 54 (2017) 38–56

[26] M. C. Cant, N. V. Rooyen, The Use of Search Engines in The Marketing of Accommodation Establishments: Adopt or Reject The Use Of Google Adwords?, International Review Of Management And Marketing, 2017, 7(3)

[27] M. Kubinaa, M. Varmusa, Irena Kubinova, Use of Big Data For Competitive Advantage Of Company, 4th World Conference On Business, Economics And Management, WCBEM, Procedia Economics And Finance 26 (2015)

[28] M. Kırlı, H. Gümüş, Y. Altıntop, Stratejik Yönetim Muhasebesinin Tamamlayıcısı Olarak Stratejik Pazarlama Muhasebesi, 4. Uluslararası Balkanlarda Sosyal Bilimler Kongresi, 2012

[29] N. Köse, D. Yeğin, Dijital Pazarlamadan Fijital Pazarlamaya Geçişe Örnek Olarak Artırılmış Gerçeklik Ve Sanal Gerçeklik Uygulamalarının Pazarlama Üzerindeki Katkılarının İncelenmesi, İstanbul Aydın Üniversitesi Dergisi - Ocak 2018 Cilt 10 Sayı 1

[30] P. Kotler, Gary Armstrong. Principles of Marketing, 8th Ed., Prentice-Hall Inc. New Jersey-USA, 1999.

[31] P. Kotler, Marketing Management, New Jersey: Pearson Education International. 2003

[32] P. Kotler, Kartajaya, Hermawan. Setiawan, Iwan. Marketing 3.0 John Wiley & Sons Publishing, 2010. USA

[33] P. Kotler, H. Kartajaya, I. Setiawan, Pazarlama 4.0 Gelenekselden Dijitale Geçiş, Optimis Yayın, 2018, İstanbul

[34] Ö. Torlak, R. Altunışık, (2018). Pazarlama Stratejileri Yönetsel Bir Yaklaşım, Beta Yayınları

[35] R. Glass, S. Callahan, The Big Data Driven Business, John Wiley & Sons Publishing, 2015. USA

[36] S. Dann, S. Dann, E-Marketing Theory And Application. New York: Palgrave Macmillan. 2011

[37] S. Ganesh, "Data Mining: Should It Be Included in The Statistics Curriculum?", The 6th International Conference On Teaching Statistics (ICOTS 6), Cape Town, Güney Afrika, 2002.

[38] T. Anna, The Reevaluation Of Communication In Customer Aproach – Towards Marketing 4.0, Internatipnal Journal of Contemporary Management, 12(4), 124-134, 2013

[39] T. Davenport, Big Data @ Work, Türk Hava Yolları Yay. Çev; Müge Çavdar. İstanbul 2014

Data Analytics and Importance in Health Sector

Duygu KARABULUT[1], Taner ERSÖZ[2]

[1]Department of Industrial Engineering,
[2]Department of Actuary and Risk Management, Karabük University, TURKEY

Abstract: Nowadays, as a result of the fast advancement of technology, large and complex databases are created. Since it may be hard to process the created complex data, there is a need for software programs. The purpose of this study is to create a substructure for data mining to be used in the health sector and to give a quick and different perspective in deciding to reach the desired data. With this purpose, based on the data from the operating room of a private hospital, the data mining method was used in this study.

Keywords: Health Sector, Data Analytics, Predictive Analysis, Classification

Giriş

Sağlık sektörü, sağlığa etkileri olan her türlü ürünü arz - talep etmek, tüketmek üzere çok farklı üretim alanlarında kurulmuş bütünleşik sistemler ile bunların içerdiği kişi, kurum, kuruluş, ürün ve benzerleri tümünü belirtmek için kullanılan bir kavramdır.

Türkiye'de sağlık hizmetleri Cumhuriyet'in kuruluşu 1923 yılından itibaren 1982 yılına kadar devletin sunması esas alınan bir hizmet olmuştur. Türkiye'de kamu dışında hizmet sunan sağlık kuruluşları sayısında hızlı artış ve özel sağlık sigortası için sağlık alanının gelişmesinin yaşandığı yıllar 1900'lerdir. Türkiye'de sağlık sektörü, 1980'lerden sonra 20 yılda yaklaşık 3 kat büyümüş, bu büyümede kamu sektörünün payı giderek belirleyici olmuştur.

Dünya ekonomisine bakıldığında günümüzde sağlık sektörü ilk sıralarda yer almaktadır. Sağlık sektörü için 1990'lı yıllarda yapılan toplam harcama yaklaşık 2985 milyar dolardır. Bu harcama dünya brüt milli hasılasının yaklaşık %8'ne denk düşmektedir. Bu rakamsal boyut, binlerce yıl öncesinde, aile üyeleri, dini örgütler ya da bazen profesyonel bir şifa verici tarafından tedavi sunulan bir kişiyle, bir hastalık arasındaki basit ve özel bir ilişkinin, geçmiş iki yüzyıl içinde nasıl genişlediğini ve bir sağlık sistemi tarafından kapsanan kompleks bir ağa nasıl girdiğini

yansıtmaktadır (WHO 2000). Sağlık sistemi performansını ele almadan önce yapılması gereken, performans kriterlerinin geliştirilmesine temel teşkil edecek şekilde sağlık sisteminin tanımlanması ve sınırlarının çizilmesidir. Günümüzde bir sağlık sisteminin ne olduğunu, nelerden oluştuğunu ve nerede başlayıp nerede bittiğini tam olarak söylemek zordur. Dünya Sağlık Örgütü (DSÖ) 2000 Raporu sağlık sistemini, temel amacı sağlığı geliştirmek, yenilemek ve sürdürmek olan tüm aktiviteleri içerecek biçimde tanımlamıştır (WHO 2000; Murray, Frenk 2000; WHO 2001; Murray, Frenk 2001; IHSD 2000a; WHO 2000a).

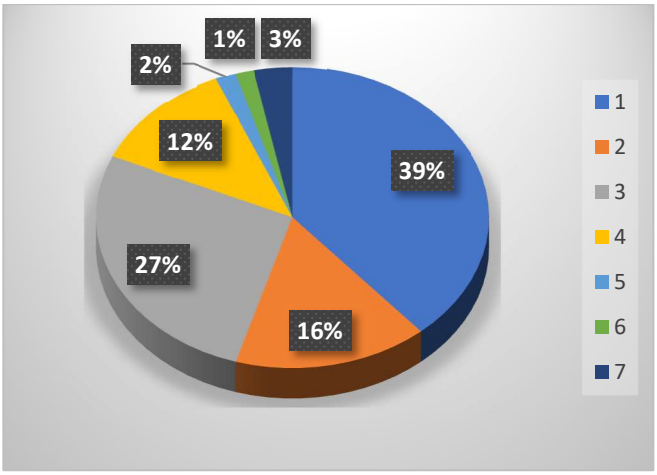
Günümüzde sağlık sistemlerinde birçok girdi, süreç ve çıktılar söz konusudur. Binlerce hastalık, teşhis ve tedavi yöntemi, ilaç, tıbbi malzeme ve işlemler mevcut olmakla birlikte, sağlık sistemi birçok kurum, kuruluş, kişi, kaynak ve finansman modelini içerdiğinden dolayı bunlar arasındaki planlama, koordinasyon ve yönetim faaliyetleri profesyonel yönetim anlayışı gerektirmektedir.

Doğru karar vermede, verilerin toplanması, hazırlanması, analiz edilmesi ve doğru yorumlanması çok önemlidir. Verilen kararın doğruluğu bilgi birikimine bağlı olduğu kadar verinin kaynağının yeterliliği ile de ilişkilidir.

Tıp alanındaki dijital verilerin artışı birçok sorunları beraberinde getirmiştir. Veri madenciliği metadolojisi günümüzde birçok farklı sektörle birlikte tıp alanında da oluşan birtakım açığa çıkan problemlerden dolayı giderek yaygınlaştığı görülmektedir.

Cerrahi müdahaleler büyük miktarda risk oluşturabileceği ihtimaline karşılık dikkatle planlanmalıdır. Cerrahi girişimin uygun seçilmesi için hasta değerlendirmesi temel alınmalıdır. Müdahale seçiminde hastanın öyküsü, fiziksel durumu ve tanısal verilerin önemi kadar müdahalenin hasta için riskleri ve yararları da son derece önemlidir. Müdahale seçiminde, başvuru bulguları, tanısal test ve diğer kaynaklardan elde edilen bilgiler de göz önüne alınmalıdır. Günümüzde cerrahi müdahalelerin en önemlileri arasında yer alan Ortopedi ve Travmatoloji, insanın doğumundan ölümüne kadar olan kas ve iskelet sistemi hastalıklarının neredeyse tamamıyla dolaylı ya da direk olarak ilgilenen bilim dalıdır. Ortopedi, kelime anlamı düzgün çocuktur. Travmatoloji ise travma sonucu yaralanmalara bakan bölüm olarak adlandırılır.

Bu çalışmada özel bir hastanenin 2010- 2014 yılları arasında gerçekleştirilen ameliyatların dağılımı Şekil 1'de verilmiştir. Genel ortopedinin yanında artroskopik cerrahi, travma cerrahisi ve eklem protezleri ameliyatların da istatistiki dağılımı verilmektedir.



Şekil 1. 2010- 2014 Yılları Arası Ameliyatların Dağılımı

1.Genel ortopedi; 2.Diz ve kalça protezi; 3.Artroskopik cerrahi; 4.Travma cerrahisi; 5.Omurga cerrahisi; 6.Tümör cerrahisi;7.Tümör cerrahisi

Ortopedi ve Travmatoloji, teknoloji ve bilimsel ilerlemelerin ışığında en hızlı gelişen, değişen ve hatta kapsamı artan dalların başında gelmektedir.

Alt Dalları:

• Artroplasti cerrahisi (eklem protezleri)

• Spor travmatolojisi

• Boy uzatma ve bacak eşitsizlikleri

• Çocuk ortopedi ve travmatolojisi

• El cerrahisi ve mikrocerrahi

• Diz cerrahisi ve artroskopik cerrahi

• Omuz ve dirsek cerrahisi

• Ortopedik onkoloji

• Ayak ve ayak bileği cerrahisi

• Kemik İltihapları (Osteomyelit) tedavisi

Bu çalışmada özel bir hastane ameliyathane verileri kullanılarak veri madenciliği uygulamaları gerçekleştirilmiştir. Hastanenin gelecek aylardaki hasta yoğunluklarının, ameliyathane çeşitlerinin cinsiyet bazlı dağılımı esas alınarak tahmin edilmesinde IBM SPSS Modeler üzerinde çalışılmış ve böylece en kestirimci istatistiki verilerin elde edilmesi esas alınmıştır.

Literatür

Selma Altındiş ve İlknur Kıran Morkoç'un yapmış oldukları 2018 yılındaki çalışmada Sağlık Hizmetlerindeki Büyük Verilere değinilmektedir. Çalışmanın amacı, sağlık hizmetlerinde büyük veri ve kullanım alanları hakkında bilgi vermektir. Sonuç olarak Sağlıkla ilgili verilerin devasa miktarlara ulaşması, bu verilerin geleneksel veri işleme yöntemleri tarafından işlenmesini zorlaştırmış ve büyük veri kavramının sağlık hizmetlerine girmesine neden olduğuna varılmaktadır. Daha önce geleneksel veri işleme yöntemleri ile depolanamayan, yönetilemeyen ve analiz edilemeyen yüksek hacimli, hızlı ve çeşitli veri kümelerinin anlamlı ve değer yaratacak sonuçlara dönüşmesi Büyük Veri ile mümkün olabileceği görüşü savunulmaktadır. Ülkemizde de sağlık sisteminin performansını artırmak amacıyla büyük hacimlerdeki sağlık veri setlerini toplamak ve analiz etmek üzere Büyük Veri Araştırma Enstitülerinin kurulması önerilmektedir.

Sezgin Irmak, Can Deniz Köksal, Özcan Asilkan'ın yapmış oldukları 2012 yılındaki çalışmada hali hazırda işleyen bir hastane veritabanında bazı önemli veri madenciliği teknikleri ile hasta yoğunluklarının tahmin edilmesi uygulamaları yapılmış ve sonuçları karşılaştırmalı olarak aktarılmıştır. Bu çalışmada bir hastane veritabanı kullanılarak veri madenciliği uygulamaları gerçekleştirilmektedir. Hastanenin gelecek aylardaki hasta yoğunluklarının tahmin edilmesinde üstel düzgünleştirme, ARIMA ve yapay sinir ağları teknikleri önce kendi içinde farklı modellerle karşılaştırılmış sonrasında da her tekniğe ait en iyi modeller kendi aralarında kıyaslanmıştır. Böylece en kestirimci model belirlenmeye çalışılmıştır. Çalışmada kestirimci analiz olarak gelecekteki hasta yoğunluklarının tahmin edilmesi amaçlanmış ve üç farklı veri madenciliği tekniği ve bunların da kendi içinde farklı modelleri üretilerek gelecekteki hasta yoğunluklarının tahmin edilmesi ve bu konuda en iyi modellerin belirlenmesi amaçlanmıştır. Üstel düzgünleştirme yöntemleri arasında en kestirimci model Winters Additive modeli olmuştur. ARIMA süreçleri içinde en kestirimci model ARIMA(3,1,0)(1,0,0)12 modeli olmuştur. Yapay sinir ağları yöntemleri arasında ise en kestirimci model Prune yöntemiyle elde edilen model olmuştur. Yapay sinir ağları modelleri arasında veriye daha fazla uyum gösteren modeller olmasına rağmen bunlarda

aşırıöğrenme probleminin gerçekleştiği görülmüştür. Her yöntemin en kestirimci modellerinin birbirleriyle kıyaslanması, uyum iyiliği kriterleri ve modellerin tahmin değerleri ile hastane veritabanından elde edilen gerçekleşen hasta sayısı değerlerinin karşılaştırılması yöntemleri kullanılarak gerçekleştirilmiştir. Her iki konuda da Winters Additive üstel düzgünleştirme modeli en kestirimci model olmuştur. Uyum iyiliği kriterleri bakımından ve ilk 7 aydaki tahminlerin gerçeğe yakınlığı bakımından ARIMA(3,1,0)(1,0,0)12 modeli ikinci en iyi olsa da 8. ve 9. aylardaki tahmin değerleri kötüye gitmiş ve yapay sinir ağları modeli bu aylarda Winters Additive üstel düzgünleştirme modelinden sonra ikinci en iyi tahminleri gerçekleştirmiştir. Gerçekleşen sayıların tahminlere oldukça yakın olması bu tekniklerin hastanenin yoğunluk tahminleri için kullanılabileceğini göstermektedir. Ayrıca veri madenciliği teknikleri kullanılarak sağlık sektörü veritabanları veya veri ambarlarından birçok amaç için faydalı bilgilerin elde edilmesi de mümkün gözükmektedir.

Ali Serhan Koyuncugil, Nermin Özgülbaş'ın 2009 yılında yapmış oldukları çalışmada sağlıkta Veri Madenciliğinin kullanımı konusunda bir altyapı oluşturmak ve sağlık profesyonellerine sağlık sektöründe Bu çalışmada sırasıyla Veri Tabanlarında Bilgi Keşfi, Veri Ambarı, Veri Madenciliği, İş Zekası ve Veri Madenciliği Yöntemleri konularında tanımlayıcı bilgilere yer verilmekte; ülkemizdeki sağlık sektöründe öncelikli konu ve sorun alanları dikkate alınarak Veri Madenciliği uygulamalarına örnekler verilmektedir. Geleceğin sayısal karar verme ve iş zekası yöntemi olan Veri Madenciliğinin konunun uzmanı kişiler tarafından sağlık sektöründe kullanımı, sağlık hizmetlerinin daha etkin sunumu, kaynakların daha verimli kullanımı ve bilimsel, karşılaştırılabilir, şeffaf bilgi erişimi açısından önerilmektedir.

Şebnem Aslan, Mete Sezgin, Selçuk Burak Haşıoğlu'nun 2008 yılında yapmış oldukları çalışmada özel sağlık kuruluşlarında müşteri memnuniyetine etki eden faktörlerin ve müşterilerin sağlık kuruluşu tercihlerindeki ölçütleri araştırılmaktadır. Araştırma, Konya ilinde yedi özel sağlık kuruluşundan yararlanan 200 katılımcıyla gerçekleştirilmiştir. Araştırmada Analitik Hiyerarşi Süreci (AHS) yönteminden yararlanılmıştır. Araştırmanın sonucunda, müşteri memnuniyetine etki eden faktörlerden en etkili ölçütün, algılanan kalite olduğu tespit edilmiştir. Diğer ölçütler sırasıyla; fiyat, kolaylık, referans ve yakınlık bulunmuştur. Ayrıca çalışmada, doktor hizmetinin, müşteri memnuniyetini en yüksek düzeyde etkileyen değişken olduğu tespit edilmiştir.

Mehtap Çakmak, M. Kemal Öktem, Uğur Ömürgönülşen'in yapmış oldukları 2008 yılındaki çalışmanın amacı, genelde Türk kamu hastanelerinin etkinlik sorununu irdelemek ve özelde ise

T.C. Sağlık Bakanlığı'na bağlı kadın doğum hastanelerinin teknik etkinliklerini ölçmektir. İkinci basamak sağlık hizmeti sunan ve birden çok girdisi ve çıktısı bulunan kadın doğum hastanelerinin teknik etkinliklerinin ölçümünde, girdi ve çıktı çeşitliliği sorunu karşısında birden fazla girdi ve çıktıyı aynı anda hesaba katarak ölçüm yapan Veri Zarflama Analizi (VZA) tekniği kullanılmıştır. Yapılan ölçüm sonucunda, araştırma kapsamındaki hastanelerin yaklaşık 1/3'nün etkin, 2/3'ünün ise etkinsiz faaliyet gösterdiği saptanmıştır.

Gizem Gülsevin, Ayça Hatice Türkan'ın 2012 yılında yapmış oldukları çalışma, Afyonkarahisar'daki Sağlık Bakanlığı'na bağlı hastanelerin etkinlik düzeylerinin veri zarflama analizi (VZA) yöntemi ile belirlenmesini amaçlamaktadır. VZA, parametrik olmayan bir etkinlik ölçüm yöntemi olup diğer etkinlik ölçüm yöntemlerine göre daha gerçekçi ve doğru sonuçlar ortaya çıkarır. Bu nedenle çalışmada analiz için VZA tekniği tercih edilmiştir. Çalışma, Afyonkarahisar İl Sağlık Müdürlüğünden sağlanan Afyonkarahisar'daki Sağlık Bakanlığına bağlı hastanelerin 2011 yılına ait verilerini içermektedir. Hastane yönetiminin girdiler üzerinde kontrol gücü vardır, ancak çıktılar üzerinde kontrol gücü oldukça zordur. Bu nedenle çalışmada girdileri minimize etmeyi amaçlayan, ölçeğe göre sabit getiri varsayımına dayanan girdi yönlü Charnes Cooper Rhodes (CCR) modeli kullanılmıştır. Kurulan model WinQSB ve EMS programlarından yararlanılarak çözülmüştür. Yapılan analiz sonucunda 8 hastane %100 etkinlik skoru ile etkin olarak bulunmuştur.

Şeyda Gür, Buse Uslu, Tamer Eren, Nesrin Akca, Ali Yılmaz, Seda Sönmez'in yapmış oldukları 2018 yılındaki çalışmada, ameliyathanelerin performanslarının artırılmasında etkili olan kriterlerin belirlenmesi için çok ölçütlü karar verme yöntemlerinden Analitik Ağ Süreci yöntemi kullanılmıştır. Kriterler arasındaki bağımlılıkları, etkileşimleri ve geri bildirimleri dikkate alan bu yöntem ile kriterlerin birbiri üzerindeki ilişkileri incelenmiştir. Çalışmada kullanılan Analitik Ağ Süreci yöntemi ile ameliyathane performanslarının artırılmasında etkili olan kriterlerin önem dereceleri hesaplanmıştır. Bu önem derecelerine göre hangi kriterin performans üzerinde ne derece etkisi olduğu gösterilmiştir. Elde edilen sonuçlara bakıldığında; iş gören kriteri, maliyet kriteri ve öğrenme ve büyüme kriterinin ön plana çıktığı görülmektedir.

Müberra Terzi'nin 2018 yılında yapmış olduğu çalışmada veri madenciliği hakkında bilgi edinip veri madenciliği metotlarının ülkemiz sağlık sektöründe nasıl kullanıldığına ve hangi alanlarda kullanılabileceğine değinilmiştir. Bu derlemede veri madenciliği hakkındaki bilgiler tek çatı altında toplanmaya çalışılmış, ülkemiz sağlık sektöründe veri madenciliği uygulamalarına ve

kullanım alanlarına göz atılmıştır. Derlemedeki örnek uygulamalar ülkemizde yapılan çalışmalardan seçilmiş olup hastalık riski üzerine yapılan çalışmaların daha fazla olduğu gözlemlenmiştir. Hastalık riski üzerine olan çalışmalardan sonra ilaç kullanımlarıyla ilgili olan çalışmalar dikkat çekmektedir. Veri madenciliğinin sağlık sektöründeki çalışmaları hastalık riski ve ilaç dozu ile sınırlı olmayıp sağlık sektöründe çalışanların beklentileri derleme içerisinde verilmiştir. Çalışmaların sonuçları yüksek başarı oranı içermekte olup veri madenciliğinin sağlık çalışanlarına verileri yorumlamalarında yardımcı bir araç olabileceği düşünülmektedir.

Selma Altındiş'in 2018 yılında yapmış olduğu çalışma ile hasta memnuniyeti, sürekli geliştirme, etkililik, verimlilik ve hasta güvenliği kavramları üzerinden günümüzde önemli bir konu haline gelen büyük veri ile sağlık hizmetleri kalitesi ilişkisinin değerlendirilmesi amaçlanmaktadır. Sonuç olarak, büyük verilerin kullanımında deneyimsizlik, analitik geliştirme maliyeti gibi birtakım zorlukların bulunmasına rağmen büyük veri teknolojilerinin benimsenmesi, uygulanması ve kullanılmasının, sağlık hizmetlerinde olumlu bir etkiye sahip olacağı düşünülmektedir. Bu anlamda sağlık hizmeti sunan kurumlara, büyük veri analizi için gerekli kaynaklara (belirli teknoloji, analitik yöntemler vs) ve veri bütünleştirmesine yönelik yatırımlar yapmaları önerilmektedir.

Hüdaverdi Bircan, Selim Çam'ın 2016 yılında yapmış oldukları çalışmada, Cumhuriyet Üniversitesi Hastanesi'ne 2011 yılında başvurmuş olan hastaların 2006-2011 arasındaki kayıtlar, hasta başvuru davranışlarının belirlenmesi amacıyla incelenmiştir. Oluşturulan veri seti yasalar tarafından yetişkin sayılan 18 yaş ile emeklilik sınırı olan 65 yaş arasında bulunan hastalara indirgenmiştir. Böylece veri seti 78.239 hastanın hastane veri tabanından alınan verileri ile oluşturulmuştur. Çalışmanın amacı, hastaların verilerinin bulunduğu çok boyutlu bir veri tabanının kümeleme analizi yöntemleriyle incelenmesi ve veri madenciliği yöntemleri ile çok boyutlu ve büyük hacimlerdeki veri tabanlarında başarılı sonuçlar üretilmiştir.

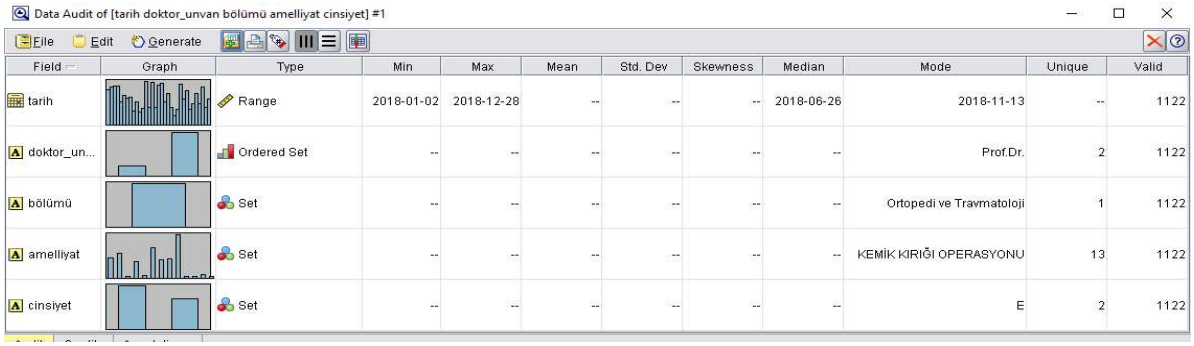
Sabri Erdem, Güzin Özdağoğlu'nun yapmış olduğu 2008 yılındaki çalışmalarında belirli bir dönem boyunca Ege Bölgesi'ndeki bir araştırma ve uygulama hastanesinin acil servisine başvuruda bulunan 214 bin hasta verisi ele alınarak, Veri Madenciliğinde sıklıkla kullanılan birliktelik kuralı yöntemiyle, veri setindeki gizli ancak anlamlılık içeren ilişkiler ortaya çıkarılmaya çalışılmıştır. Çalışma sonuçları, bölgesel özellikler taşıyabileceği düşünülen acil servislere hastaların başvuru nedenleri ve hasta profilleri açısından bir fikir vermekte ve acil servis

bölümlerinin yeniden yapılanma çalışmalarına da farklı bir açıdan yol göstererek katkıda bulunmaktadır.

Uygulama

Bu çalışmada özel bir hastane ameliyathane verileri alınarak veri madenciliği uygulamaları gerçekleştirilmektedir. Hastanenin gelecek aylardaki hasta yoğunluklarının, ameliyathane çeşitlerinin cinsiyet bazlı dağılımı ve ameliyatı gerçekleştiren doktorun ünvanı baz alınarak değişkenlerin önem sırası ve hangi istatistiklerin etkilediği belirlenmektedir.

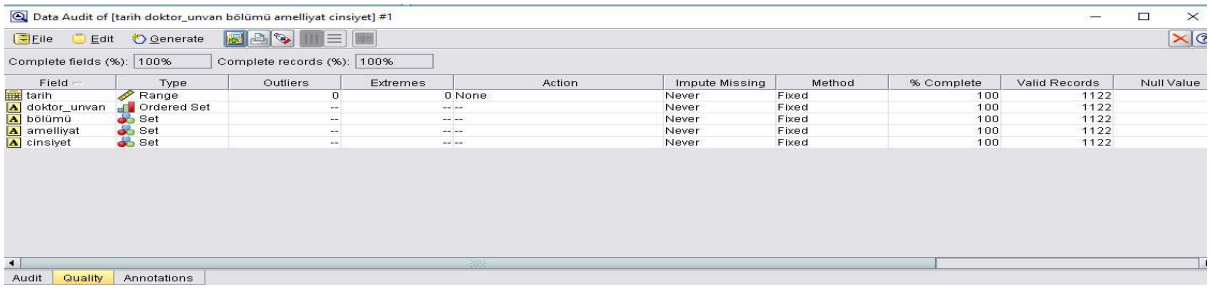
Bunun sonucunda ortopedi bölümündeki birden fazla rahatsızlığın cinsiyete ya da diğer değişkenlerden hangilerine bağlı olduğu istatistiğine varılmaktadır. Tanımlayıcı istatistikler aşağıdaki gibi doktorun ünvanı, ameliyatın türü, hastanın cinsiyeti olarak tanımlanmaktadır.


Data Audit of [tarih doktor_unvan bölümü amelliyat cinsiyet] #1

File Edit Generate

Field	Graph	Type	Min	Max	Mean	Std. Dev	Skewness	Median	Mode	Unique	Valid
tarih		Range	2018-01-02	2018-12-28	--	--	--	2018-06-26	2018-11-13	--	1122
doktor_un...		Ordered Set	--	--	--	--	--	--	Prof.Dr.	2	1122
bölümü		Set	--	--	--	--	--	--	Ortopedi ve Travmatoloji	1	1122
amelliyat		Set	--	--	--	--	--	--	KEMİK KIRIĞI OPERASYONU	13	1122
cinsiyet		Set	--	--	--	--	--	--	E	2	1122

Veri Madenciliği modeli öncesinde verinin modele hazırlanması gerekir. Aşağıdaki tabloda verinin kalitesinin iyi olduğu ve modele hazır olduğunu gösterilmektedir.


Data Audit of [tarih doktor_unvan bölümü amelliyat cinsiyet] #1

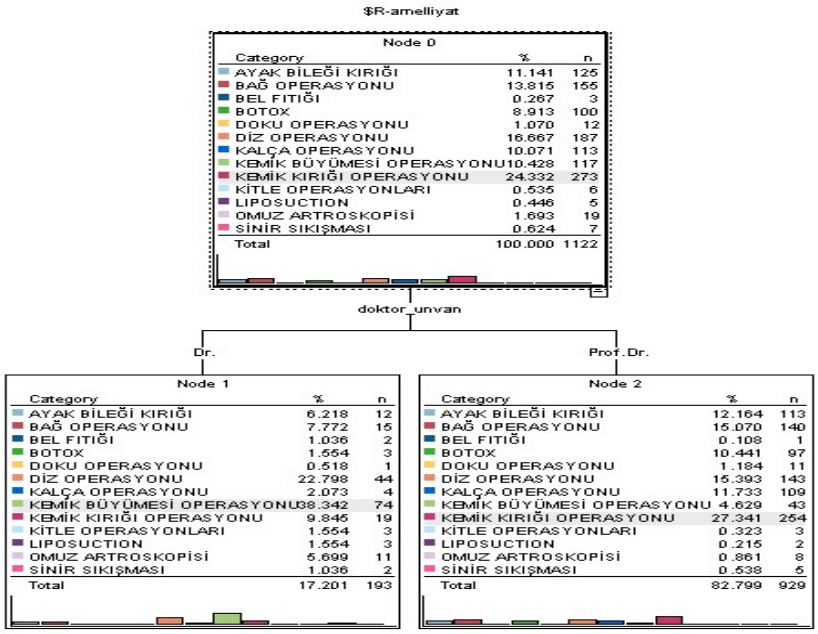
File Edit Generate

Complete fields (%): 100% Complete records (%): 100%

Field	Type	Outliers	Extremes	Action	Impute Missing	Method	% Complete	Valid Records	Null Value
tarih	Range	0	0	None	Never	Fixed	100	1122	
doktor_unvan	Ordered Set	--	--	--	Never	Fixed	100	1122	
bölümü	Set	--	--	--	Never	Fixed	100	1122	
amelliyat	Set	--	--	--	Never	Fixed	100	1122	
cinsiyet	Set	--	--	--	Never	Fixed	100	1122	

Audit Quality Annotations

Sınıflayıcı Quest algoritma ekran görüntüsü Şekil 1'de verilmiştir.

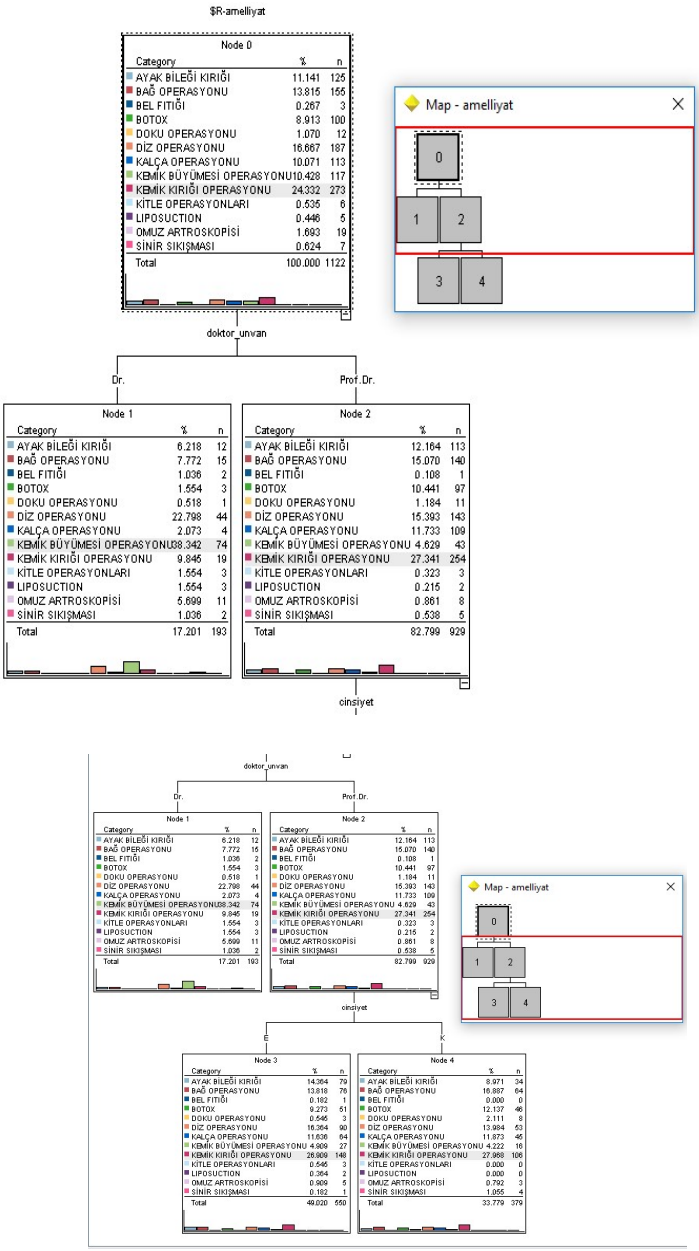


Şekil 1.Sınıflayıcı Model Quest Karar Ağacı Algoritması

Yukarıda görüldüğü gibi sonuç değişkeni, ameliyatı etkileyen en önemli değişken doktorun ünvanı olduğu görülmektedir. Diğer değişkenler Quest karar ağacı algoritmasında sonucu etkilememektedir.

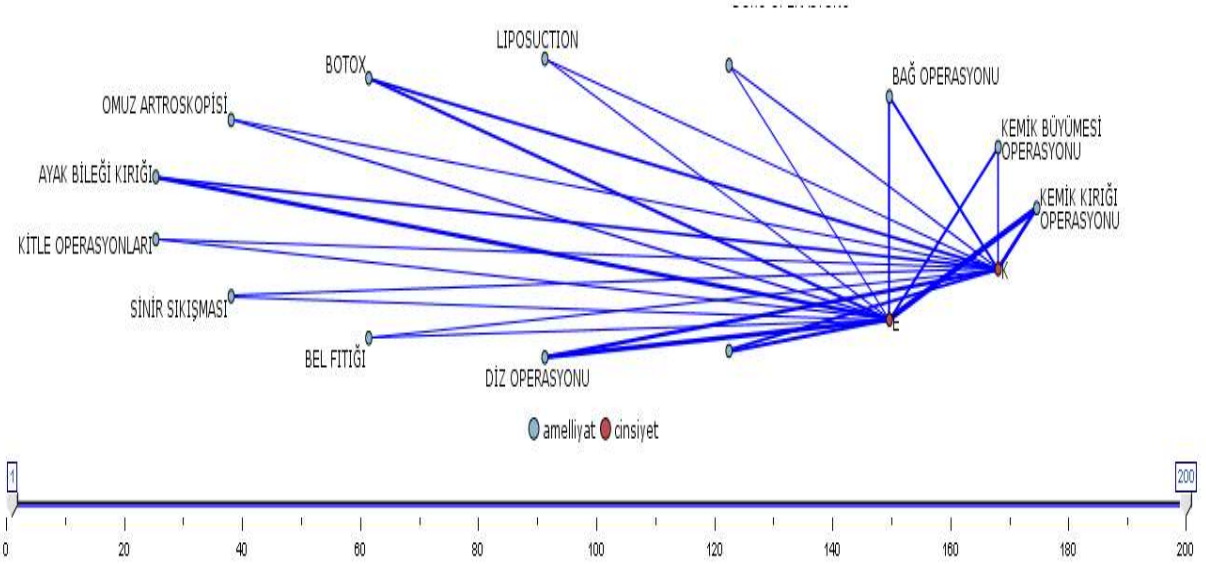
Operasyonlarda yüksek oranda profesör doktor tercih edildiği ve vakaların %24 oranında kemik kırığı operasyonu olduğu istatistiğine ulaşılmaktadır.

Sınıflayıcı Chaid algoritma ekran görüntüsü Şekil 2'de verilmiştir.

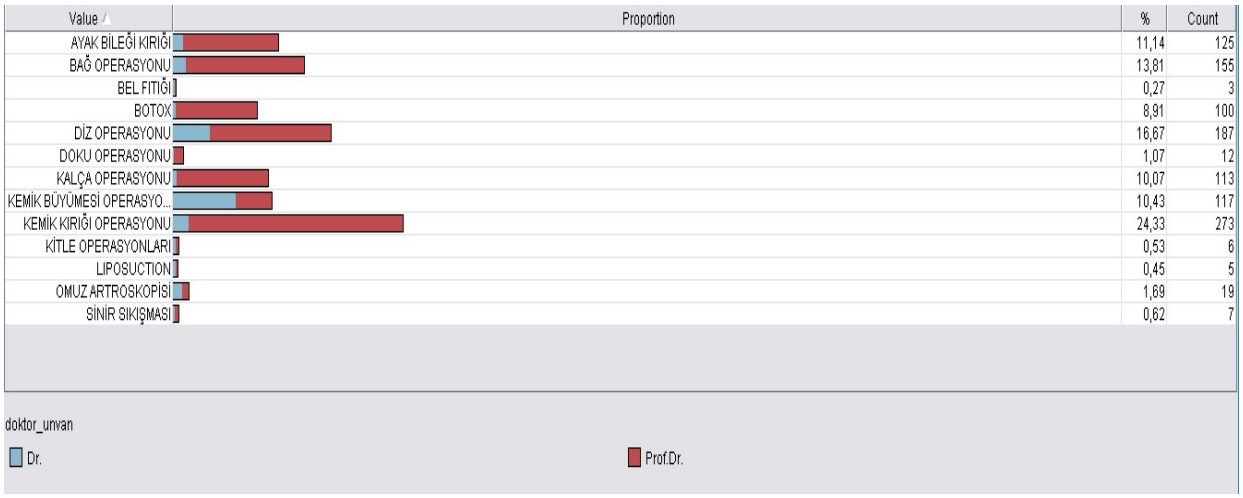


Şekil 2.Chaid Karar Ağacı Algoritması

Sonuç değişkeni ameliyatı etkileyen en önemli değişken burada da doktor ünvanı olduğu izlenmektedir. İkinci sırada cinsiyet faktörünün etkisi verilmektedir.



İlişki analizi yukarıda grafiksel verilmektedir. Kalın çizgiler yüksek kolerasyonu, ince çizgiler düşük kolerasyon göstermektedir. Grafikte de görüldüğü gibi kemik kırığı operasyonlarının yüksek kolerasyon gösterdiği sonucuna varılmaktadır.



Yukarıdaki grafikte doktor ünvanı değişkenleri alınarak, her bir operasyon için ayrı ayrı istatistik verilmektedir. Genel ortalamaya bakıldığında operasyonda tercihlerin profesör doktor yönünde olduğu görülmektedir.

Sonuç ve Öneriler

Bu çalışmada bir özel hastanenin ortopedi ameliyatları veri madenciliği ile araştırılmıştır. Araştırma sonucuna göre, erkeklerde ortopedi operasyonların kadınlar ile karşılaştırıldığında

sonucunda yüzdelik olarak daha fazla yapıldığı görülmektedir. Aynı zamanda doktor ünvanının da operasyon çeşidinde ve sayısında büyük farklılık yarattığı sonucuna varılmaktadır.

Elde edilen veriler, hasta profilinin gelecek aylar için oluşturulmasında önemli rol alacaktır. Bu bağlamda ameliyat planlaması yapılması ve steril ürün hazırlama süreçlerinde de periyotların belirlenmesinde büyük katlı sağlayacağı kaçınılmazdır. Cinsiyet bazlı elde edilen verilerin dağılımındaki kök nedene inildiğinde çok daha farklı istatistiki bulgulara rastlanabileceği öngörülmektedir.

Ameliyathane süreçlerindeki birçok belirsizliğe neden olan büyük ve anlaşılması güç veriler, işlenmesi ve kullanılabilir hale getirilmesi sonucunda anlam kazanarak birçok istatistiki bulgu elde etmemize olanak sağlamaktadır. Bu bulgular zaman içerisinde süreçlerin doğru ve daha hızlı ilerlemesine yardımcı olacağı öngörülmektedir. Hasta profilinin belirlenmesi, memnuniyet derecesini arttırmakta da oldukça önemlidir. Sağlık sektöründeki operasyon başarılarını arttırmanın temelinde hasta profilinin rolü bu açıdan kesişmektedir. Bu nedenle istatistiki bulgulara operasyonun her alanında daha fazla önem verilmesi gerektiği düşünülmektedir.

Hastaların verilerinin bulunduğu çok boyutlu bir veri tabanının incelenmesi ve veri madenciliği yöntemleri ile çok boyutlu ve büyük hacimlerdeki veri tabanlarında başarılı sonuçlara ulaşılmaktadır.

KAYNAKÇA

[1] Aydın, Sinan; Özkul, Ekrem (2015), Data Mining and An Aplication in Anadolu University Open Education System, Journal of Research in Education and Teaching, 4.

[2] Baykal, Abdullah (2006), Application Fields of Data Mining, D.Ü. Ziya Gökalp Eğitim Fakültesi Dergisi, 7, 95-107.

[3] Irmak, Sezgin; Köksal, Deniz Can; Asilkan, Özcan (2012), Predicting Future Patient Volumes of The Hospitals By Using Data Mining Methods, International Journal of Alanya Faculty of Business, 4, 101-114.

[4] Doğan, Onur (2017), Türkiye'de Veri Madenciliği Konusunda Yapılan Lisansüstü Tezler Üzerine Bir Araştırma, Gazi Üniversitesi İktisadi ve İdari Bilimler Fakültesi Dergisi, 19/3, 929-951.

[5] Seker, Sadi Evren (2015), Data Mining on Social Networks, http://www.YBSAnsiklopedi.com , 2, 1-5.

[6] Koyuncugil, Ali Serhan; Özgülbaş, Nermin (2009), Data Mining: Using and Applications in Medicine and Healthcare, Data Mining: Using and Applications in Medicine and Healthcare, Bilişim Teknolojileri Dergisi, 2.

[7] J. J. Fried (2006), "Data Mining For Quality Care", Commun Oncol, 3(1), 51.

[8] I. E. Allen (2006), C. A. Seaman, "Data Mining For Quality", Quality Progress, 39 (2), 70-74.

[9] A. Milley (2000), "Healthcare And Data Mining", Health Management Technology. 21(8), 44-46.

[10] Ma, Y., et.al. (2001). Targeting the Right Students Using Data Mining. The Seventh ACM SIGKDD International Conference on Knowledge Discovery and Data Mining. (pp.457–464), USA: San Francisco, California, 26-29 Ağustos.

[11] Irmak, S. (2009), Veri Madenciliği Yöntemleri ile Sağlık Sektörü Veritabanlarında Bilgi Keşfi: Tanımlayıcı ve Kestirimci Model Uygulamaları, Akdeniz Üniversitesi S.B.E. Yayımlanmamış Doktora Tezi.

[12] Kadılar, C. (2005), SPSS Uygulamalı Zaman Serileri Analizine Giriş, Bizim Büro Basımevi, Ankara.

Program Development for Cost Calculation in Different Hole Drilling Operations

Tugay ÜSTÜN, Yakup TURGUT

Department of Manufacturing Engineering, Gazi University, TURKEY

Abstract: Drilling operation is one of the most frequently used process in manufacturing industry. Especially in machine manufacturing sector, drilling operations take one third of the total production time. Drilling operation cost a lot, not only because of taking some serious time of total production time, but also because of the drill usage for this process. In this study, the cost comparison of three different methods evaluated for drilling operation in CNC machines. These methods are drilling cycle (G81), high-speed peck drilling cycle (G73) and deep hole peck drilling cycle (G83). The experiments were performed according to the cutting parameters suggested by the cutting tool company and the machining times measured in these three different methods. A novel program coded on Microsoft Visual Studio 2017 C#, which is able to calculate from machine amortization to workmanship, the whole process cost. Process costs can be calculated according to the number of holes in these different methods, through this program. Furthermore, drilling operation costs can be calculated for different cutting parameters too.

Keywords: Drilling, Cost Accounting, Programming

REFERENCES

[1] S. Kalpakjian And S. R. Schmid, "Manufacturing Engineering And Technology," Prentice Hall, Vol. 8th Editio, 2009.

[2] İ. Çiftçi And H. Gökçe, "Ti6al4v Titanyum Alaşımının Delinmesinde Delme Yönteminin Aşınmaya Etkisinin İncelenmesi," J. Polytech., Vol. 0900, Pp. 0–2, 2018.

[3] H. L. Tonshoff, W. Spintig, A. Neises, And W. Konig, "Machining Of Holes Developments in Drilling Techonolgy," Ann. Cırp, P. 43, 1994.

[4] S. Yagmur, A. Acir, U. Seker, and M. Gunay, "An Experimental Investigation of Effect of Cutting Parameters On Cutting Zone Temperature in Drilling," J. Fac. Eng. Archit. Gazi Univ., Vol. 28, No. 1, Pp. 1–6, 2013.

[5] N. Bıçakcı, "Ti6al4v Titanyum Alaşımının Delinebilirliğinin Araştırılması," Karabük Üniversitesi, 2015.

[6] S. Sharif And E. A. Rahim, "Performance of Coated and Uncoated Carbide Tools When Drilling Titanium Alloy Ti6al4v," J. Mater. Process. Technol., Vol. 185, Pp. 72–76, 2007.

[7] P. K. Shetty, R. Shetty, D. Shetty, F. Rehaman, And T. K. Jose, "Machinability Study on Dry Drilling Of Titanium Alloy Ti6al4v Using L9 Orthogonal Array"," Procedia Mater. Sci., Vol. 5, Pp. 2605–2614, 2014.

[8] F. Bıçakçı And A. Tekin, "Sanayi İşletmelerinde Faaliyet Tabanlı Maliyet Sistemine Geçiş Çabaları ve Bir Uygulama," Selçuk Üniversitesi, 2006.

[9] S. Baysan And M. B. Durmuşoğlu, "Değişen Rekabet Koşullarında Değişmeyen Maliyet Muhasebesine Yeni Bir Soluk: Yalın Maliyet Muhasebesi," Altı Sigma Yalın Konf., Pp. 1–5, 2008.

[10] U. Pamukoğlu And C. Göloğlu, "Delik İşlemleri için Maliyet Merkezli Bir Sistem," J. Eng. Sci., Vol. 10, Pp. 31–39, 2004.

[11] "Makine Takım Endüstrisi A.Ş." [Online]. Available: Https://Www.Makinatakim.Com.Tr/Hss-Din-338-Silindirik-Sapli-Kisamatkap-Ucu-Haddeli/.

[12] G. Meral, "Aısı 1050 Malzemenin Delinmesinde Delme Parametrelerinin Kesme Kuvvetleri ve Delik Kalitesi Üzerindeki Etkisinin Araştırılması," Gazi Üniversitesi, 2010.

[13] O. E. Çelik, "Farklı Yazılım Sistemleri Kullanarak Bir Yolcu Uçağı Panelinin Üretimi için Otomatik Perçin Tezgâhı NC Kodlarının Oluşturulması ve Karşılaştırılması," Hacettepe Üniversitesi, 2014.

[14] P. Smid, Cnc Programming Handbook, 3rd Ed. New York: Industrial Press Inc., 2008.

[15] M. Akkurt, Talaş Kaldırma Yöntemleri ve Takım Tezgâhları. İstanbul: Birsen Yayınevi, 2004.

Multi-Objective Optimization of Hard Turning: Non-Dominated Sorting Genetic Algorithm-II Approach

Ahmet KOCATÜRK, Bülent ALTUNKAYNAK

Department of Statistics, Gazi University, TURKEY

Abstract: Multi-objective optimization problems allow multiple purpose to be simultaneously optimized. The nondominated sorting genetic algorithm II (NSGA-II), which is one of the most effective multi-objective heuristic methods in the solution of multi-objective optimization problems, is widely used in the literature. NSGA-II obtains a Pareto optimal solutions, known as a set of dominant solutions without requiring any prior knowledge in one run. The NSGA-II is more useful than the classical genetic algorithm, minimizing the computational complexity by calculating the fast dominated sorting approach and the crowded distance without having to repeat for each solution. In this study, NSGA-II method was used to optimize the cutting parameters of hard materials turning. In the experimental studies, the regression models based on the cutting velocity, feed rate and depth of cut parameters represent three different objective functions. This optimization problem, which has five objective functions with three variables, has been discussed by NSGA-II method. The optimal solution of these functions is to use the NSGA-II method to find the most suitable set of Pareto solutions. The solutions obtained by using NSGAII method have been found to be successful in multi-objective optimization problems. In addition, decision makers from the optimal solutions can choose the most suitable solution according to their importance in the objective functions.

Keywords: Multi-Objective Optimization, NSGA-II, Pareto Solution Set, Hard Turning

Introduction

The turning of hard materials in the machinery industry changes the operational cost due to various parameters. It is also desirable to optimize surface roughness in turning. Thus, the parameters used in the turning of hard materials could be linked to a common functionality and could be optimized

based on the economic criteria. This kind of optimization problems with multiple purposes is very common in real life. Simultaneous optimization of multiple purposes is called multi-objective optimization problems [1].

The solution of multi-objective problems is more complex and difficult than single-objective problems. The local minimum or local maximum points are mathematically the best solutions for deciding on optimization problems with a single purpose. However, it is not possible to achieve the best solution value in optimizing conflicting goals [2].

Therefore, intuitive methods such as artificial neural networks, genetic algorithms or simulated annealing are widely used in the solution of multi-objective problems [3-5]. However, these methods need to be repeated many times to find each solution. As a result, the computational complexity of classical heuristic methods is considerably greater for multi-objective optimizations. Multi-objective heuristic methods have been developed to prevent this [6-10].

In literature studies, one of the most effective methods for solving multi-objective optimization problems is the Non-Dominated Sorting Genetic Algorithm-II (NSGA-II) [11-15].

NSGA-II finds a set of non-dominated solutions (Pareto solution set) in single run without requiring any prior knowledge. The basic operating principle of the algorithm is based on the fast dominant sequence and to calculate the crowded distance according to the order of dominance in determining the Perato surface. First, a randomly selected starting population is generated. Then, genetic algorithm operators are determined and a new generation population is obtained. The current and new population are combined and sorted by dominant sequence and crowded distance. Individuals who determine the pareto surface are selected and the Pareto solution set is obtained. The search process of the algorithm continues until the generation number.

In this study, NSGA-II method was used to optimize the cutting parameters of hard materials turning. After some preliminary studies at [16] used in the study of hardened steel turning, parameters were estimated for two different hardness levels. In experimental studies, regression models based on cutting velocity, feed rate and depth of cut parameters represent three different objective functions. These functions are: tangential force (P1), axial force (P2) and radial force

(P3). In addition, the surface roughness (R) and tool life (T) models are also to be optimized simultaneously. This optimization problem, which has five objective functions with three variables, has been discussed by NSGA-II method.

The optimal solution of these functions is to use the NSGA-II method to find the most suitable set of Pareto solutions. Pareto optimal solution set was obtained for problems. According to the results obtained from the study conducted by [1], it was shown that the results obtained using NSGA-II method were better than the results obtained by genetic algorithm.

Methodology

Non-Dominated Sorting Genetic Algorithm-II (NSGA-II) is an effective heuristic method that finds Pareto optimal solutions for multi-objective optimization [16]. The NSGA-II algorithm was developed by revising the deficiencies of the NSGA algorithm developed by [9]. The NSGA examines whether individuals are dominating each other for all purposes. For this comparison $O(MN^3)$ calculation complexity is required. Where M is the target number and N is the number of population. In the NSGA-II algorithm, computational complexity was reduced to $O(MN^2)$. When looking at the dominance of individuals for this process, instead of comparing with all populations, the first surface in the Pareto solution compares with the individuals in the solution. Since the NSGA-II algorithm has low computational complexity and takes into account effectiveness, it is implemented in many areas.

NSGA-II is designed based on genetic algorithm. In addition to the steps of the genetic algorithm, the dominant sequence and crowded distance calculation are also applied. First, the initial population is created in NSGA-II. Solutions from the initial population are ranked according to their superiority in the Pareto solution. In this ranking, the fast dominant sorting method is used.

- **Fast Non-Dominated Sorting Approach**

In order to identify individuals in a population at the dominant surface in the first row, the individual dominance of each solution is compared with the other solutions in the population. This

process continues for all solutions on the first dominant surface. For the second-order dominant surface solution, the first-order solutions are ignored and the same process is repeated. Thus, individuals are classified into different sets of dominance according to their degree of dominance.

The number of dominant and the dominant set of solutions is determined for each p element in the P population. The dominance value of each element on the first dominant surface is set to zero. It is then compared with other elements in the solution set. After that, the element with the dominance number zero is placed in another set of Q. The Q set forms the second dominant surface. A comparison is made with the elements in the Q set and the process continues in this way. The fast non-dominated sorting algorithm is as follows:

Step 1: For each $p \in P$;

- $S_p = \emptyset$, The solution set in which the p solution is dominant is defined.
- $n_p = 0$, The number of solutions that dominate the p solution is defined.
- For each $q \in P$;

If p solution is dominant in q solution, q solution is added to a set of solutions where p solution is dominant, $S_p = S_p \cup \{q\}$.

If q solution is dominant in p solution, the dominance counter for the solution p is increased by one, $n_p = n_p + 1$.

If $n_p = 0$, there is no solution to the p solution and p solution belongs to the first surface, $p_{rank} = 1$. The solution p is added to the surface 1 and the first surface is updated, $F_1 = F_1 \cup \{p\}$.

Step 2: For each $p \in P$; repeat Step 1.

Step 3: The surface counter is taken as 1, $i = 1$.

Step 4: As long as the i. surface is different from the null set, $F_i \neq \emptyset$, the following process is repeated:

- For the surface $(i + 1)$, the cluster where the elements are collected is taken as a null, $Q = \emptyset$.
- For each $p \in F_i$ and $q \in S_p$;

Reduction of the dominance counter, $n_q = n_q - 1$.

If $n_q = 0$, no solution on the sequential surface dominant the q solution. $q_{rank} = 1$ and Q set is updated, $Q = Q \cup \{q\}$.

- Surface counter is incremented by 1, $i = i + 1$.

- New surface is assigned to Q set, $F_i = Q$.

Because the dominance counter and the dominant solution set are used for each solution, the order of Pareto solutions is faster than NSGA. Therefore, the sorting algorithm defined is called the fast dominant sorting algorithm.

- **Crowded Distance**

The NSGA-II algorithm uses the crowded distance approach so that the solutions within the Pareto solution set are spread and varied. The aim here is to determine the Euclidean distance between the functional values of the solutions on a predetermined dominated surface.

The crowded distance of a solution is the distance between the related solution and neighboring solutions on the surface on which the solution is located. It is used to estimate the perimeter of cuboid, which is formed by using corners which are the closest neighbors, and this is called crowded distance. For the calculation of the crowded distance, the solutions must be on the same optimal surface. The crowded distance algorithm is as follows:

Step 1: The starting distance for all element j on element F_i is defined as zero, $F_i(d_j) = 0$.
Step 2: For each objective function m;

- The elements on the surface of F_i are sorted based on the objective function m, $I = rank(F_i, m)$
- The boundary value elements on the surface F_i are assigned an infinite distance, $I(d_1) = \infty$ and $I(d_n) = \infty$
- All other points are calculated,

$$I(d_k) = I(d_k) + \frac{I(k+1).m - I(k-1).m}{f_m^{max} - f_m^{min}},$$

$k = 2,3, \dots, n-1.$

where $I(k+1).m$, indicates the m. function value of element k. in set of I.

New population Q_0 is generated using crossover and mutation operators for the first population P_0, sorted by the fast non-dominant sorting algorithm and calculated by the crowded distance algorithm.

- **Tournament Selection Operator**

The selection of the crowded tournament is the process of selecting units for the match pool. The solution is selected up to the size of the match pool. In this method, two individuals are randomly selected. Two individuals are selected according to their dominance order and crowd distance in the population.

This tournament process is concluded with the completion of the population size initially defined for the next iteration.

- **Crossover and Mutation Operator**

The crossover process is a genetic operator designed to exchange genes between parents to produce two new chromosomes and to obtain the best properties from each. This process takes place during the evolution period according to the user-definable crossover probability. The crossover operator usually uses single-point crossover, two point and k point crossover, crossover for ordered lists, and uniform crossover.

Mutations are hereditary changes that occur due to other reasons than gene change in a chromosome. It occurs when one or more gene values on a chromosome take a different value than the first state. In this process, the addition of new gene values to the gene pool can be achieved by a better solution than the previously formed solution. It is also used to prevent the problem from being stuck in local solutions in the population.

- **Application**

Multi-objective optimization problems can be formulated as follows:

$$\min_x f(x) = [\, f_1(x),\ f_2(x), \dots,\ f_m(x)\,] \quad (1)$$

where $x = (x_1, x_2, \dots, x_n)$ is the *n* decision variables and $f(x)$ is the *m* objective functions.

Reference [17] used three decision variables and five objective functions. So the model for optimization problem:

$$\min_x f(x) = [\, f_1(x),\ f_2(x),\ f_3(x),\ f_4(x),\ f_5(x)] \quad (2)$$

where;

$f_1(x)$: surface roughness (R),

$f_2(x)$: tangeltial force (P1),

$f_3(x)$: axial force (P2),

$f_4(x)$: radial force (P3) and

$f_5(x)$: tool life (T) function.

Also x_1 refers to cutting velocity, x_2 refers to feed rate and x_3 refers to depth of cut variables.

Regression model for surface roughness (R):

$$\begin{aligned} f_1(x) = {} & 12.7937 - 0.03118 * x[1] - 28.8786 * x[2] - \\ & 2.8599 * x[3] + 0.0358 * x[1] * x[2] + \\ & 0.00236 * x[1] * x[3] + 11 * x[2] * x[3] + \\ & 0.0000381 * x[1] * x[1] + 32.039 * x[2] * x[2] + \\ & 0.2853 * x[3] * x[3] \end{aligned}$$

Regression model for tangeltial force (P1):

$$\begin{aligned} f_2(x) = {} & -373.0294 + 0.5308 * x[1] + 788.3909 * x[2] + \\ & 697.2733 * x[3] - 7.2420 * x[1] * x[2] - \\ & 1.9860 * x[1] * x[3] + 235 * x[2] * x[3] + \\ & 0.0075 * x[1] * x[1] + 6659.8597 * x[2] * x[2] + \\ & 0.5986 * x[3] * x[3] \end{aligned}$$

Regression model for axial force (P2):

$$\begin{aligned} f_3(x) = {} & 375.4951 - 2.971 * x[1] - 360.2475 * x[2] + \\ & 76.6834 * x[3] + 7.9052 * x[1] * x[2] - \\ & 0.4 * x[1] * x[3] - 145 * x[2] * x[3] + \end{aligned}$$

$$0.00398 * x[1] * x[1] - 1528.4596 * x[2] * x[2] + 66.7154 * x[3] * x[3]$$

Regression model for radial force (P3):

$$f_4(x) = 239.6985 - 2.4094 * x[1] + 755.0606 * x[2] + 133.1862 * x[3] - 0.05559 * x[1] * x[2] + 0.2472 * x[1] * x[3] - 585 * x[2] * x[3] + 0.0041565 * x[1] * x[1] + 2593.51 * x[2] * x[2] + 22.9351 * x[3] * x[3]$$

Regression model for tool life (T):

$$f_5(x) = (x[1]\text{^}0.5937) * (x[2]\text{^}0.4697) * (x[3]\text{^}0.4743)$$

The parameter values used for the NSGA-II method were determined by the number of variables, the number of objective functions, the lower and upper limits of the variables for each optimization problem. The lower and upper limits of the variables are shown in Table I. The parameters used in the NSGA-II method are shown in Table II.

TABLE VIII. LOWER LIMIT AND UPPER LIMIT VALUES OF CUTTING PARAMETERS

Parameters	**Lower Bound**	**Upper Bound**
cutting velocity (x_1)	142	265
feed rate (x_2)	0.15	0.25
depth of cut (x_3)	1	2

TABLE IX. PARAMETERS USED IN NSGA-II METHOD

Parameters	**Value**
Population size	100
Generation size	100
Crossover rate	0.7
Mutation rate	0.2

Pareto solution set was obtained for optimal solution of these functions. The results obtained from the studies conducted by [1] and [17] were compared with the most appropriate results in the NSGA-II method set and shown in Table III.

TABLE X. COMPARISON OF OPTIMIZATION RESULTS

	x_1	x_2	x_3	R	P1	P2	P3	T
[17]	235	0.15	1.00	4.85	445.13	115.39	199.14	40.34
[1]	225	0.15	1.15	4.90	479.09	135.15	224.28	38.74
NSGA-II	235	0.15	1.00	**4.84**	445.97	**114.71**	**198.96**	**40.26**

According to the results of the analysis, the optimal results of the objective functions in the study by [17]: 4.85 for the surface roughness (R), 445.13 for the tangeltial force (P1), 115.39 for the axial force (P2), 199.14 for the radial force (P3) and 40.34 for tool life (T) was found. In case the variables have the same value, the objective function values obtained by NSGA-II method are: 4.84 for the surface roughness (R), 445.97 for the tangeltial force (P1), 114.71 for the axial force (P2), 198.96 for the radial force (P3) and 40.26 for tool life (T) was found. Better results were obtained with the objective function values obtained by NSGA-II method. In the study conducted by [1], although the results obtained by the genetic algorithm method of the optimization problem are close to the results obtained with the other two methods, it can not be said that it is better.

- **Results**

In this study, NSGA-II approach which is effective in heuristic methods has been used for the solution of multi-objective optimization problems. The surface roughness (R), tangential force (P1), axial force (P2), radial force (P3) and tool life (T) functions are optimized for estimation of the cutting parameters effecting the turning of hard materials as a multi-objective optimization problem. In addition, the results were compared with the results of the study by [1] and [17] and it was shown that the results obtained with the NSGA-II method were better.

According to these results, the solutions obtained by using NSGA-II method have been successful in multi-objective optimization problems. When classical methods for optimization problems with

multiple objective functions are insufficient, an efficient method NSGA-II can be used in heuristic methods. In addition, decision makers from the optimal solutions can choose the most suitable solution according to their importance in the objective functions.

REFERENCES

[1] O. Manav, And S. Chinchanikar, "Multi-Objective Optimization of Hard Turning: A Genetic Algorithm Approach" Materials Today: Proceedings, Vol. 5, No. 5, Pp. 12240-12248, 2018.

[2] R. C. Peralta, A. Forghani, And H. Fayad, "Multiobjective Genetic Algorithm Conjunctive Use Optimization For Production, Cost, And Energy With Dynamic Return Flow" Journal of Hydrology, Vol. 511, Pp. 776-785, 2014.

[3] Ç. Elmas, "Artificial Intelligence Applications, Artificial Neural Networks - Fuzzy Logic - Genetic Algorithm", 2nd Ed., Ankara: Seçkin, 2011.

[4] D. E. Goldberg, "Genetic Algorithms İn Search, Optimization, and Machine Learning", 1st Ed., Usa: Addison-Wesley, 1989.

[5] S. Kirkpatrick, C. D. Gelatt, And M. P. Vecchi, "Optimization By Simulated Annealing" Science, Vol. 220, No. 4598, Pp. 671-680, 1983.

[6] D. Kalyanmoy, "Multi Objective Optimization Using Evolutionary Algorithms" 1st Ed., New York: John Wiley & Sons, Pp. 124, 2001.

[7] C. M. Fonesca, And P. J. Fleming, "Genetic Algorithms For Multiobjective Optimization: Formulation, Discussion And Generalization" in Proc Of The Fifth Int. Conf. On Genetic Algorithms, Pp. 415-423, July 1993.

[8] J. Horn, N. Nafploitis, And D. E. Goldberg, "A Niched Pareto Genetic Algorithm For Multiobjective Optimization" Nj: Ieee Press, Pp. 82-87, 1994 [Proceedings Of The First Ieee Conference On Evolutionary Computation].

[9] N. Srinivas, And K. Deb, "Muiltiobjective Optimization Using Nondominated Sorting in Genetic Algorithms" Evolutionary Computation, Vol. 2, No. 3, Pp. 221-48, 1995.

[10] E. Zitzler, And L. Thiele, "Multiobjective Optimization Using Evolutionary Algorithms-A Comparative Case Study" International Conference On Parallel Problem Solving From Nature, Berlin, Germany: Springer, Pp. 292-301, 1998.

[11] B. Wang, Y. Liang, T. Zheng, M. Yuan, And H. Zhang, "Multiobjective Site Selection Optimization of The Gas-Gathering Station Using Nsga-Iı" Process Safety And Environmental Protection, Vol. 119, Pp. 350-359, 2018.

[12] S. M. Wang, S. Ma, And W. Y. Duan, "Seakeeping Optimization of Trimaran Outrigger Layout Based On Nsga-Iı" Applied Ocean Research, Vol. 78, Pp. 110-122, 2018.

[13] Y. Yang, L. Cao, C. Wang, Q. Zhou, And P. Jiang, "Multi-Objective Process Parameters Optimization Of Hot-Wire Laser Welding Using Ensemble Of Metamodels And Nsga-Iı" Robotics And Computerıntegrated Manufacturing, Vol. 53, Pp. 141-152, 2018.

[14] T. Vo-Duy, D. Duong-Gia, V. Ho-Huu, H. C. Vu-Do, And T. Nguyenthoi, "Multi-Objective Optimization of Laminated Composite Beam Structures Using Nsga-Iı Algorithm" Composite Structures, Vol. 168, Pp. 498-509, 2017.

[15] N. Alikar, S. M. Mousavi, R. A. R. Ghazilla, M. Tavana, And E. U. Olugu, "Application Of The Nsga-Iı Algorithm To A Multi-Period Inventory-Redundancy Allocation Problem İn A Series-Parallel System" Reliability Engineering & System Safety, Vol. 160, Pp. 1-10, 2017.

[16] K. Deb, A. Pratap, S. Agarwal, And T. A. M. T. Meyarivan, "A Fast And Elitist Multiobjective Genetic Algorithm: Nsga-II" IEEE Transactions On Evolutionary Computation, Vol. 6, No. 2, Pp.182-197, 2002.

[17] S. Chinchanikar, And S. K. Choudhury, "Effect of Work Material Hardness And Cutting Parameters on Performance Of Coated Carbide Tool When Turning Hardened Steel: An Optimization Approach" Measurement, Vol. 46, No. 4, Pp. 1572-1584, 2013.

CO^2 Emission and Energy Consumption for Different Climate and Building Materials

Salih HİMMETOĞLU, Yılmaz DELİCE, Emel KIZILKAYA AYDOĞAN

Department of Industrial Engineering, Erciyes University, TURKEY

Abstract: With the development of technology from past to present, the types, properties and product range of the materials used in the buildings are quite developed. Therefore, the effects of climate, environmental conditions and energy consumption cannot be ignored for selecting these materials used in the buildings. Usage of materials with the same characteristics for buildings to be built on different climate may lead to adverse effects about energy-saving and green gasses. Furthermore, the use of the same materials may not be a proper approach even in buildings with a different purpose. In this study, forecasting of energy consumption and CO2 emission is analyzed by utilizing artificial neural network structure according to different climate criteria and material characteristics for public buildings built in recent years. The Effect levels to energy consumption and CO2 emission of the building materials and the climate criteria are determined for buildings serving the same using purpose in different climate characteristics. For the study, different pilot regions where the public buildings are located are chosen according to climatic characteristics and five different building materials used in these public buildings are taken into account. When the results compare according to CO2 emission and energy consumption, it was observed that the conditions which obtain most efficient results are different.

Keywords: Artificial Neural Network, Energy Consumption, CO2 Emission, Data Mining, Forecasting, Public Buildings

REFERENCES

[1] M. Bayraktar, ―A Methodology for Energy Optimization of Buildings Considering Simultaneously Building Envelope HVAC and Renewable System Parameters, Ph. D. Thesis. 2015.

[2] M. Tekin, Production Management, 6th ed., vol. 1. Konya: Günay Ofset, 2009.

[3] R. M. Kirby, "A Comparison of Short And Medium Range Statistical Forecasting Methods" Management Science Vol. 13 Pp. 4, B-202, 1966.

[4] R. L. Carlson And M. Umble, "Statistical Demand Functions For Automobiles and Their Use For Forecasting İn An Energy Crisis," The Journal Of Business, Pp. 193 – 204, 1980.

[5] S. L. Zhou, T. A. Mcmahon, A. Walton And J. Lewis, ―Forecasting Operational Demand For An Urban Water Supply Zone, Journal Of Hydrology, Vol. 259(1-4), Pp. 189-202, 2002.

[6] E. E. Cahow, ―Forecast Of Demand For Chronic Care Nursing Home Services: 2005-2025, Brandeis University, Waltham Massachusetts, USA, Ph. D. Thesis, 2003.

[7] K. Matuyama, T. Sumita And D. Wakayama, ―Periodic Forecast and Feedback To Maintain Target Inventory Level, International Journal Of Production Economics, Vol. 118(1), Pp. 298-304, 2009.

[8] R. Fildes, P. Goodwin, M. Lawrence And K. Nikolopoulos, ―Effective Forecasting and Judgmental Adjustments: An Empirical Evaluation And Strategies For İmprovement in Supply-Chain Planning, International Journal Of Forecasting, Vol. 25, Pp. 3–23, 2009.

[9] X, Xu, Y. Qi And Z. Hua, ―Forecasting Demand of Commodities After Natural Disasters, Expert Systems With Applications, Vol. 37, Pp. 4313-4317, 2010.

[10] W. E. Griffiths, L. S. Newton And C. J. O'Donnell, ―Predictive Densities For Models With Stochastic Regressors And Inequality Constraints: Forecasting Local-Area Wheat Yield, International Journal of Forecasting, Vol. 26(2), Pp. 397-412, 2010.

[11] M. Ghodsi, ―A Brief Review Of Recent Data Mining Applications İn The Energy İndustry, International Journal Of Energy And Statistics, Vol. 2(01), Pp. 49 - 57, 2014.

[12] M. W. Gardner And S. R. Dorling, ―Artificial Neural Networks (The Multilayer Perceptron)―A Review of Applications in The Atmospheric Sciences, Atmospheric Environment, Vol. 32(14-15), Pp. 2627-2636, 1998.

[13] X. Sun, D. K. Gauri And S. Webster, ―Forecasting For Cruise Line Revenue Management, Journal of Revenue and Pricing Management, Doi:10.1057/Rpm.2009.55, 2010.

[14] S. Chung, ―Demand Modeling And Analysis For The Management of Underground Infrastructure Systems, Purdue University, USA, Ph. D. Thesis, 2001.

[15] H. Shuai, G. Qingwu And J. Wu, ―A New Multi-Method Combination Forecasting Model For ESDD Predicting, IEEE Asia-Pacific Power And Energy Engineering Conference Pp. 1 – 4, March 2010.

[16] M. R. Devi And R. Manonmani, " Electricity Forecasting Using Data Mining Techniques in Tamil Nadu and Other Countries - A Survey," International Journal of Emerging Trends in Engineering And Development, Vol. 6(2), Pp. 295–302, 2012.

[17] L. Fugon, J. Juban And G. Kariniotakis, ―Data Mining For Wind Power Forecasting, European Wind Energy Conference & Exhibition EWEC 2008, Brussels, Belgium, Pp. 1–6, 2008.

[18] S. R. Mohandes, X. Zhang And A. Mahdiyar, ―A Comprehensive Review on the Application of Artificial Neural Networks in Building Energy Analysis. Neurocomputing, Vol. 340, Pp. 55-75, 2019.

[19] X. Cipriano, J. Carbonell And J. Cipriano, "Monitoring And Modeling Energy Efficiency of Municipal Public Buildings: Case Study in Catalonia Region," International Journal Of Sustainable Energy, Vol. 28(1-3), Pp. 3-18, 2009.

[20] D. Hawkins, S. M. Hong, R. Raslan, D. Mumovic, And S. Hanna, ―Determinants of Energy Use in UK Higher Education Buildings Using Statistical And Artificial Neural Network Methods, International Journal of Sustainable Built Environment, Vol. 1(1), Pp. 50-63, 2012.

[21] H. Son, S. Lee, C. Kim, "An Artificial Neural Network-Based Prediction Of Government-Owned Building Energy Consumption With Design Variables," Pro-Ceedings Of The ICSDEC 2012 Developing The Frontiers Of Sustainable Design Engineering Construction, Pp. 1–10, 2013.

[22] C. Buratti, M. Barbanera And D. Palladino. "An Original Tool For Checking Energy Performance and Certification Of Buildings By Means Of Artificial Neural Networks," Applied Energy, Vol. 120, Pp. 125- 132, 2014.

[23] S. Mishra And V. K. Singh, ―Monthly Energy Consumption Forecasting Based on Windowed Momentum Neural Network, IFAC-Papersonline, 2015, Vol. 48(30) Pp. 433-438, 2015.

[24] Q. Dong, K. Xing And H. Zhang, ―Artificial Neural Network For Assessment of Energy Consumption and Cost For Cross Laminated Timber Office Building İn Severe Cold Regions, Sustainability, Vol. 10(1), Pp. 84, 2017.

[25] C. Deb, S. E. Lee And M. Santamouris, ―Using Artificial Neural Networks to Assess HVAC Related Energy Saving İn Retrofitted Office Buildings, Solar Energy, Vol. 163, Pp. 32-44, 2018.

Bitcoin Price Forecasting with Multivariate Long Short Term Memory (LSTM) Deep Learning Method

Ali Osman ÇIBIKDİKEN[1], Ebru Şeyma KARAKOYUN[2]

[1]Department of Computer Engineering, Necmettin Erbakan University,

[2]Department of Industrial Engineering, Karabük University, TURKEY

Abstract: Long Short Term Memory (LSTM) is one of the deepest learning methods capable of learning along a chain. The method has a chain of modules able to repeating information and transferring it to the next module. Due to this feature, it is a convenient method for data sets consisting of time-dependent information such as finance. Bitcoin, using blockchain technology, has become one of the most popular cryptocurrencies today. Bitcoin data is a time series. In this study, price estimation model is proposed by using Long-Short Term Memory method for a Bitcoin price estimation for multivariate time series consisting of opening price, closing price, highest price, lowest price, Bitcoin volume, Purchasing volüme and weighted price variables. In addition, the application has been developed in Python programming language.

REFERENCES

[1] S. Hochreiter, J. Schmidhuber, "Long Short-Term Memory", Neural Computation, 9(8):1735-1780, 1997.

[2] S. Nakamoto, "Bitcoin: A Peer-To-Peer Electronic Cash System", 2008.

[3] A. Navon, Y. Keller, "Financial Time Series Prediction Using Deep Learning", Https://Arxiv.Org/Abs/1711.04174, 2017.

[4] F. Thomas, K. Christopher, "Deep Learning With Long Short-Term Memory Networks For Financial Market Predictions", Fau Discussion Papers İn Economics, No. 11/2017, 2017.

[5] W. W. S. Wei, Time Series Analysis Univariate and Multivariate Methods, Pearson Education, 2006.

[6] F. Chollet, Keras, Https://Github.Com/Fchollet/Keras, 2016.

Covering-Based Generalized IF-Rough Set Models For A Selecting HVAC System

Salih HİMMETOĞLU, Emel KIZILKAYA AYDOĞAN, Yılmaz DELİCE

Department of Industrial Engineering, Erciyes University, TURKEY

Abstract: In today's urban life, most of the people's time is spent at home or work office. For this reason, design and selecting of building materials that control heating (H), ventilation (V) and air conditioning (AC) are essential. The building materials called, in short HVAC, must simultaneously provide high comfort, low cost and high energy productivity. Furthermore, HVAC must be appropriately designed to prevent adverse effects on the environment and climate. In this study, nine different HVAC systems were examined according to nine different criteria over cost, pollution, comfort and energy which are considered as four main factors in the selection of HVAC systems. Since some of the criteria for HVAC systems are described as linguistic, it is not possible to evaluate the systems with traditional methods using crisp values. Therefore, we propose generalized intuitionistic fuzzy (IF) - rough set models which are a new and flexible method. IF-neighborhoods are formed by using IF-implicator and IF-t norms, and upper and lower approximations in rough set theory are calculated according to the neighborhoods. Covering-based generalized IF rough set models are generated by using the approximations and IF-TOPSIS method. According to the obtained results, we can see that the proposed method is an appropriate decision-making method which considers the uncertainties in the linguistic expressions for selecting the most suitable HVAC system.

Keywords: HVAC, IF-Covering, Rough Set Theory, Fuzzy Sets, Approximations

REFERENCES

[1] M. Bayraktar, "A Methodology for Energy Optimization of Buildings Considering Simultaneously Building Envelope HVAC and Renewable System Parameters," Ph. D. Thesis, 2015.

[2] E. Burak, F. E. Boran, and M. Kurt, "Ergonomic Product Concept Selection Using Intuitionistic Fuzzy TOPSIS Method," Journal of Engineering Sciences and Design, vol. 3(3), pp. 433-440, 2015.

[3] L. Zhang, J. Zhan, And Z. Xu, "Covering-Based Generalized If Rough Sets With Applications To Multi-Attribute Decision-Making," Information Sciences, Vol. 478, Pp. 275-302, 2019.

[4] J. S. Dodgson, M. Spackman, A. Pearman, And L. D. Phillips, Multicriteria Analysis: A Manual, Department For Communities And Local Government, London, 2009.

[5] J. P. Huang, K. L. Poh, And B. W. Ang, "Decision Analysis in Energy And Environmental Modeling," Energy, Vol. 20(9), Pp. 843-855, 1995.

[6] J. D. Balcomb, And A. Curtner, "Multi-Criteria Decision-Making Process For Buildings. In Collection Of Technical Papers." Ieee 35th Intersociety Energy Conversion Engineering Conference And Exhibit, Vol 1, Pp. 528-535, 2000.

[7] S. De Wit, And G. Augenbroe, "Analysis Of Uncertainty İn Building Design Evaluations And İts İmplications," Energy And Buildings, Vol. 34(9), Pp. 951-958, 2002.

[8] C. J. Hopfe, Uncertainty And Sensitivity Analysis İn Building Performance Simulation For Decision Support And Design Optimization, Eindhoven University, Ph. D. Thesis, 2009.

[9] S. H. Kim, And G. Augenbroe, Ventilation Operation in Hospital İsolation Room: A Multi-Criterion Assessment Considering Organizational Behavior in Building Simulation, Pp. 27-30, July 2009.

[10] Y. J. Kim, K. U. Ahn, And C. S. Park, "Decision Making of Hvac System Using Bayesian Markov Chain Monte Carlo Method," Energy And Buildings, Vol. 72, Pp. 112-121, 2014

[11] P. Huang, G. Huang, And Y. Wang, "Hvac System Design Under Peak Load Prediction Uncertainty Using Multiple-Criterion Decisionmaking Technique," Energy And Buildings, Vol. 91, Pp. 26-36, 2015.

[12] K. T. Atanassov, Intuitionistic Fuzzy Sets. In Intuitionistic Fuzzy Sets, Physica, Heidelberg, 1999, Pp. 1-137.

[13] Z. H. Xu, "Intuitionistic Fuzzy Aggregation Operators," IEEE Transactions on Fuzzy Systems, Vol. 15(6), Pp. 1179-1187, 2007.

Determination of Production Defects in Iron and Steel Sector by Data Mining

İsmail Burak AKINCI, Filiz ERSÖZ

Industrial Engineering Department, Karabük University, TURKEY

Abstract: The studies related to the production industry are limited in the world and in our country. Especially in iron and steel sector, quality levels of different types of products need to be monitored. Iron and steel products obtained from the studies have prolonged their use and price and sales superiority has been achieved. At the same time, the market value of the products increases and there is a minimum loss of product. Therefore, studies in this field should be focused on. On the basis of quality, instead of debugging errors is the approach of not making mistakes. Instead of using your earnings as a philosophy, we should adopt an understanding of gaining from our losses. Understanding the importance of quality work and improvements, the primary purpose of enterprises is to support quality production by preventing or reducing errors in production. Data mining has started to be used effectively in enterprises. Data mining involves the process of selecting, organizing and modeling the most necessary data for business executives. At this point, it is possible to define data mining as a set of techniques and concepts that produce new information for decision-making processes. In this study, firstly the VM process is defined and then the VM studies which are selected from the literature covering 2010-2018 and applied to certain quality improvement problems in the manufacturing sector are evaluated. The definition of process and product quality, estimation of quality, classification of quality and optimization of quality parameters are discussed. In addition, the application of decision trees, one of the most widely used and effective VM techniques, in order to determine the variables and levels that cause production errors in an industrial organization is also included.

Keywords: Production, Manufacturing Defect, Data Mining

REFERENCES

[1] Liu, H.T., Tsai, Y.L. "A Fuzzy Risk Assessment Approach For Occupational Hazards in The Construction Industry", Safety Science. 50,1067-1078 (2012).

[2] Dünya Çelik Derneği [Online]. Erişim adresi: http://celik.org.tr/2017-yilinda-dunya-ham-celik-uretimi-5-3-oraninda-artti/ Erişim tarihi: 3 Kasım 2018.

[3] Türkiye Çelik Üreticileri Derneği [Online]. Erişim adresi: http://celik.org.tr/harita/ Erişim tarihi: 5 Kasım 2018.

[4] Güner, M., Akman, Ü., Yücel, Ö. "Improving The Man Shirt Production Processusing The Six Sigma Method", Tekstil ve Konfeksiyon, 20 (1), 75-82 (2010).

[5] Uyran, B. "Toplam Kalite Yönetimi", Mevzuat Dergisi, 55 (2002).

[6] Çevik, O., Aran, G. Kalite İyileştirme Sürecinde Hata Türü Etkileri Analizi (Fmea) Ve Piston Üretiminde Bir Uygulama. Sosyal Ekonomik Araştırmalar Dergisi, 8 (16), 241-265 (2009).

[7] Ersöz F., "Veri Madenciliği Teknikleri ve Uygulamaları" Ankara, (2015).

[8] Kayaalp, F., Başarslan, M., "Open Source Data Mining Programs: A Case Study on R", 6, 42-45 (2018).

[9] Savaş, S., Topaloğlu, N., Yılmaz, M., "Veri Madenciliği ve Türkiye'deki Uygulama Örnekleri" (2012).

[10] Sumathi, S., Sivanandam, S.N., "Introductionto Data Miningand Its Applications" (2006).

[11] Veri Madenciliği ve İşletmelerde Kullanım Alanları [Online]. Erişim adresi: https://dralabay.wordpress.com/2014/03/07/veri-madenciligi-ve-isletmelerde-uygulama-alanlari/Erişim tarihi: 5Kasım 2018.

[12] Veri Madenciliği Aşamaları [Online]. Erişim adresi:https://mevlutcanvar.com.tr/veri-madenciligi-asamalari/309Erişim tarihi: 5Kasım 2018.

[13] Veri Madenciliği Yöntemleri [Online]. Erişim adresi:https://emraharslanbm.wordpress.com/2012/07/17/veri-madenciligi-yontemleri/Erişim tarihi: 5 Kasım 2018.

[14] Yaralıoğlu, K., "Dokuz Eylül Üniversitesi Veri Madenciliği Ders Notu".

[15] ISO Nedir [Online]. Erişim adresi: https://www.muhendisbeyinler.net/iso-nedir/ Erişim tarihi: 5Ocak 2019.

[16] ISO Standartları [Online]. Erişim adresi: https://www.iso.org/obp/ui/#iso:std:iso:9001:ed-5:v1:enErişim tarihi: 5Ocak 2019.

[17] ISO 9001 Genel Bilgi [Online]. Erişim adresi: https://www.iso.org/iso-9001-quality-management.html Erişim tarihi: 5Ocak 2019.

Connected Employee Platform and A Case Study in A Global Company

Batuhan Burak ERSÖZLÜ, Ufuk CEBECİ, Şahen TOKATLIOĞLU

Department of Industrial Engineering, Istanbul Technical University, TURKEY

Abstract: In the modern era, enterprises are facing a variety of difficulties because of today's emerging technologies. One of the numerous difficulties for a business is to fulfill its employees in order to adapt to the consistently changing business processes and to make progress and stay in competition. In order to build proficiency, viability, efficiency and occupation responsibility of employee, the business must fulfill the requirements of its employee by giving great working conditions. The target of this paper is to dissect the effect of workplace on employee work fulfillment. This paper may profit society by urging individuals to contribute more to their occupations and may help them in their daily work life. Consequently, it is fundamental for an association to support their employee to snap down for accomplishing the hierarchical objectives and goals. The investigation and the item are changing the advanced endeavor, expanding representative commitment over the whole workforce, including forefront, field, remote and outside laborers to enhance execution, efficiency and unwaveringness. Our stage is making ground-breaking employee encounters, where every single representative feels some portion of an option that is more noteworthy than themselves, are glad to be a piece of your image and effectively advance the positive characteristics of the association.

Keywords: Employee Platforms, Employee Satisfaction, Employee Experience, Case Study, Mobile Application, User Experience, Software, Connected Workspace, Human Resources, İnformation Technologies, İnnovation, Visualized Applications

REFERENCES

[1] Forrester Research: Build Mobile Apps That Drive Engagement by Jeffrey S. Hammond, Michael Facemire, and Rowan Curran with Christopher Mines and Eric Wheeler, February 23, 2015

[2] Forrester Research: Trends 2015: The Future Of Customer Service by Kate Leggett with Stephen Powers, Ian Jacobs, and Arelai Ephraim, March 2, 2015

[3] Predictions on the Future of the Employee Experience. (2019). Retrieved from https://www.hrtechnologist.com/articles/employeeengagement/5-predictions-on-the-future-of-the-employee-experience/

[4] The Connected Employee Experience. (2019). Retrieved from https:// www.capgemini.com/service/technology-operations/infrastructureservices/connected-employee-experience/

[5] Your own branded employee app for Internal Communications, Engagement and HR. (2019). Retrieved from https://staffbase.com/en/

[6] Trends Shaping the Future of Employee Experience in 2019. (2019). Retrieved from https://www.itagroup.com/insights/trends-shapingfuture-of-employee-experience

[7] Baca, B., & Cassidy, A. (1999, September). Intranet Development And Design That Works. In Proceedings Of The Human Factors And Ergonomics Society Annual Meeting (Vol. 43, No. 13, pp. 777-781). Sage CA: Los Angeles, CA: SAGE Publications.

[8] Bersin, J. Becoming irresistible: A New Model For Employee Engagement, Deloitte Review 16, January 26, 2015, Retrieved from https://www2.deloitte.com/insights/us/en/deloitte-review/issue-16/ employee-engagement-strategies.html .

[9] Bottazzo, V. (2005). Intranet: A Medium of Internal Communication and Training. Information Services & Use, 25(2), 77-85.

[10] Buniyamin, N., & Barber, K. D. (2004). The intranet: A Platform For Knowledge Management System Based On Knowledge Mapping. International Journal of Technology Management, 28(7-8), 729-746.

[11] Cole, C. (2019). BMC's 'Connected' Employee Experience Boots Digital Workplace. [online] DiversityQ. Available at: https://diversityq.com/ bmcs-connected-employee-experience-boots-digitalworkplace-1004791/ [Accessed 4 Feb. 2019].

[12] Damsgaard, J., & Scheepers, R. (2001). Harnessing Intranet Technology For Organisational Knowledge Creation. Australasian Journal of Information Systems, 9(1).

[13] Damsgaard, J., & Scheepers, R. (2001). Using Intranet Technology To Foster Organizational Knowledge Creation. ECIS 2001 Proceedings, 32.

[14] Dukes, E. (2019). The Business Case for Employee Experience Software. [online] Iofficecorp.com. Available at: https:// www.iofficecorp.com/blog/the-business-case-for-employeeexperience-software [Accessed 4 Feb. 2019].

[15] Grover, R. (2019). How to Boost Employee Engagement with Awesome Internal Communication. Retrieved from https:// insights.staffbase.com/blog/18-ideas-to-improve-internalcommunication-in-2018-and-boost-your-employee-engagement.

[16] Horton, R., Buck, T., Waterson, P. (2001). Explaining Intranet Use With The Technology Acceptance Model. Journal of Information Technology. 16, 237.

[17] Intranet Software That Connects Your Organization | Interact Software. (2019). Retrieved from https://www.interact-intranet.com

[18] Kaplan, R. S., & Norton, D. P. (1995). Putting the Balanced Scorecard. Performance measurement, management, and appraisal sourcebook, 66.

[19] Lockley, S. (2019). How Employee Experience Is the Way to a More Engaging Workplace. Retrieved from https://insights.staffbase.com/ blog/the-anatomy-of-an-employee-experience-refining-todaysworkplace

[20] MacCormick J.S., Dery K., Kolb D.G. (2012). Engaged or Just Connect? Smartphones And Employee Engagement. Organizational Dynamics, 41, 194.

[21] Mahapatra, R., & Lai, V. (1998). Intranet-based Training Facilities ERP System Implementation: A Case Study. AMCIS 1998 Proceedings, 366.

[22] Murgolo-Poore, M. E., Pitt, L. F., & Ewing, M. T. (2002). Intranet Effectiveness: a Public Relations Paper-And-Pencil Checklist. Public Relations Review, 28(1), 113-123.

[23] Mphidi, H., & Snyman, R. (2004). The Utilisation of An Intranet As A Knowledge Management Tool in Academic Libraries. The Electronic Library, 22(5), 393-400.

[24] The employee experience: Helping People Get Excited To Do Their Best At Work. (2019). Retrieved from https://www.pwc.com/us/en/library/ workforce-of-the-future/employee-experience.html

[25] Plaskoff, J. (2017). Employee Experience: The New Human Resource Management Approach. Strategic HR Review,16 (3), 136-141.

[26] Scott, J.E. (1998). Organizational Knowledge And The Intranet. Decision Support Systems, 23(1), 3–17.

[27] Scarbrough, H. (1999) Feature Knowledge Management: System Error. People Management, April, 68–74.

[28] The Connected Employee Experience. (February, 2014). Retrieved from https://www.pwc.com/mt/en/publications/assets/hrs-pwctechnology-connected-employee-experience.pdf

[29] Sollenberger, D., Darpel, D., Talevski, S., Shumaker, L., Meyer, K., Journey, J., ... & Gimbert, N. (2002). U.S. Patent Application No. 09/833,433.

[30] Ruppel, C. P., & Harrington, S. J. (2001). Sharing Knowledge Through İntranets: A Study of Organizational Culture And İntranet İmplementation. IEEE transactions on professional communication, 44(1), 37-52.

[31] Urbach, N., Smolnik, S., & Riempp, G. (2009). A Conceptual Model For Measuring The Effectiveness of Employee Portals. AMCIS 2009 Proceedings, 589.

[32] Urbach, N., Smolnik, S., & Riempp, G. (2010). Improving the Success of Employee Portals: A Causal and Performance-Based Analysis. In ECIS (p. 54).

[33] Tojib, D. R., Sugianto, L. F., & Sendjaya, S. (2008). User Satisfaction With Business-To-Employee Portals: Conceptualization And Scale Development. European Journal of Information Systems, 17(6), 649-667.

[34] What is iBeacon? A Guide to iBeacons. (2019). Retrieved from http:// www.ibeacon.com/what-is-ibeacon-a-guide-to-beacons/

[35] Bluetooth | Definition of Bluetooth in English by Oxford Dictionaries. (2019). Retrieved from https://en.oxforddictionaries.com/definition/ bluetooth

[36] Android (operating system). (2019). Retrieved from https:// en.wikipedia.org/wiki/Android_(operating_system)

[37] IOS. (2019). Retrieved from https://en.wikipedia.org/wiki/IOS 38. GPS (Global Positioning System) Definition. (2019). Retrieved from https://techterms.com/definition/gps

Investigation of the Effects of Normal Distribution or Nonnormal of Data on Machine Capability Analysis

Erkan IŞIĞIÇOK, Gözde TÜRK

Department of Econometrics, Uludağ University, TURKEY

All processes in the business world have two fundamental disease, including deviation and variability from the average (target). One of the statistical process control graphs used for quantitative variables is to keep the average and the other to control variability. Apart from the normal distribution or nonnormal of quantitative data, the average and variability are controlled or not, and then the capability of process or machine is checked. The desired outcome is in addition to the normal distribution of data, the process is under control and capable. On the other hand, capability analysis is defined as the machine capability analysis when it is performed for the machine, while the process capability analysis takes its name when it is done for the process. In this study, machine capability analysis has been applied.

The aim of this article is to investigate the effects of the normal distribution or nonnormal of data on machine capability analysis. For this purpose, data on the lengths measured by surface of the shock absorber body pipe cut by a CNC machine in a company in the automative industry were used. In the study, 50 observations values were used, and the lower specification limit was 124.5 and upper specification limit was 125.5, and the CNC (pipe cutting) machine was sufficient or not.

The analysis first started with the implementation of the normality test and the data was not distributed normally. It is concluded that assuming this data, which does not have normal distribution, is normally distributed, and the machine is under control and is also sufficient with the I-MR control charts in the Minitab program. The same analysis was applied with the nonnormal command under the assumption that the data was nor normally distributed, and even in this case the machine was sufficient. In addition, the data that does not have normal distribution has been transformed into normality, and I-MR control charts and machine are under control and also sufficient. According to the findings, the average and specification limits of the values of I-MR control charts are the same and the machine capability results are different. In this study, these

similarities and differences were examined comparatively. Let us add it right away; these findings are specific to the machine and cut pipes we take into consideration and should not be generalized.

Keywords: Normal Distribution, The Data Does Not Have A Normal Distribution, Control Charts for Measurements, Under Control, Machine Capability Analysis.

Portfolio Selection with the Possibilistic Mean – Variance Model: An Application on the Borsa Istanbul

Furkan GÖKTAŞ, Süleyman DÜNDAR

Department of Business Administration, Karabük University, TURKEY

Abstract: The possibilistic mean – variance (MV) model enables the practitioners to incorporate the expert knowledge and robust statistics into the portfolio selection. Hence, it is a considerable alternative in decision making under uncertainty. In this study, we will examine the possibilistic mean – variance model theoretically under the assumption that the possibility distributions of the asset returns are given with the triangular fuzzy numbers. Here, the triangular fuzzy numbers will be determined based on the box plots. According to this, the possibilistic mean depends on the data set's median, interquartile range and skew. Furthermore, the possibilistic variance depends only on the data set's interquartile range. Then, we will illustrate this model based on the weekly returns of ten sector indices in 2017. Moreover, we will compare the risk adjusted performance and profitability of the possibilistic MV model and Markowitz's traditional MV model where the trading and testing periods cover the complete year of 2017 and 2018 respectively.

Keywords: Portfolio Selection, Decision Making Under Uncertainty, Triangular Fuzzy Numbers, Box Plot, Robust Statistics, Linear Programming

Building Digital Assistant (ChatBot) with SAP Conversational Artificial Intelligence

Metehan KOCAOĞLU[1], Kağan ÖZDEMİR[1], Alper KARABULUT[1],

Rüştü Orkun KORKMAZ[1], Ufuk CEBECİ[2]

[1]Solvia Digital Solutions, [2]Department of Industrial Engineering,
Istanbul Technical University, TURKEY

Abstract: Human-Computer Speech is gaining momentum as a technique of computer interaction. A chatbot is a software, which can “chat" with a human user in natural language such as English. Conversational artificial intelligence technology (CAI) enables learners to engage in spoken conversations with the non-player characters. In this study, SAP's digital assistant can be used to run business processes using SAP Conversational Artificial Intelligence. The benefits of the digital assistant (chatBot) will be discussed. An example of the request for personnel’s leave request, which is one of the areas of use, will be explained.

Keywords: Chatbot, Digital Asistant, SAP CAI, Conversational Artificial Intelligence, Leave Request.

Introduction

Nowadays, many innovations have started to be created and used with the digitalizing world. In addition, companies are trying to integrate innovations with their own systems in order to manage their processes quickly. It can be given 11 examples about chatbot. WeChat, one of them, was created by Chinese holding company Tencent in 2013. The product was created by a special projects team within Tencent (who also owns the dominant desktop messaging software in China, QQ) under the mandate of creating a completely new mobile-first messaging experience for the Chinese market [3]. In this study, digital assistant (ChatBot) application which can be created in SAP by using artificial intelligence was performed. The aim of ChatBot is to be a digital assistant. Today, time saving is a must-have. Humanity is looking for ways to make standardized tasks simpler, faster and more useful throughout the day. Many new applications are emerging new

products with this quest. SAP CAI is an innovative and self-enhancing bot creation platform. SAP CAI provides some significant gains [1].

- Saving Time

One of the biggest benefits of using chatBot in your business is to save time. When used on the website, it can provide quick and automatic answers to most questions and prevents customers from waiting for a day or longer to get a response as in the past. This allows an enterprise to serve more people while increasing productivity and reducing costs.

- Saving Money

The use of Chatbot may be cheaper than employing more workers. Thanks to algorithms to be created for frequently asked questions, you can meet the needs of your staff through chatBot without having to find a person.

- Customer Satisfaction

Another benefit of using the conversations in your business is that they provide more customer satisfaction. Chat robots do not work for 8 hours and do not need sleep, this means they are always available. Customers using the website in the evening can ask questions and get immediate answers. This could increase your profits. Frustrated customers who can't get a quick reply may not leave your website and come back. Chat boots can eliminate this scenario.

- Increasing Customer Base

Another benefit of using the conversations in your business is that they can help you reach more people to increase your customer base. Since chatBots can be used in many applications, it can be used to help the business grow.

- Reducing Errors

Unfortunately, people who handle customer service questions can make mistakes in some cases. He could really forget the information he knew. ChatBot will always give you the right answers based on the questions asked.

SAP CAI

Sign in at https://cai.tools.sap/ “Fig. 1”.

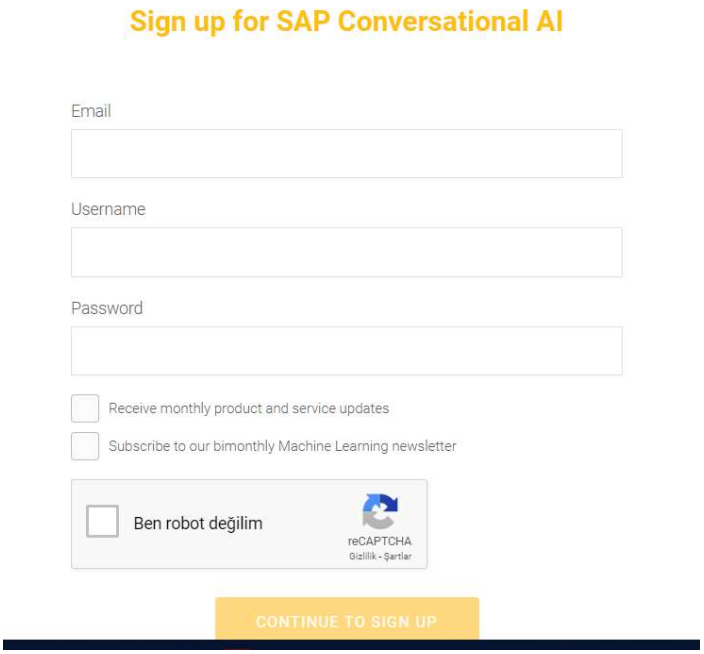


Fig. 1. SAP CAI Registration Screen

After registering, the bots that were previously created are displayed in the next screen. A new chatBot project is started with the “New Bot” button. On the screen, there are templates with ready answers against standard sentences available to users such as “Hello”, “How are you”. These templates can also be imported into the system [2].

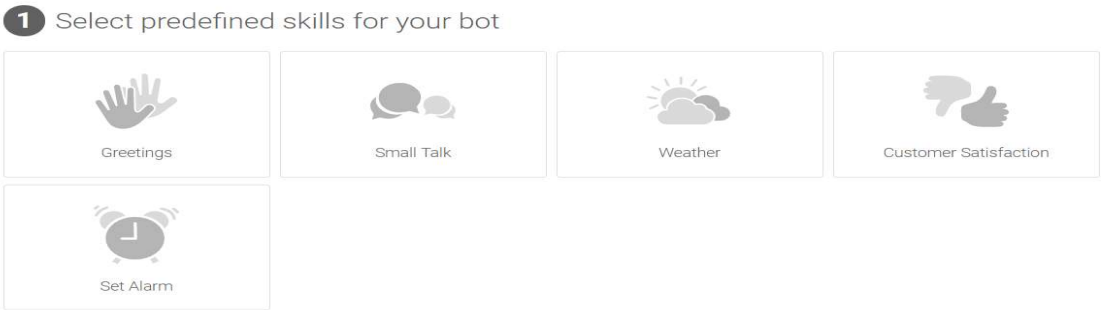


Fig 2. SAP CAI Register Screen

One of the language options as in Figure 3 is selected.

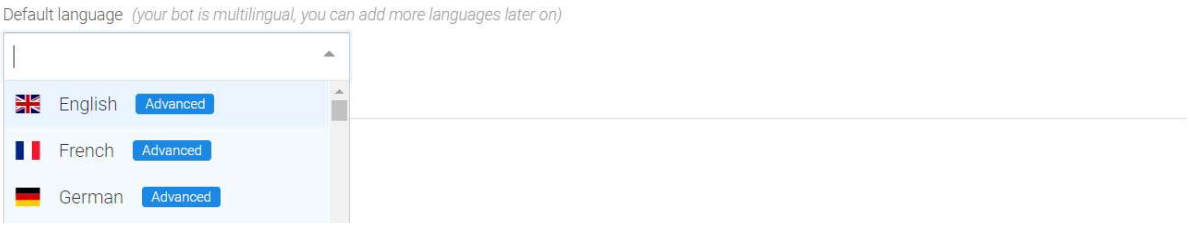


Fig 3. SAP CAI Language Options

A new bot is created with “CREATE A BOT” button.

Intents

An intent is a box of expressions that mean the same thing but are constructed in different ways. Intents are the heart of your bot’s understanding. Each one of your intents represents an idea your bot is able to understand “Fig. 4”.

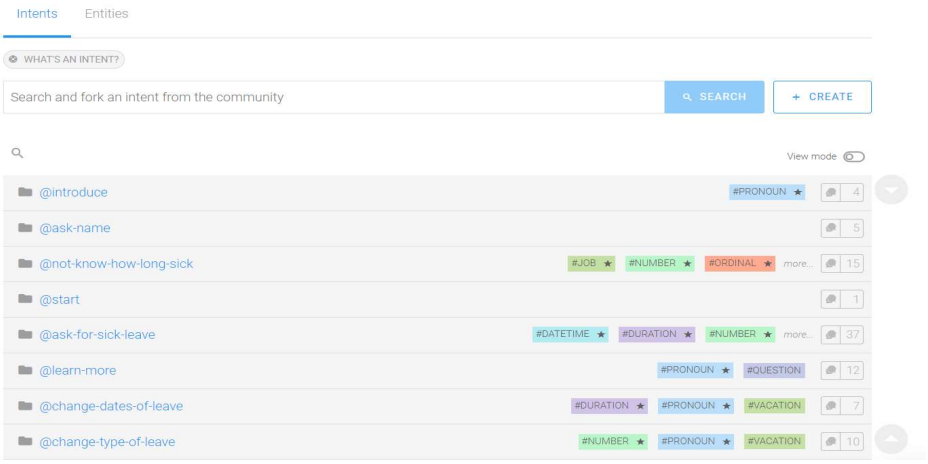


Fig. 4. Intents

You want your bot to understand when someone asks for help. Just create an intent called “help” and fill it with every expression a user would say when asking for guidance.”Fig. 5”:

- Could you help me?
- I’m lost, give me a hand please.
- Can you help?
- What can you do for me?

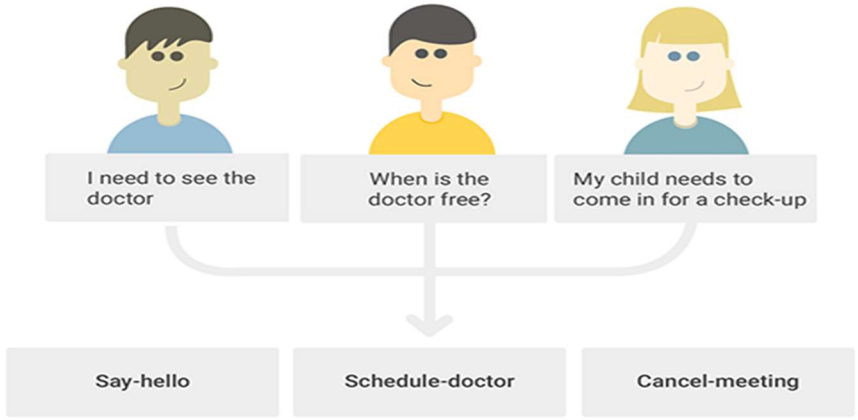


Fig. 5. Conversation

As in "Figure 6", the intent, named "ask-leave", which specifies the permission request, is created. To use the chat screen, the message written by the staff is trained.

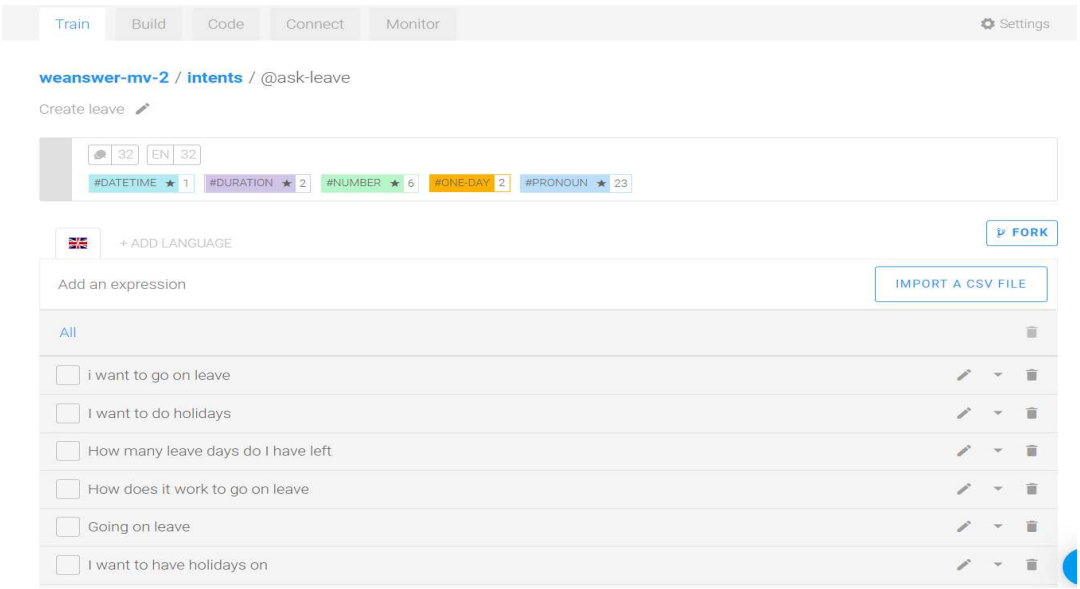


Fig. 6 Training Ask-Leave Intent

Entity

An entity is a keyword that is extracted from an expression. SAP CAI automatically detect 28 different entities such as Datetime, Location, Person, and so on. They are called gold entities.

However, there is no limit. It can also be tagged own custom entities to detect keywords depending on bot's context, such as subway stations if it is being built a transport assistant.

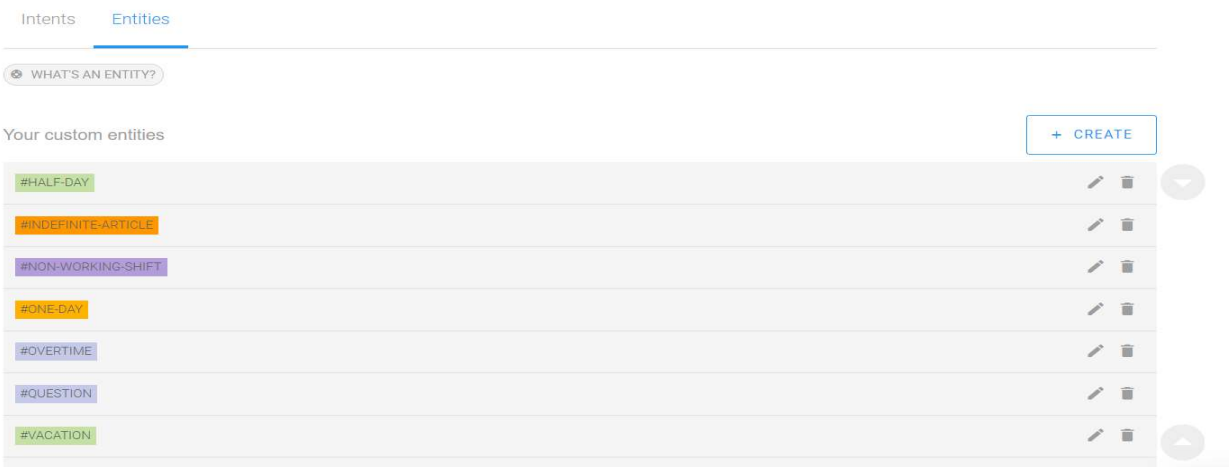


Fig. 7. Entity List

"In Fig. 8", if the request for a permit is selected, a half-day entity is created to be asked for a half-day or a full-time option.

Fig 8. Training #half-day entity

Condition, Api's

A condition is a test that can be evaluated to either true or false, "Fig. 9".

There's a finite list of operators can be used:

- Is to test equality between two values,

- Is-not to test inequality between two values,
- In to check if a value is in a list of elements,
- Not-in to check if a value is not in a list of elements,
- Matches to match a value with a regular expression,
- Matches-not to check if the value doesn't match with a regular expression,
- Lower-than to test if the value is lower than another,
- Greater-than to test if the value is greater than another,
- Is-present to test if the value is present in the conversation state,
- Is-asbent to test if the value is asbent in the conversation state.

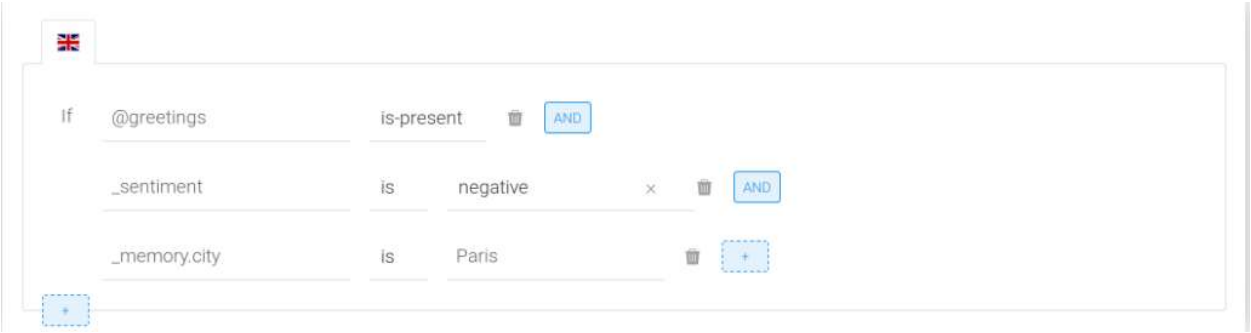


Fig. 9. Usage of Condition

At many points in conversation, you most likely want to retrieve business information or connect to an external system to perform actions. You can do this through webhooks. A webhook is a simple HTTP call to your backend. To configure your HTTP call, click CALL WEBHOOK in the Bot Builder "Fig. 10".

You can provide the full URL or route (starting with a '/') to be called by the Bot. Builder. If you provide a route, the Bot webhook base URL (configurable in your bot's settings) will be prepended to it.

You can specify the HTTP method to use in your webhook call (GET, POST, PUT, or PATCH).

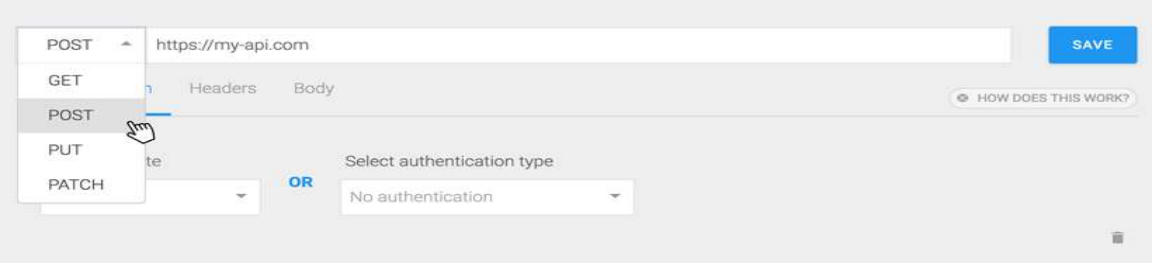


Fig 10. Api

Test Bot

The generated chatbot can be tested from the bottom right window ”Fig. 11”.

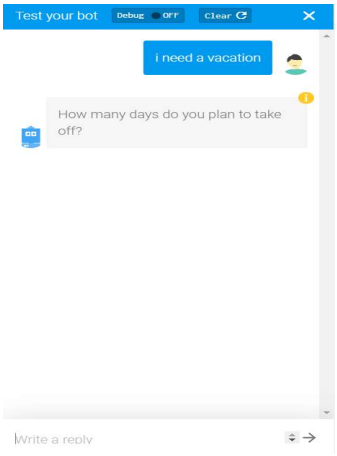


Fig 11. Testing ChatBot

Building Bot

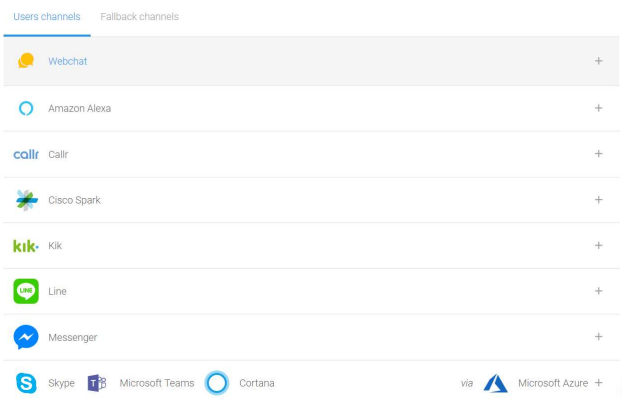


Fig 12. Bot Channels

The Webchat channel is developed by the SAP Conversational AI team and is an open-source project on GitHub “Fig. 12”.

You can use the default version of the webchat that we provide in the platform or customize the open-source version by forking it and deploying it on your side.

Fig. 13. chatBot Script

Webchat channel give a script and when it is placed at html page, it can be used “Fig. 13”.

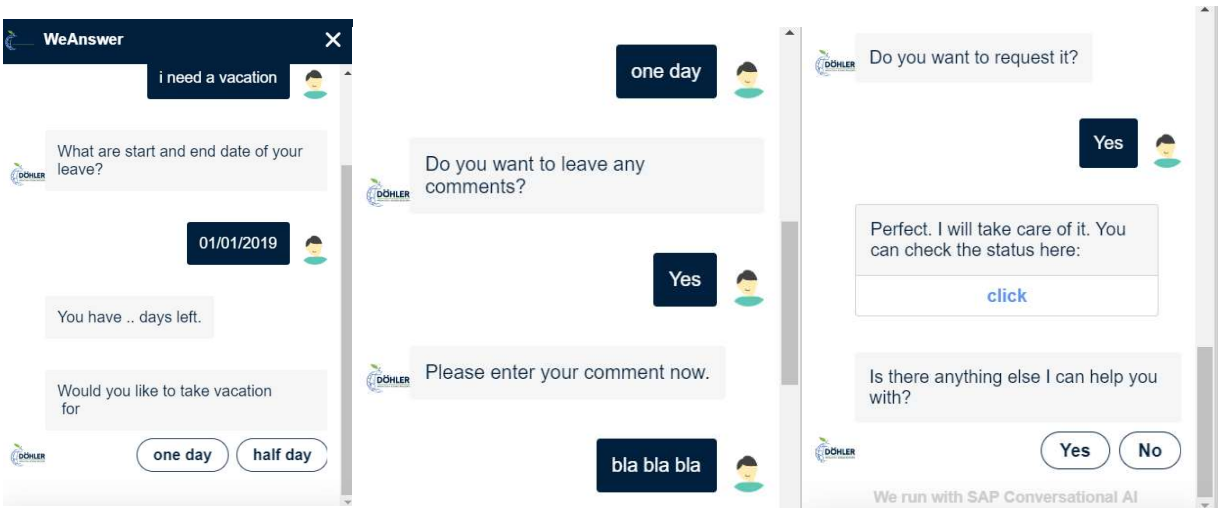


Fig. 14. ChatBot Script

Conclusion

SAP CAI product is described. The digital assistant was made. The permission request of the personnel has been processed “Fig. 14”.

REFERENCES

[1] Kayla Sloan, 6 Benefits of Using Chatbots in Your Business, May 9, 2018
Https://Due.Com/Blog/Chatbots-Business-Benefits

[2] Create your chatbot. Https://Cai.Tools.Sap/Docs/Concepts/Create-Builder-Bot

[3] Michael Quoc, 11 Examples of Conversational Commerce and Chatbots, Jun 1, 2016
Https://Chatbotsmagazine.Com/11-Examples-Of-Conversational-Commerce-57bb8783d332

[4] Matt Mayer, Building a WeChat (Weixin) robot, March 22, 2014
Https://Blog.Reigndesign.Com/Blog/Building-A-Wechat-Weixin-Robot/

Map Ranking, Map-Reduce and Application in Big Data Analysis

Safiye TURGAY, Suat ERDOĞAN

Industrial Engineering Department, Sakarya University, TURKEY

Abstract: The map method works with a certain algorithm and the inputs which send to a value list as a parameter. All values in the list converted to the intermediate result list. We sort all of the data then obtain the map list. The proposed and developed map structure was tested with quick-sort approach. The sorting process depends on the byte situation of the each data. Small data can do it easily side by side. Thus, small data do not need to applying of the reduce process. Sample selection havbe to will be easier. The goal is to give an intermediate operationtothe map-reduce structure. The more accurate is to get a ranking. In large data analysis, the data mapping sequence and reduction works with a certain algorithm structure, introducing and sending inputs as a parameter to a value list. An intermediate result list is created by converting all the values in the list, which are included in the entered system. In the structure developed after the mapping (Map) process, the mapping list is divided and obtained. The order depends on the byte value that is generated by each data. In case of large volume data, the data will be used without using a single line operation. In short, the data may be side-by-side, so there is no need to apply the reduction to each data. Therefore, the process of selecting the sample will be easier.

Keywords: Big Data, Map Reduce, Sorting, Styling.

REFERENCES

[1] Schönberger, Viktor Mayer ve Kenneth Cukier. Büyük Veri - Yaşama, Çalışma ve Düşünme Şeklimizi Dönüştürecek Bir Devrim. Çev. Banu Erol. İstanbul: Paloma, 2013.

[2] Demchenko Y, de Laat C, Membrey P (2014) Defining Architecture Components Of the Big Data Ecosystem. In International Conference on Collaboration Technologies and Systems (CTS) pp. 104-112.

[3] Rabie A. Ramadan, Big Data Tools-An Overview, International Journal of Computer & Software Engineering Volume 2 (2017), Article ID 2: IJCSE-125, 15 pages

[4] http://www.buyukveri.co/hadoop-hdfs-mimarisi/?

[5] https://hadoop.apache.org/docs/r1.2.1/mapred_tutorial.html

[6] MapReduce örnek – datascience.istanbulwww.datascience.istanbul

[7] https://dzone.com/articles/word-count-hello-word-program-inmapreduce

[8] Map ReducewithExamplesdatascienceguide.github.io

[9] http://www.dreamsyssoft.com/java-8-lambda-tutorial/map-reducetutorial.php

[10] java -Hadoop map-reduceoperation is failing on writingoutput - Stack Overflowstackoverflow.com

[11] Source codefororg.apache.hadoop.mapred.MapRunner.java - zGrepCodezgrepcode.com

[12] howto set configurationsfor Map-reducejobsforHadoop? - Stack Overflowstackoverflow.com

[13] Introductionto MapReducefor .NET Developers – DeveloperZendeveloperzen.com

[14] MapReduce Nedir? | B3LAB PORTALwww.b3lab.org

[15] Introductiontothe MapReduce Life Cycle | SUPINFO, ÉcoleSupérieure d'Informatiquewww.supinfo.com

[16] How to: Launching MapReduceJobs | MapRmapr.com

[17] http://salsahpc.indiana.edu/tutorial/

[18] Word Count Program With MapReduceandJava -DZoneBig Datadzone.com

[19] http://www.idearastirma.com/icerik/36/orneklemenin-onemi.aspx

[20] http://dresneradvisory.com/

[21] https://www.guru99.com/create-your-first-hadoop-program.html

[22] ApacheHadoop- AzureHDInsight için Java MapReduce oluşturma | Microsoft Docsdocs.microsoft.com

[23] MapReduceTutorial | MapreduceExample in ApacheHadoop | Edurekawww.edureka.co

[24] https://www.kadinyazilimci.com/siralama-algoritmalari-ii/

Benchmarking of OECD Countries in Views of Value-Added Manufacturing Using DEA

Billur ECER, Ahmet AKTAŞ, Mehmet KABAK

Industrial Engineering Department, Gazi University, TURKEY

Abstract: Value-added is an important term which indicates the efficiency level of economic activities. Valueadded term describes the difference between the value of produced goods and the total cost of production. Some areas related to value-added manufacturing are pharmaceuticals, automotive, computer and communications equipment manufacturing, aviation, etc. To develop a value-added product some inputs are needed. Some of these inputs are energy, labor force and research and development activities. The main aim of this study is to present a benchmarking analysis of 29 OECD countries in views of value-added manufacturing values. To do so, an output oriented BCC model is used to evaluate efficiency of countries and obtained results of the analysis provide some improvement ways for countries and their position against the other countries.

Keywords: OECD Countries, Benchmarking, Value-Added Manufacturing, DEA

Introduction

Value-added term is important for manufacturing operations, because the main aim of manufacturing is to make profit. Value-added describes the difference between the selling price of the product and the total cost to produce it. The leading value-added manufacturing areas can be listed as, pharmaceuticals, organic and inorganic chemical manufacturing, plastics manufacturing, semiconductors, computer manufacturing, communications equipment manufacturing, surgical and medical instruments manufacturing, automotive parts and aviation parts.

Almost all of the areas related with value-added manufacturing are hi-tech products. In this regard, research and development (R&D) activities are correlated with value-added manufacturing. There is no debate that research and development activities are conducted by manufacturers to discover some improvement strategies for their processes to develop their products or reduce costs.

Moreover, manufacturing is an activity that requires labour and energy inputs. The usage of greater amount of energy per capita and the greater ratio of employment in industry in a country can give us an idea for understanding the importance of manufacturing and industry in a country.

The main aim of this study is to present a benchmarking analysis of value-added manufacturing values of 29 OECD countries by considering R&D activities, industrial employment rates and energy usage. To perform this analysis Data Envelopment Analysis (DEA) technique is used. Output-oriented BCC (Banker, Charnes, Cooper) model is chosen for the analysis and obtained results of the model are discussed.

The rest of the paper organized as follows: a literature review for recent studies on benchmarking of countries by using DEA models is presented in the second part. Next, basic definitions for DEA and BCC model are given in the third part. The fourth part contains the information and the results about the benchmarking application of countries. Finally, the paper is concluded in the fifth part by presenting conclusions and future research directions.

Recent Literature of Benchmarking of Countries

Benchmarking of countries is a popular research area among researchers. In this part, a summary of 12 studies published between 2016 and 2019 is presented. DEA models are utilized in all of these studies. Details of these studies are presented as follows:

Chen and Hung analyzed efficiency of R&D activities in 25 countries by using network DEA model [1]. They develop a three-stage efficiency model consisting research, translation and economic efficiency stages.

Timmer et al. [2] developed a decomposition framework based on DEA to analyze labor efficiency difference between Germany and USA in the early 20th century. Results of their analyses lead them to obtain some findings to understand the drivers behind the productivity differences.

Environmental performance efficiency of European countries is evaluated in Chodakowska and Nazarko's study [3]. By using DEA models, environmental performance and technological competition indicators are integrated. They found out the diversification of environmental performance of European countries.

Efficiency of CO_2 emission reduction techniques applied on 12 European countries are compared by a two-stage DEA model [4]. Kwon et al.'s analysis shows benchmarks for inefficient countries to improve their CO_2 reduction strategies.

Marti et al. [5] developed a DEA based logistics performance index (LPI) approach to benchmark countries in terms of logistics performance. They analyzed a number of scenarios by considering different combinations of factors that affecting LPI.

Storto and Goncharuk's study presents a benchmarking analysis of European countries in views of efficiency and effectiveness of healthcare services [6]. They used slack based measure model of DEA to identify shortcomings of national healthcare systems.

Faghih et al. [7] used DEA models to determine benchmarks of 55 MENA countries in views of national entrepreneurial efficiency. The data used in the analysis is gathered from Global Entrepreneurship Monitor data.

Road safety policies across Europe is evaluated by Nikolaou and Dimitriou [8]. They used data of 2005 – 2014 for 23 EU countries and discover some managerial implications from the analysis via DEA.

Output oriented BCC model is used in See and Yen's study to analyze efficiency of health services [9]. Data of 121 countries are evaluated in the analysis and benchmarks are presented for policy makers.

Network DEA model is used in Wanke et al.'s study to investigate drivers of railway performance in selected Asian countries [10]. Different improvement suggestions for countries are provided as a result of the analysis.

Efficiency of government excellence in 45 low and low-middle income countries are evaluated by Choi and Park [11]. Effect of government excellence on social progress is investigated in the study and countries with same economic conditions are benchmarked with each other.

Logistic performance evaluation of OECD countries is considered in Rashidi and Cullinane's study [12]. Top performers and inefficient countries are presented with benchmarking countries at the end of the analysis.

It is seen that different DEA models are used to compare efficiency of countries. Therefore, DEA is an appropriate method for our analysis.

Methodology

Efficiency of a process is defined as the ratio of outputs to the inputs. It describes the degree of transformation of inputs to the outputs. The higher value that efficiency takes, the more efficient the process is. Efficiency of process with more than one input and/or output is calculated by considering importance degree of the inputs and outputs. DEA is a technique, which is used for measuring the relative efficiency of a number of systems with the same inputs and outputs. DEA models try to determine the weights of inputs and outputs for determining the best efficiency score of each decision-making unit (DMU).

There are two different problem assumptions for measuring efficiency with DEA models. The system that is under consideration can aim the same output level with lower input (input-oriented) or obtain more output with the same input level (output-oriented).

Moreover, systems may follow constant return to scale (CRS) or variable return to scale (VRS) assumptions. CRS is about existence of a constant efficiency frontier, while VRS accepts efficiency frontier may change related to the input level. Based on the scale assumption, there are different models named as CCR (Charnes, Cooper, Rhodes) and BCC (Banker, Charnes, Cooper).

In this study, value added manufacturing efficiency of countries is measured by using DEA model. It is thought that VRS assumption is appropriate for this problem, because of the different development levels of countries. Also, a DMU with higher output is more efficient than a DMU which has the same level of input. So, the DEA model for this study is determined as output-oriented BCC model. The mathematical formulation of this model is given as follows [13]:

$$\max z_k$$

$$-\sum_{j=1}^{n} Y_{rj}\lambda_{jk} + Y_{rk}z_k \le 0 \qquad \forall r$$

$$\sum_{j=1}^{n} x_{ij}\lambda_{jk} \le X_{ik} \qquad \forall i$$

$$\sum_{j=1}^{n} \lambda_{jk} = 1$$

$$\lambda_{jk} \ge 0 \qquad \forall j$$

Benchmarking Analysis of OECD Countries

In this study, value added manufacturing efficiency of 29 OECD member countries are measured by using a DEA model. BCC model is used to measure efficiency of countries. The output of this efficiency analysis is determined as percent ratio of value-added manufacturing in the country's GDP (MVA). On the other hand, percentage of employment in industry of total employment (EiI), energy use per capita (kg of oil equivalent) (EU) and research and development expenditure percentage of GDP (RD) are considered as inputs of value-added manufacturing. Data of 29 OECD member countries are retrieved from World Bank's database of World Development Indicators [14] for 2015. These data are the most recent data of countries considered in the analysis. The other member countries' data are not available, that is why they are out of consideration. Countries that taken into consideration in this study are Austria, Belgium, Chile, Czech Republic, Denmark, Estonia, Finland, France, Germany, Greece, Hungary, Ireland, Israel, Italy, Japan, Korea Republic, Luxembourg, Mexico, Netherlands, Norway, Poland, Portugal, Slovak Republic, Slovenia, Spain, Sweden, Turkey, United States and United Kingdom. Data of the efficiency analysis are given in Table 1 as follows:

TABLE 1. DATA FOR THE ANALYSIS

Country	Indicator			
	MVA	***EiI***	***EU***	***RD***
AUT	16.56	25.76	3800.30	3.05
BEL	12.83	21.43	4687.79	2.47
CHL	11.56	23.28	2028.89	0.38
CZE	24.13	38.03	3860.00	1.93
DNK	12.41	19.29	2816.61	2.96
EST	13.83	30.66	4173.33	1.49
FIN	14.80	21.70	5924.70	2.90
FRA	10.43	20.38	3689.52	2.27
DEU	20.76	27.69	3817.55	2.92
GRC	8.31	14.94	2182.07	0.97
HUN	20.52	30.33	2432.75	1.36

Country	Indicator			
	MVA	***EiI***	***EU***	***RD***
IRL	34.32	19.08	2819.88	1.20
ISR	11.78	17.71	2777.88	4.27
ITA	14.39	26.60	2481.75	1.34
JPN	20.72	25.91	3428.56	3.29
KOR	27.09	25.07	5413.35	4.22
LUX	4.86	12.52	6548.41	1.27
MEX	17.14	25.06	1488.02	0.53
NLD	10.70	16.43	4233.04	2.00
NOR	6.86	20.12	5815.81	1.93
POL	17.64	30.54	2490.21	1.00
PRT	12.16	24.35	2131.68	1.24
SVK	19.75	36.11	3003.66	1.18
SVN	19.96	32.01	3174.87	2.20
ESP	12.90	19.90	2571.34	1.22
SWE	13.69	18.29	5102.79	3.27
TUR	16.71	27.23	1656.80	0.88
USA	11.92	18.85	6797.62	2.74
GBR	8.97	18.66	2763.98	1.67

DEAP software is used measure efficiency of countries. Efficiency analysis results of output-oriented BCC model is given in Table 2.

According to efficiency analysis results in Table 2, Ireland and Mexico are the efficient countries under constant return to scale assumption among 29 countries. Efficient countries for variable return to scale assumption are determined as Chile, Greece, Ireland, Luxemburg and Mexico. As it mentioned before, under VRS assumption, efficiency frontier changes for different input levels and number of efficient DMUs are generally greater than number of efficient DMUs in CCR models. Scale efficiency is the ratio of CCR efficiency to BCC efficiency and it shows the DMU has increasing return to scale or decreasing return to scale.

TABLE 2. EFFICIENCY VALUES

Country	EFFICIENCY SCORE		
	CCR	***BCC***	***SCALE***
AUT	0.358	0.483	0.742 (DRS)
BEL	0.333	0.374	0.890 (DRS)
CHL	0.941	1.000	0.941 (IRS)
CZE	0.514	0.703	0.731 (DRS)
DNK	0.362	0.362	1.000 -
EST	0.318	0.403	0.790 (DRS)
FIN	0.379	0.431	0.879 (DRS)
FRA	0.285	0.304	0.936 (DRS)
DEU	0.447	0.605	0.739 (DRS)
GRC	0.313	1.000	0.313 (IRS)
HUN	0.693	0.700	0.990 (IRS)
IRL	1.000	1.000	1.000 -
ISR	0.370	0.451	0.820 (IRS)
ITA	0.476	0.480	0.992 (IRS)
JPN	0.497	0.604	0.822 (DRS)
KOR	0.601	0.789	0.761 (DRS)
LUX	0.216	1.000	0.216 (IRS)
MEX	1.000	1.000	1.000 -
NLD	0.362	0.482	0.751 (IRS)
NOR	0.190	0.200	0.948 (DRS)
POL	0.593	0.604	0.981 (DRS)
PRT	0.469	0.478	0.981 (IRS)
SVK	0.554	0.584	0.949 (DRS)
SVN	0.517	0.582	0.888 (DRS)
ESP	0.412	0.425	0.970 (IRS)
SWE	0.416	0.445	0.935 (IRS)
TUR	0.829	0.865	0.958 (IRS)
USA	0.352	0.358	0.982 (IRS)

Country	EFFICIENCY SCORE		
	CCR	***BCC***	***SCALE***
GBR	0.267	0.283	0.945 (IRS)
MEAN	0.485	0.586	0.857

Another important result obtained with DEA is the reference sets and target values for inefficient DMUs. These values are presented in Table 3. For example, Turkey should be similar to Ireland and Mexico with weight values of 0.127 and 0.873. That means Turkey has to change their input and output values to sum of 0.127 of Ireland's scores and 0.873 of Mexico's scores. That change will bring Turkey's scores into an output value of 19.317 and input values of 24.302, 1656.800 and 0.615, respectively.

TABLE 3. REFERENCE SETS AND TARGET VALUES

Country	*REFERENCE SET*	Indicator			
		MVA	***EiI***	***EU***	***RD***
AUT	IRL (1.000)	34.320	19.080	2819.880	1.200
BEL	IRL (1.000)	34.320	19.080	2819.880	1.200
CHL	CHL (1.000)	11.560	23.280	2028.890	0.380
CZE	IRL (1.000)	34.320	19.080	2819.880	1.200
DNK	IRL (0.998), MEX (0.002)	34.278	19.095	2816.610	1.198
EST	IRL (1.000)	34.320	19.080	2819.880	1.200
FIN	IRL (1.000)	34.320	19.080	2819.880	1.200
FRA	IRL (1.000)	34.320	19.080	2819.880	1.200
DEU	IRL (1.000)	34.320	19.080	2819.880	1.200
GRC	GRC (1.000)	8.310	14.940	2182.070	0.970
HUN	IRL (0.709), MEX (0.291)	29.326	20.818	2432.750	1.005
IRL	IRL (1.000)	34.320	19.080	2819.880	1.200
ISR	GRE (0.274), IRL (0.690), LUX (0.036)	26.132	17.710	2777.880	1.139
ITA	IRL (0.746), MEX (0.254)	29.958	20.598	2481.750	1.030
JPN	IRL (1.000)	34.320	19.080	2819.880	1.200

Country	*REFERENCE SET*	Indicator			
		MVA	***EiI***	***EU***	***RD***
KOR	IRL (1.000)	34.320	19.080	2819.880	1.200
LUX	LUX (1.000)	4.860	12.520	6548.410	1.270
MEX	IRL (1.000)	17.140	25.060	1488.020	0.530
NLD	GRE (0.031), IRL (0.585), LUX (0.384)	22.189	16.430	4233.040	1.220
NOR	IRL (1.000)	34.320	19.080	2819.880	1.200
POL	IRL (0.701), MEX (0.299)	29.192	20.865	2422.310	1.000
PRT	IRL (0.483), MEX (0.517)	25.443	22.170	2131.680	0.854
SVK	IRL (0.970), MEX (0.030)	33.807	19.259	2780.123	1.180
SVN	IRL (1.000)	34.320	19.080	2819.880	1.200
ESP	GRE (0.042), IRL (0.791), MEX (0.166)	30.363	19.900	2571.340	1.079
SWE	IRL (0.880), LUX (0.120)	30.772	18.290	3268.895	1.208
TUR	IRL (0.127), MEX (0.873)	19.317	24.302	1656.800	0.615
USA	IRL (0.965), LUX (0.035)	33.287	18.850	2950.606	1.202
GBR	GRE (0.099), IRL (0.900), LUX (0.002)	31.703	18.660	2763.980	1.177

Conclusion

Value added manufacturing percentage of GDP is one of the important indicators for competitiveness in global trade competition between countries. In this study, efficiency of value-added manufacturing values of 29 OECD member countries are measured by using output-oriented BCC model. Mean efficiency of 29 countries is measured as 0.586 and efficient countries are found as Chile, Greece, Ireland, Luxemburg and Mexico. Benchmarks for inefficient countries are presented and target values are calculated by DEAP software.

In further studies, this study can be extended by considering different indicators in the model. Moreover, Malmquist total factor productivity index can be calculated by considering different years' data to analyze changes of efficiency in a multi-period view.

REFERENCES

[1] P. C. Chen And S. W. Hung, "An Actor-Network Perspective on Evaluating The R&D Linking Efficiency Of İnnovation Ecosystems," Technological Forecasting & Social Change, 112, 303–312, 2016.

[2] M. P. Timmer, J. Veenstra And P. J. Woltjer, "The Yankees of Europe? A New View On Technology and Productivity in German Manufacturing İn The Early Twentieth Century," The Journal Of Economic History, 76, 874-908, 2016.

[3] E. Chodakowska and J. Nazarko, "Environmental Dea Method For Assessing Productivity Of European Countries," Technological And Economic Development of Economy, 23, 589–607, 2017.

[4] D. S. Kwon, J. H. Cho And S. Y. Sohn, "Comparison of Technology Efficiency For Co2 Emissions Reduction Among European Countries Based on Dea With Decomposed Factors," Journal Of Cleaner Production, 151, 109-120, 2017.

[5] L. Martí, J. C. Martín And R. Puertas, "A Dea-Logistics Performance Index," Journal of Applied Economics, 20, 169-192, 2017.

[6] C. L. Storto And A. G. Goncharuk, "Efficiency Vs Effectiveness: A Benchmarking Study on European Healthcare Systems. Economics And Sociology, 10, 102-115, 2017.

[7] N. Faghih, M. R. Zali And N. Vafaei, "Entrepreneurial National Efficiency Based on Gem Data: Benchmarks For The Mena Countries" In: Faghih N., Zali M. (Eds) Entrepreneurship Ecosystem in The Middle East And North Africa (Mena). Contributions To Management Science. Springer, Cham.

[8] P. Nikolaou And L. Dimitriou, "Evaluation of Road Safety Policies Performance Across Europe: Results From Benchmark Analysis For A Decade," Transportation Research Part A, 116, 232–246, 2018.

[9] K. F. See And S. H. Yen, "Does Happiness Matter To Health System Efficiency? A Performance Analysis," Health Economics Review, 8, 33, 2018.

[10] P. Wanke, Z. Chen, W. Liu, J. J. M. Antunes And M. A. K. Azad, "Investigating The Drivers Of Railway Performance: Evidence From Selected Asian Countries," Habitat International, 80, 49–69, 2018.

[11] H. Choi And M. J. Park, "Evaluating The Efficiency Of Governmental Excellence For Social Progress: Focusing on Low And Lower-Middle-İncome Countries," Social Indicators Research, 141, 111–130, 2019.

[12] K. Rashidi And K. Cullinane, "Evaluating The Sustainability of National Logistics Performance Using Data Envelopment Analysis," Transport Policy, 74, 35–46, 2019.

[13] A. Tarım, "Veri Zarflama Analizi: Matematiksel Tabanlı Göreli Etkinlik Ölçüm Yaklaşımı," Ankara: T.C. Sayıştay Başkanlığı, Araştırma Çeviri İnceleme Dizisi, 2001.

[14] Internet, "Http://Data.Worldbank.Org/Data-Catalog/World-Developmentindicators", Access Date: 20.04.2019

A Study on Method Prediction for a Better Directed Treatment of Warts

Ahmet ÇİFCİ[1], Mehmet ŞİMŞİR[2]

[1]Deparment of Electrical-Electronics Engineering, Burdur Mehmet Akif Ersoy University,
[2]Department of Mechatronics Engineering, Karabük University, TURKEY

Abstract: Various kinds of medical treatment methods may be used to cure common types of diseases. Experience based predictions are done to choose a treatment method among the choices of cures to get better results for the patient. This condition sometimes may continue with trying another treatment method unless a satisfactory result is reached and changing the treatment method of the cure process is not a desired course for time and health. This study presents a confident way to choose the treatment method for wart disease by using feedforward neural network. The study uses two types of datasets, one for cryotherapy and other for immunotherapy treatment methods. It was observed from the experimental results that, feedforward neural network achieved 94.4% success and 85.6% success for cryotherapy and immunotherapy datasets, respectively. The results are remarkable for both doctors and patients.

Keywords: Cryotherapy, Immunotherapy, Feedforward Neural Network, Wart.

Introduction

Warts are benign skin growth caused by the virus called Human papillomavirus (HPV), which infect the top layer of the skin [1]. Most warts are harmless, but they are highly contagious. Traumas events such as cuts or damages to the skin facilitate infection. They are spread so quickly, especially in the summer. Direct contact with a wart or contacts with another person by sharing towels, razors, or other personal items may cause the virus to spread [2]. Although they are most common on the knuckles, fingers, hands, elbows, knees, they may occur in the whole body. Children, adolescents, people who bite their nails, and people with a weak immune system have a higher risk of developing warts [3]. Warts are usually flesh-coloured, hard and rough. On the other hand, there are dark (brown, grey-black), flat and soft wart types.

Types of wart include common warts (*verruca vulgaris*), plantar warts (*verruca plantaris*), flat warts (*verruca plana*), and genital warts (*condylomata acuminata*) [4]. Common warts are most commonly seen in the hands, between the fingers, legs, and around the nails [5]. Plantar warts are embedded in the skin and can be painful. Plantar warts can be confused with calluses [6]. Even though flat warts can occur on any part of the body, they are often found on the face and hands. They are small, smooth and flesh-coloured. They can occur in large numbers [4]. Genital warts are soft, flesh-coloured papules on the genitalia and breech. It is more common in people who have sexual activity without a condom [4].

There exist several methods available for the treatment of warts. Salicylic acid treatment, electrosurgery, freezing (cryotherapy), immunotherapy, and laser treatment are the main methods to get rid of warts [7]. The most common methods for the treatment of warts are cryotherapy and immunotherapy. Cryotherapy is performed by applying liquid nitrogen gas to the desired skin lesion using special devices. Liquid nitrogen gas at a temperature of -196 degree Celsius will freeze the tissue. This freezing process is a short duration of time (10 - 60 seconds); at the end of this time, the tissue will return to its normal temperature. In this short-term and immediate freezing-melting process, the cells in the targeted tissue will be destroyed and will die. These dead and abnormal cells, which no longer function, will be removed from the tissue during the healing process and replaced by fresh tissues [8,9]. Immunotherapy is a new type of treatment that aims to prevent or eliminate the growth of cancer cells that abnormally proliferate (such as HPV lesions). Immunotherapy uses the patient's own immune system to fight warts. The antigen is injected into the body to activate the immune system [10,11]. Cryotherapy is an uncomplicated method for the patients with respect to immunotherapy.

Numerous investigations have been reported in the literature on the treatment of warts. In [12], Putra et al. used AdaBoost and Random Forest as a strong learner or a weak learner for selection of wart treatment method. The results showed that the accuracy rate of cryotherapy was 96.6% and accuracy rate of immunotherapy was 91.1%. Tanyıldızı et al. [13] judged the performances of classification algorithms in cryotherapy and immunotherapy datasets. The best success rate obtained using the K-star algorithm is found as 96.66% for cryotherapy and using the Random Forest algorithm is found as 85.55% for immunotherapy. Fuzzy Rule, Naive Bayes, and Random Forest based algorithms have been carried out in cryotherapy and immunotherapy treatment for comparing the effectiveness of these algorithms by Akyol et al. in [14]. They

concluded that the random forest algorithm outperforms other classification algorithms in both accuracy and sensitivity within cryotherapy and immunotherapy datasets. Nugroho et al. proposed C4.5 algorithm combined with Random Forest Feature Weighting for wart treatment selection method [15]. The results showed that the proposed method can improve the performance of prediction. Ali et al. applied some algorithms to show that which treatment is more effective from cryotherapy and immunotherapy [16]. They concluded that cryotherapy treatment is better than immunotherapy. See also [17], where the author used the decision tree-based method to specify the rules of predicting the performance of wart treatment methods. The results obtained the level of accuracy of 94.4% on cryotherapy and 90% on immunotherapy.

In this study, feed-forward neural network was used to decide and select the cure for wart treatment. Patient data were used as input and success of cure type was used as output data. A neural network for immunotherapy and another one for cryotherapy was built to decide which methodology is appropriate for the patient. As it was mentioned before, cryotherapy is an uncomplicated method for the patient with respect to immunotherapy. When a patient data is applied to neural network for cryotherapy and gives positive outputs for applicability. Treatment method starts with cryotherapy cure, but if cryotherapy neural network gives ineffective result as an output, applying data to immunotherapy neural network is preferred and this situation supports a logical way to select the uncomplicated method at first without waste of time. Because immunotherapy method has a higher result of success percentage according to the datasets used in this study. But this method may also be abrasive for the patient.

The paper is structured in the following manner. Section 2 presents the immunotherapy and the cryotherapy datasets. Method and experimental results are given in section 3. Finally, conclusions are duly drawn in section 4.

Dataset

The immunotherapy and the cryotherapy datasets used in this study are gathered from the University of California, Irvine (UCI) Machine Learning Repository [18,19]. The datasets are collected along two years from the dermatology clinic of Ghaem Hospital in Mashhad, Iran [20,21]. The immunotherapy with candida antigen and the cryotherapy with liquid nitrogen were applied to 180 patients with plantar and common warts. Each data set contains 90 patients. Patients were randomly selected. The datasets do not have any missing value.

The immunotherapy dataset consists of eight features; gender, age, time elapsed before treatment, the number of warts, types of wart, surface area of warts, induration diameter of initial test, and response to treatment. The details of the immunotherapy dataset are presented in Table 1.

Table 1. Features of Immunotherapy Dataset

Feature No.	**Feature Name**	**Values**
1	Gender	41 Male
		49 Female
2	Age (year)	15-56
3	Time elapsed before treatment (month)	0-12
4	The number of warts	1-19
5	Types of wart (count)	47 Common
		22 Plantar
		21 Both
6	Surface area of warts (mm^2)	6-900
7	Induration diameter of initial test (mm)	5-70
8	Response to treatment	Yes or No

The cryotherapy dataset has seven features; gender, age, time elapsed before treatment, the number of warts, types of wart, surface area of warts, and response to treatment. The details of the cryotherapy dataset can be seen in Table 2.

Table 2. Features of Cryotherapy Dataset

Feature No.	**Feature Name**	**Values**
1	Gender	47 Male
		43 Female
2	Age (year)	15-67
3	Time elapsed before treatment (month)	0-12
4	The number of warts	1-12

5	Types of wart (count)	54 Common
		9 Plantar
		27 Both
6	Surface area of warts (mm^2)	4-750
7	Response to treatment	Yes or No

Method and Experimental Results

With the developments on intelligent systems, human expertise can be simulated by artificial neural networks to gain time and to get more accurate decisions. As a general approach, our study is based on previously taken and reliable reference data. After a general overview on types of cures for wart treatment, it is necessary to classify the patient data. “Gender”, “age”, “time elapsed before treatment”, “the number of warts”, “types of wart” and “surface area of warts” features for the patients were used as the classification titles for the neural network data inputs. And output data was also used as success of cure type as “0” and “1”.

A feedforward neural network was chosen for this study. Feedforward neural networks has input, output and at least one hidden layers. The neural network was constructed with one hidden layer. If too many hidden layers and neurons are used, overfitting may occur, that is although the network can be trained to work very well for the training data, it performs poorly for test data. It is essential to optimize the success percentage of the neural network with effectual numbers of neurons for the hidden layer to achieve the best performance level [22,23].

The constructed neural networks have 6 input variables, 6 neurons for the hidden layer and also 2 for output for both cryotherapy and immunotherapy as seen in Fig. 1.

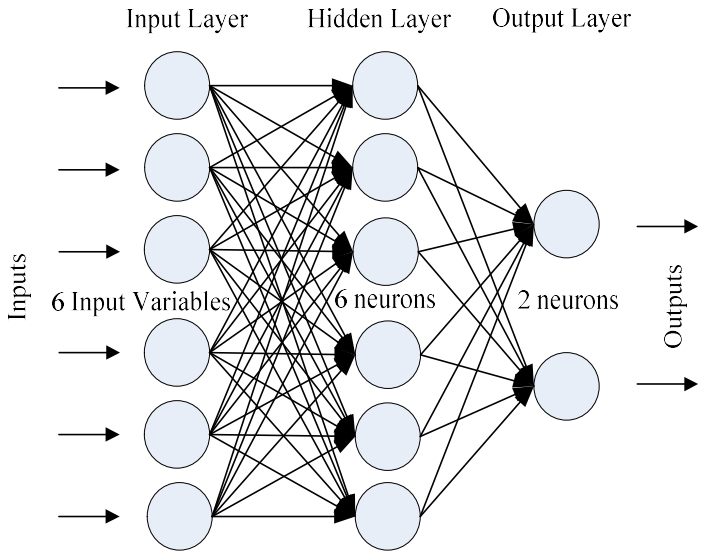


Fig. 1. General Structure Of Constructed Feed Forward Backpropagation Neural Networks For Both Cryotherapy And Immunotherapy.

To determine the best condition for a better success percentage of neural networks, several numbers of neuron numbers were experimented by MATLAB and current configuration shown in Fig. 1 was obtained because of better success rate according to the tested other choices. The training performance curves of the designed neural networks are shown in Fig. 2 for cryotherapy and Fig. 3 for immunotherapy. Best validation level of the neural networks for training process is indicated by means of training, validation and test curves. The training process of the neural network was reached to goal at 7 epochs in terms of minimum gradient for cryotherapy neural network.

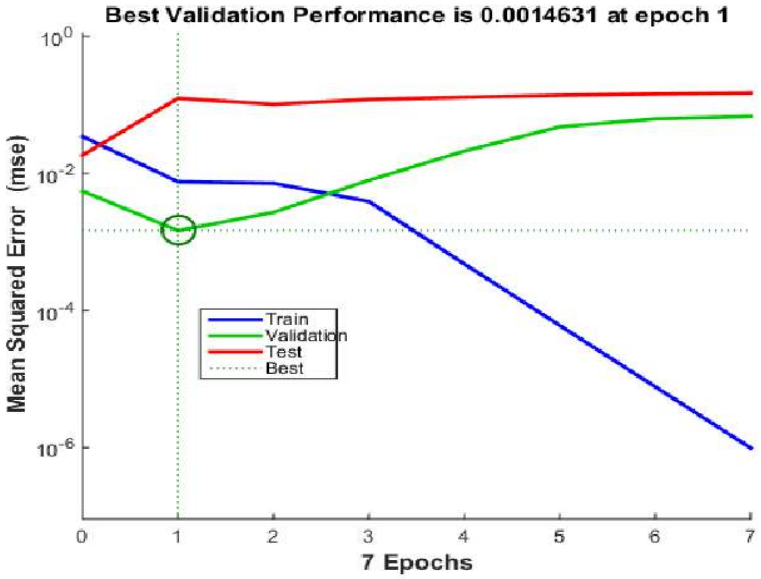


Fig. 2. Training Performance Curves of Neural Network Designed For Cryotherapy by MATLAB.

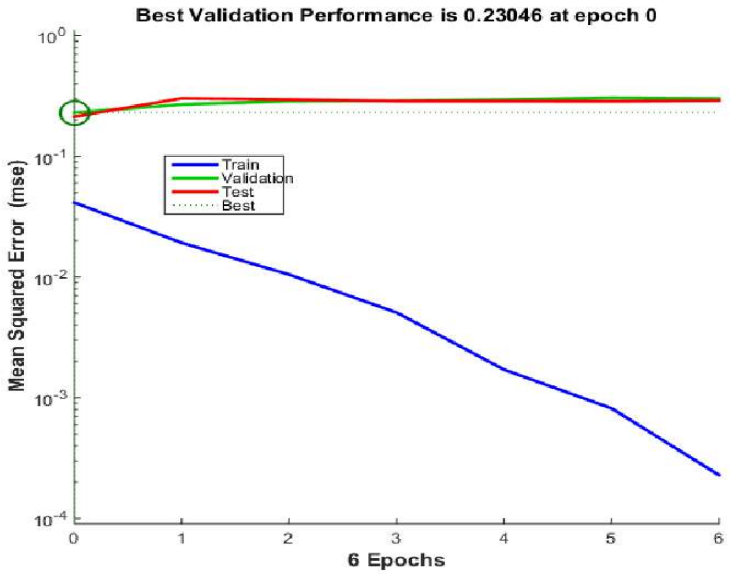


Fig. 3. Training Performance Curves of Neural Network Designed For Immunotherapy By MATLAB.

The training state throughputs of the neural network are shown in Fig. 4 for cryotherapy and Fig. 5 for immunotherapy, which indicate weight changes and validation checks for the neural networks and gradients till the determined goals were reached.

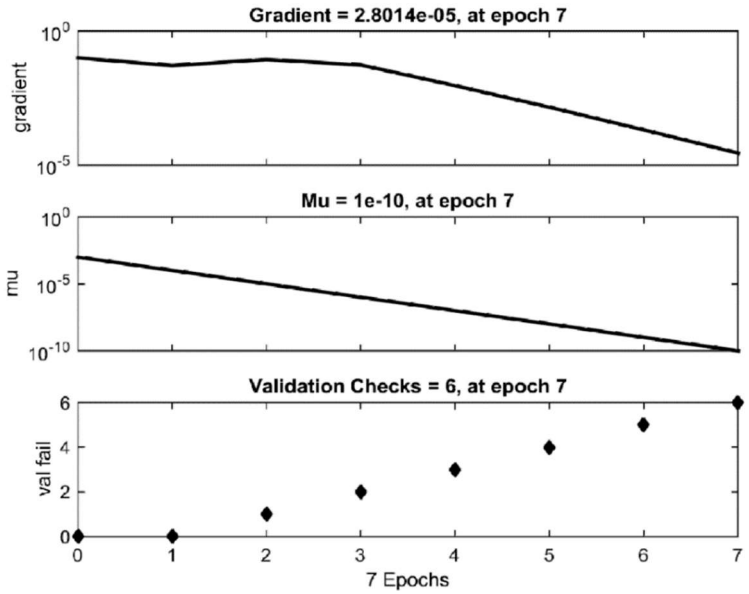


Fig. 4. Training State Data of Neural Network For Cryotherapy by MATLAB.

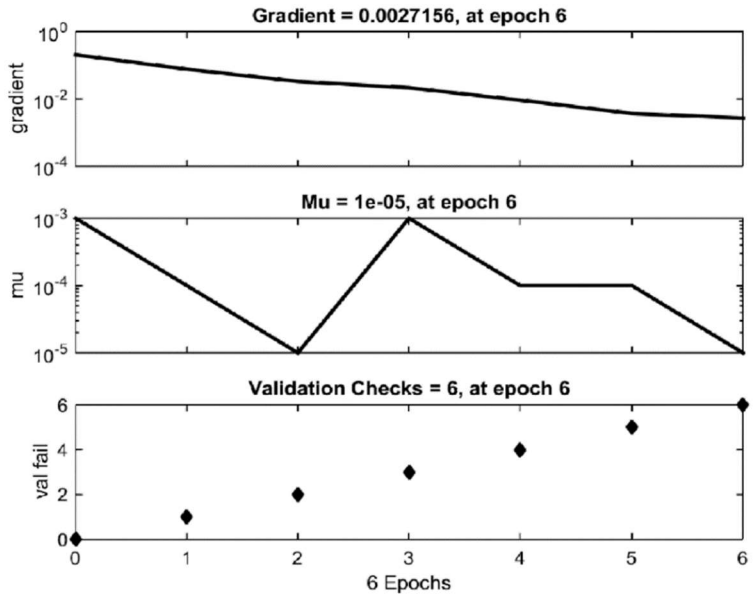


Fig. 5. Training State Data of Neural Network For Immunotherapy by MATLAB.

After acceptable satisfactory levels were reached according to the number of data used, it was necessary to test the success rate of the designed and trained neural networks by the help of MATLAB Simulink model as seen in Fig. 6. The system in Fig. 6 was designed for testing the constructed neural networks using the pre-paired patient data to achieve a performance test.

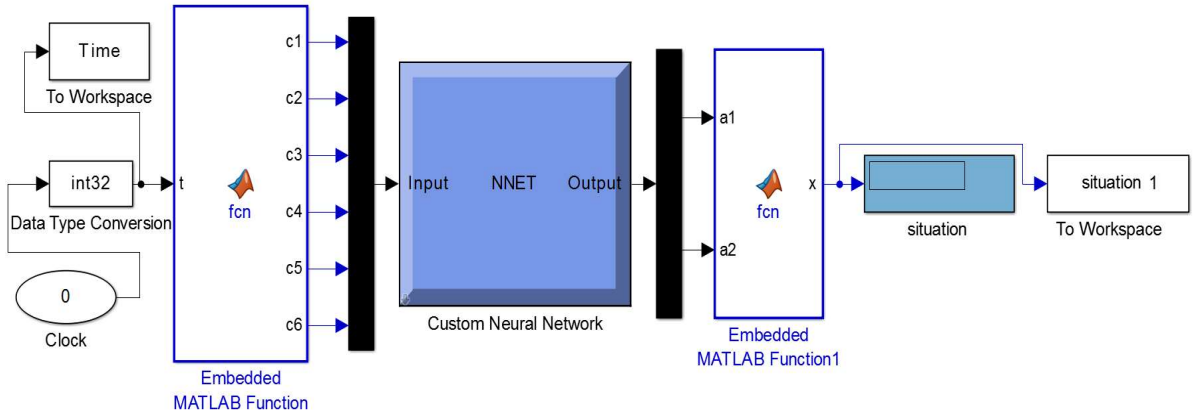


Fig. 6. MATLAB Simulink Model For Neural Network Performance Test.

After performance test for cryotherapy neural network, a success percentage of 94.44% was achieved. And the neural network test results for immunotherapy was 85.6% after the same performance test which was realized using the model seen in Fig. 6. Both cryotherapy and immunotherapy network trainings were practiced using 90 sets of patient data but there is a success percentage difference between them. The cause of this difference is the success level of immunotherapy method. Immunotherapy is more successful with respect to cryotherapy according to the datasets used in this study, it is also a more onerous cure for wart treatment. And this rate of success for cure treatment decreases the variety of outputs as a result reduction. As a result of this situation, the success percentage of the neural network was directly affected.

As it was mentioned before, to find a suitable cure for wart treatment is important. And it is also a noteworthy point that; trying the easier and more effortless cure for the patient is a big advantage. Although the onerous method has a lower level of neural network success percentage. There is a confident way to pre-test the cure type with the neural network to prevent loss of time. This situation will also increase the success percentage and usage numbers of the more effortless cure types for treatments by pre-testing them firstly instead of directly using onerous cures of treatments.

Conclusion

In this study, the most common wart treatment methods, cryotherapy and immunotherapy, were analysed for wart treatment prediction by applying a feedforward neural network. The experimental results show that the successes of feedforward neural network were 94.4% and 85.6%

for cryotherapy and immunotherapy methods respectively. Although acceptable levels of success percentages were obtained by the designed neural networks, it is also possible to achieve better levels of success percentages for the neural networks using more patient data while training the designed neural networks.

This study takes an inspiring role in the cure selecting for wart treatment by obtaining positive results for a better directed preference. Using artificial intelligence for cure selections of treatments will cause obtaining faster and well directed treatment processes.

REFERENCES

[1] K. E. Rowson, and B. W. Mahy, "Human papova (wart) virus," Bacteriological Reviews, vol. 31, no. 2, pp. 110–131, 1967.

[2] D. C. Wiley, and A. C. Cory, Encyclopedia of School Health, Thousand Oaks, CA: SAGE Publications, 2013.

[3] S. K. Loo, and W. Y. Tang, "Warts (non-genital)," BMJ Clinical Evidence, 1710, 2009.

[4] L. Allison, and M. D. Holm, "Warts," in Pediatric Clinical Advisor, vol. II, L. C. Garfunkel, J. Kaczorowski, and C. Christy, Eds. St. Louis: Mosby Elsevier, 2007.

[5] W. Bonnez, "Papillomaviruses," in Mandell, Douglas, and Bennett's Principles and Practice of Infectious Diseases, vol. VIII; J. E. Bennett, R. Dolin, and M. J. Blaser, Eds. Philadelphia: Saunders, an Imprint of Elsevier Inc., 2015.

[6] S. M. Simons, And R. Kennedy, "Foot Injuries," in Clinical Sports Medicine Medical Management And Rehabilitation, W. R. Frontera, S. A. Herring, L. J. Micheli, And J. K. Silver, Eds. Philadelphia: Saunders Elsevier, 2007.

[7] N. H., Boroujeni, And F. Handjani, "Cryotherapy Versus Co2 Laser in The Treatment of Plantar Warts: A Randomized Controlled Trial," Dermatology Practical & Conceptual, Vol. 8, No. 3, Pp. 168-173, 2018.

[8] S. Cockayne, C. Hewitt, K. Hicks, S. Jayakody, A.R. Kang'ombe, E. Stamuli, Et Al. "Cryotherapy Versus Salicylic Acid For The Treatment Of Plantar Warts (Verrucae): A Randomised Controlled Trial," Bmj, 342, P. D3271, 2011.

[9] M. Barrett, N. Harter, And A. Wysong, "Common Pediatric Dermatologic Surgery Procedures," in Pediatric Dermatologic Surgery, K. Nouri, L. Benjamin, J. Alshaiji, And J. Izakovic, Eds. Hoboken: Wiley & Blackwell, 2019, Pp. 73–92.

[10] Y. C. T. Shih, L. S. Elting, A. L. Pavluck, A. Stewart, And M. T. Halpern, "Immunotherapy İn The İnitial Treatment Of Newly Diagnosed Cancer Patients: Utilization Trend And Cost Projections For Nonhodgkin's Lymphoma, Metastatic Breast Cancer, And Metastatic Colorectal Cancer," Cancer Invest, Vol. 28, No. 1, Pp. 46–53, 2010.

[11] S. Garg, And S. Baveja, "Intralesional İmmunotherapy For Difficult to Treat Warts With Mycobacterium W Vaccine," Journal Of Cutaneous And Aesthetic Surgery, Vol. 7, No. 4, Pp. 203–208, 2014.

[12] M. A. Putra, N. A. Setiawan, And S. Wibirama, "Wart Treatment Method Selection Using Adaboost With Random Forests As A Weak Learner," Communications in Science And Technology, Vol. 3, No. 2, Pp. 52–56, 2018.

[13] E. Tanyıldızı, M. Karabatak, G. Yıldırım, And Z. Özpolat, "Siğil Tedavisinde Sınıflandırma Algoritmalarının Performans Analizi," Fırat Üniversitesi Mühendislik Bilimleri Dergisi, Vol. 30, No. 2, Pp. 249– 256, 2018.

[14] K. Akyol, A. Karacı, And Y. A. Gültepe, "Study On Prediction Success of Machine Learning Algorithms For Wart Treatment," International Conference on Advanced Technologies, Computer Engineering And Science, May 11-13, 2018 Safranbolu, Turkey, Pp. 186–188.

[15] H. W. Nugroho, T. B. Adji, And N. A. Setiawan, "Random Forest Weighting Based Feature Selection For C4.5 Algorithm on Wart Treatment Selection Method," International Journal On Advanced Science, Engineering And Information Technology, Vol. 8, No. 5, Pp. 1858–1863, 2018.

[16] Y. Ali, M. Shahzad, And A. Akbar, "Comparison Of Cryotherapy and Immunotherapy in Warts Treatment," International Journal Of Multidisciplinary Sciences And Engineering, Vol. 9, No. 7, Pp. 14–18, 2018.

[17] S. B. Akben, "Predicting The Success Of Wart Treatment Methods Using Decision Tree Based Fuzzy İnformative Images," Biocybern. Biomed. Eng., Vol. 38, Pp. 819–827, 2018.

[18] Https://Archive.İcs.Uci.Edu/Ml/Datasets/Immunotherapy+Dataset

[19] Https://Archive.İcs.Uci.Edu/Ml/Datasets/Cryotherapy+Dataset+

[20] F. Khozeimeh, R. Alizadehsani, M. Roshanzamir, A. Khosravi, P. Layegh, And S. Nahavandi. "An Expert System For Selecting Wart Treatment Method," Computers in Biology And Medicine, Vol. 81, Pp. 167–75, 2017.

[21] F. Khozeimeh, F. Jabbari Azad, Y. Mahboubi Oskouei, M. Jafari, S. Tehranian, R. Alizadehsani, and P. Layegh, "Intralesional İmmunotherapy Compared To Cryotherapy in The Treatment of Warts," International Journal Of Dermatology, Vol. 56, Pp. 474–478, 2017.

[22] L. Wang, H. C. Quek, Hou, K. H. Tee, N. Zhou, And C. Wan, "Optimal Size Of A Feedforward Neural Network: How Much Does it Matter?," Joint International Conference on Autonomic And Autonomous Systems And International Conference On Networking And Services Icas/Icns, October 23-28, 2005 Papeete, Tahiti, French Polynesia, P. 69.

[23] M. Şimşir, R. Bayır, And Y. Uyaroğlu, "Real-Time Monitoring and Fault Diagnosis of A Low Power Hub Motor Using Feedforward Neural Network," Computational Intelligence and Neuroscience, Vol. 2016, 2016.

Estimation of the Demand for the Blood Bank Using Hybrid PCA-ANFIS Method

Seda Hatice GÖKLER, Semra BORAN

Department of Industrial Engineering, Sakarya University, TURKEY

Abstract: Blood is a vital product that is needed by thousands of people every day due to diseases, surgeries or injuries. Blood banks should accurately determine the amount of blood they should have in their stock to meet blood needs. Therefore, having less blood than necessary in hospitals creates important problems such as not meet need for blood and loss of life. On the other hand, storing large amounts of blood causes deteriorating the blood and causes stock out in other hospitals. The aim of this study is to determine the criteria affecting blood demand and to forecast the blood demand by the machine learning algorithm Adaptive Network Based Fuzzy Inference System (ANFIS) method. However, since the number of impact criteria is high, principal component analysis (PCA) method has been used in order to decrease criteria and eliminate the dependencies between the criteria. The developed hybrid method was applied in a regional blood center.

Keywords: ANFIS, PCA, Demand Forecasting, Blood Banking

REFERENCES

[1] Özdamar K. Paket Programlar ile İstatistiksel Veri Analizi-2 (Çok Değişkenli Analizler), 5.Baskı, Eskişehir, Kaan Kitabevi, 2004.

[2] Jang J. ANFIS: Adaptive-Network-Based Fuzzy İnference System. IEEE. 23(3), 665-685, 1993.

[3] Tabachnick BG, Fidell LS. Using Multivariate Statistics. Boston: Allyn and Bacon. 2001.

[4] Çokluk Ö, Şekercioğlu G, Büyüköztürk Ş. Sosyal Bilimler için Çok Değişkenli İstatistik: SPSS ve Lisrel Uygulamaları, Pegem Akademi Yayıncılık, 2012.

[5] Ismail FH, Aziz MA, Hassanien AE. Optimizing The Parameters of Sugeno Based Adaptive Neuro Fuzzy Using Artificial Bee Colony: A Case Study on Predicting The Wind Speed. Proceedings Of The Federated Conference On Computer Science And Information Systems. 2015; 8: 645–651.

[6] Moosavi V, Vafakhah M, Shirmohammadi B, Ranjbar M. Optimization of Wavelet-ANFIS and Wavelet-ANN Hybrid Models by Taguchi Method for Groundwater Level Forecasting Arabian Journal For Science And Engineering. 2014; 39(3): 1785–1796.

[7] Fragiadakis NG, Tsoukalas VD, Papazoglou VJ. An Adaptive Neurofuzzy Inference System (ANFIS) Model For Assessing Occupational Risk İn The Shipbuilding İndustry. Safety Science 2014; 63: 226–235.

[8] Alizadeh M, Lewis M, Zarandi MHF, Jolai F. Determining Significant Parameters in the Design of ANFIS. 2011 Annual Meeting of the North American Fuzzy Information Processing Society 18-20 March 2011 El Paso, TX, USA

[9] Qu N, Zhu M, Ren Y, Dou S. Adaptive Neuron-Fuzzy İnference System Combined with Principal Components Analysis For Determination Of Compound Thiamphenicol Powder On Near-İnfrared Spectroscopy, Journal of the Taiwan Institute of Chemical Engineers 2012; 43:566–572.

[10] Rençber ÖF. Sınıflandırma Problemlerinde Çoklu Lojistik Regresyon, Yapay Sinir Ağı ve Anfis Yöntemlerinin Karşılaştırılması, Seçkin Yayıncılık 2018.

[11] Phootrakornchai W., Jiriwibhakorn S. Online Critical Clearing Time Estimation Using An Adaptive Neuro-Fuzzy İnference System (ANFIS), International Journal of Electrical Power & Energy Systems 2015, 73: 170-181.

Analysing of Multivariate Processes with Machine Learning Algorithms

Deniz DEMİRCİOĞLU DİREN, Semra BORAN, Seda Hatice GÖKLER

Department of Industrial Engineering, Sakarya University, TURKEY

Abstract: It is often not easy to obtain results from complex processes multi variables. Additional techniques and methods are needed to guide. In this study, after the detecting the out of control and under control samples with Hotelling T2 control chart in a multivariate manufacturing process then machine learning algorithms was used to predict the quality of future samples. Four machine learning algorithms were trained and tested by shifts of different magnitude from the process average. The performances of the algorithms were compared according to the accuracy and error rates of the predictions and the most appropriate one was chosen as Multilayer Perceptron.

Keywords: Multivariate Process., Multivariate Control Chart, Machine Learning Algorithm

REFERENCES

[1] H. Hotelling, "Multivariate Quality Control-Illustrated by the Air Testing of Sample Bombsights," Techniques of Statistical Analysis, vol. 2(5), 1947, pp. 110-112.

[2] R. L. Mason, N. D. Tracy and J.C. Young, "Decomposition of T2 for Multivariate Control Chart İnterpretation" J. Qual. Technol. Vol. 27 (2), 1995, pp.99–108.

[3] C. Zhao, C. Wang, L. Hua, X. Liu, Y. Zhang and H. Hu, "Recognition of Control Chart Pattern Using Improved Supervised Locally Linear Embedding And Support Vector Machine", Procedia Engineering vol. 174, 2017, pp. 281 – 288.

[4] M. Zhang and W. Cheng, "Recognition of Mixture Control Chart Pattern Using Multiclass Support Vector Machine and Genetic Algorithm Based on Statistical and Shape Features", Mathematical Problems in Engineering, 2015, pp. 1-10.

[5] S. T. A. Niaki and B. Abbasi, "Fault Diagnosisin Multivariate Control Charts Using Artificial Neural Network, Quality and Reliability Engineering International, vol. 21, 2005, pp.825-840.

[6] M. G. Parra, P. R. Loaiza, "Application of the Multivariate T Control Chart and the Mason–Tracy–Young Decomposition Procedure to the Study of the Consistency of Impurity Profiles of Drug Substances", vol. 16(1), 2003, pp. 127-142.

[7] M. Misra, H. H. Yue, S.J. Q. and C. Ling, "Multivariate Process Monitoring And Fault Diagnosis By Multi-Scale PCA", Computers And Chemical Engineering, vol. 26, 2002, pp. 1281-1293.

[8] F. Aparisi, and J. Sanz, "Interpreting the Out-of-Control Signals of Multivariate Control Charts Employing Neural Networks", International Journal of Computer and Information Engineering, vol. 61, 2010, pp. 226-230.

[9] L.H. Chen, and T.Y. Wang, "Artificial Neural Networks To Classify Mean Shifts From Multivariate X2 Chart Signals", Computers And Industrial Engineering., vol.47, 2004, pp.195-205.

[10] He, S., Wang, G.A., Zhang, M., and Cook, D.F., "Multivariate Process Monitoring and Fault Identification Using Multiple Decision Tree Classifiers", International Journal of Production Research, vol.51 (11), 2013, pp. 3355–3371.

[11] E. Alpaydın, "Yapay Öğrenme", Boğaziçi Üniversitesi Yayınevi, 3. Baskı, 2012, sf. 8-10.

[12] J. Han, M. Kamber, J. Pei, "Data Mining Concepts and Techniques" Third Edition, 2012, pp. 18-20

[13] T. M. Mitchell, "Machine Learning", McGraw-Hill Science/Engineering/Math, 1997, pp.81-85

[14] K-M. Tsai and H-J. Luo, "An Inverse Model for Injection Molding of Optical Lens Using Artificial Neural Network Coupled with Genetic Algorithm", Journal of Intelligent Manufacturing, vol.28, 2017, pp. 473–487.

Comparison of Two Different Social Groups on Twitter with Network Analysis

Adem AKSAN, Ayşe OĞUZLAR

Department of Econometrics, Uludağ University, TURKEY

Abstract: The social communities created by social media, where people have shown great interest, have led to the analysis of social network. In this sense, many techniques have been developed and these techniques have been practised in various fields. By means of the softwares, developed for investigation of complex network analysis, detailed surveys and research can be made about social media. In this study, a social network analysis has conducted on the social media. The study was carried out through the Nodexl program for two different social communities via Twitter to draw the network graphs and to compare the network analysis results by using , "#EytHepBirlikteAnkaraya and #BirlikOlFenerbahçe" According to the results obtained from the analysis, it is observed that although the EYT members are a more recent social group than Fenerbahçe supporters, the ties in their network are stronger and the network density of EYT members is four times more than in Fenerbahçe supporters. In addition, although the hashtags are addressing different topics, the value of the network characteristics such as clustering and centralization were found to be similar to each other.

Keywords: Social Media, Social Network Analysis, NODEXL

Giriş

Bilgisayarın icadı ve yaygınlaşmasının ardından internetin ortaya çıkışı ve bilgisayar ile buluşmasıyla dünyada teknolojik bir devrime sebep oldu. 6 Ağustos 1991 yılında Fransa ve İsviçre sınırında CERN'de World Wide Web projesi kamuoyuna açıklandı. Tim Berner's Lee nin "İnternetteki gezintiniz dünyayı etkiliyor" sözü ile aslında olacakları önceden görmüş gibiydi. (Gürsakal, 2009) Taşınabilir dizüstü bilgisayarlar akabinde akıllı telefonların ortaya çıkması ile internet kullanımı dünyada hızla artmaya başladı. İnternette harcanan vakit arttıkça internet ortamında da çok büyük gelişmeler oldu.

Twitter'da ilk tweet 21 Mart 2006'da Jack Dorsey tarafından atıldı. Günümüzde Twitter'da her dakika ortalama yaklaşık 98 bin tweet atılmaktadır. Atılan tweetler o kadar önem kazanmış ki ünlü Amerikan TV yıldızı Kim Kardeshian attığı her tweet için markalardan 8 bin dolar aldığı Ülkemiz basınında geniş yer buldu. Spor dünyası ile ilgili sosyal medya üzerinde yapılan araştırmada Cristiano Ronaldo'nun Facebook, Instagram ve Twitter'da toplam 262 milyon takipçisi var ve ünlü futbolcunun sosyal medya sayesinde 500 milyon dolar gelir elde ettiği söyleniyor (Altan, 2017).

Hatta bireysellikten çıkarak oluşturulan işletme profilleri ile bir markanın, belirli bir grubun görüşlerini insanlara ulaştırma çabasındadır. Atılan bir tweet beğeni yoluyla binlerce, milyonlarca insana ulaşabilmektedir. Ücretsiz üye olunan bu ve benzeri siteler sayesinde insanlara en hızlı ulaşmanın yolu bilinmektedir. Bireyler, sosyal gruplar, şirketlerin yanı sıra siyasi partilerde sosyal medya aracılığı ile seçmenlerine ulaşmaya çalışmaktadır.

Obama 2008'deki seçim kampanyasın da web teknolojilerini ve sosyal medyayı şüphesiz en iyi kullanan başkan adayıydı. Kampanyanın yürütülmesi için Facebook'un ortak kurucusu olan Chris Hughes ile anlaşıldı. Chris Hughes, Barack Obama'nın seçim kampanyasına odaklanmak için Facebook'taki görevinden istifa etti ve kısa bir sürede kendine bir ekip kurdu. Bu ekiple birlikte Barack Obama'nın sosyal medya stratejilerini geliştirdi. (Stelter, 2008)

Sosyal Medyada yaşanan bu hızlı gelişmeler, sosyal ağ analizinin sosyal medyaya yönelmesine neden oldu. Bu çalışmada Twitter üzerinde trend topic olan farklı konularda ki iki hashtag'in Nodexl programı aracılığıyla ağ grafikleri çizilerek, ağların karşılaştırılması ve yorumlaması yapılmıştır.

1. Sosyal Ağ Analizi

İnsanoğlu var olduğu günden bu yana sosyal ağların bir parçası olmuştur (Hansen, Shnejderman, & Smith, 2011). İnsanoğlu akrabalık, dil, din gibi çeşitli nedenlerle birbiriyle sayısız bağlantı içindedir ve bu bağlar sayesinde oluşan bir ağ dünyası mevcuttur.

Sosyal Ağ Analizi (Social Network Analysis) belirli konularda kişiler, kurumlar, sosyal grup veya topluluklar arasında ki bağlantıların incelenmesi, tanımlanmasında ve yorumlanmasında kullanılan kullanılan bir yaklaşım olarak tanımlanmaktadır (John Scott & Carrington, 2011).

Bireyler ve kuruluşlar gibi sosyal aktörler arasında ki bilgi akışıyla oluşan bağlantıların incelenmesi olarak tanımlanmıştır (Himelboim, Smith, Rainie, Shneiderman, & Espina, 2017). Sosyal ağ: kişilerin ve diğer toplulukların aralarındaki etkileşimi yardımlaşmayı ve etkilerini gösteren bir yapı olarak da tanımlanmaktadır (Karagöz & Kozak, 2014).

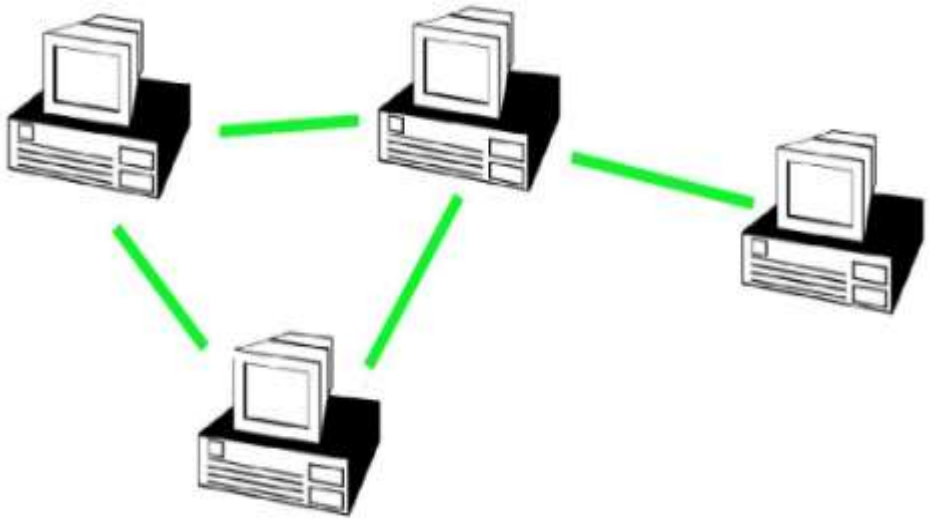
Sosyal ağ analizi, sosyal yapıyı aktörler ve bu aktörleri birbirine bağlayan bir ağ olarak görüp, etkilerini incelemektedir (Gürsakal, 2009). Bu ağlar köşeler ve düğümlerden oluşur.

Sosyal ağ analizinde odak nokta; sosyal varlıklar arasındaki ilişki ve bu ilişki modellerine odaklanmasıdır (Karagöz & Kozak, 2014).

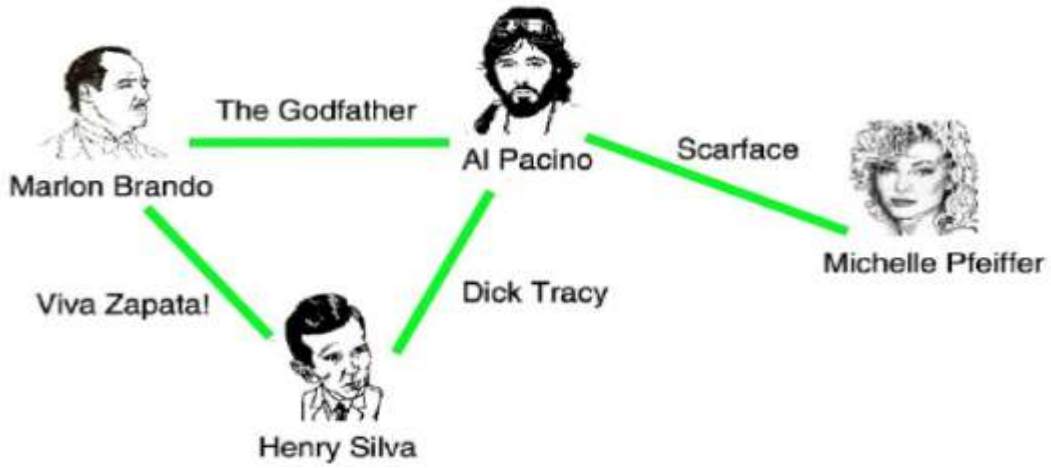
Modern sosyal ağ analizi dört temel özelliğe sahiptir. Bunlar:

1. Sosyal ağ analizi, sosyal aktörleri birbirine bağlayan bağlara dayanan yapısal bir sezgiyle motive edilir,
2. Sistematik ampirik verilere,
3. Büyük ölçüde grafik görüntülerden yararlanır ve
4. Matematiksel ve / veya hesaplamalı modellerin kullanımına dayanır (Freeman L. C., 2004).

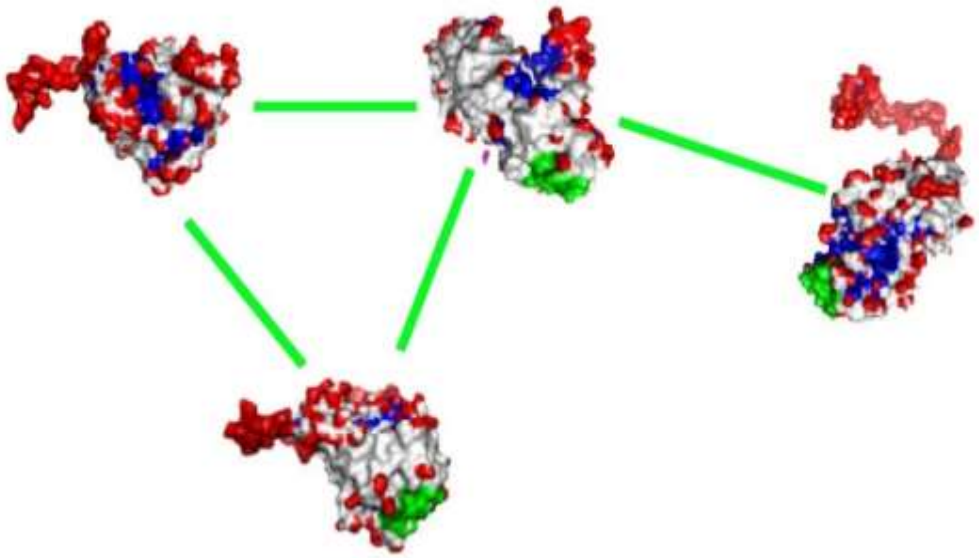
Ağ olarak incelenen sosyal yapı örnekleri, okuldaki çocuklar arasındaki dostluk, sosyal seçkinlerin üyeleri arasındaki aile ilişkileri, şirketlerin ortak yönetim kurulu üyeleri, ülkeler arasındaki ticari ilişkiler ve web siteleri arasındaki köprülerdir (Barabási, 2002).



Şekil 1. İnternet Ağı (Barabási, 2002)



Şekil 2. Hollywood Aktör Ağı (Barabási, 2002)



Şekil 3. Protein Etkileşim Ağı

Yukarıda gösterilen üç farklı ağlardan;

- 1'nci ağ; yönlendiricilerin (uzman bilgisayarların) birbirine bağlandığı küçük bir internet ağını,
- 2'nci ağ; filmde oynarlarsa iki aktörün birbirine bağlandığı Hollywood aktör ağı,
- 3'ncü ağ (c) Hücrede birbirlerine bağlanabilecekleri deneysel kanıt varsa, iki proteinin bağlandığı bir protein-protein etkileşim ağını göstermektedir.

Düğümlerin ve bağlantıların niteliği farklı olsa da bu ağlar, 4 düğüm ve 4 bağlantıdan oluşan aynı grafiksel gösterime sahiptir (Barabási, 2002).

Sosyal ağ analizini ağlar olarak kavramsallaştırılabilecek verilere uygulamanın her zaman yararlı veya gerekli olduğunu söylemek doğru değildir. Örneğin, eğer bir araştırmacı, bir kişinin yardım için başvurabileceği kişi sayısını bilmekle ilgileniyorsa, bağların yapısı yerine bağların sayısı önemlidir ve ağ analizine gerek yoktur. Ağ analizinin uygulanabilmesi için, sosyolojiden ya da diğer sosyal ve davranış bilimlerinden gelen teori, bağların yapısının ağ üyelerinin davranışları, görüşleri ya da sosyal konumlarıyla bağlantılı olduğuna inanmak için sebepler vermelidir (Barabási, 2002).

Son yıllarda geliştirilen teknikler sayesinde sosyal ağ analizi değişik alanlarda da uygulanmaya başlamıştır. Karmaşık ağ analizlerinin incelenmesi için geliştirilen yazılımlar sayesinde ağlar konusunda detaylı araştırma ve incelemeler yapılabilmektedir.

Twitter, Facebook ve Instagram gibi sosyal medya sitelerinde gösterilen yoğun ilgi sosyal medya üzerinde Sosyal Ağ Analizi çalışmalarına ivme kazandırmıştır. 2018 yılı global digital raporu verilerine göre (DIGITAL AROUND THE WORLD IN 2018) dünyada 3 milyar 196 milyon kişi sosyal medya kullanmaktadır. Ortalama bir internet kullanıcının internette geçirdiği süre ise günde ortalama 6 saat (Kemp, 2018).

Milyarlarca insanın günde ortalama 6 saat vakit geçirdiği ortamda şirketler, siyasi parti ve liderler, popüler kişi ve gruplar sosyal medya sayesinde tanıtımlarını yaparak hedef kitlelerine ulaşmaya çalışmaktadırlar.

2. Nodexl

Sosyal Medya Araştırma Vakfı (The Social Media Research Foundation) tarafından Temmuz 2008'de kullanıma sunulan Nodexl programı Microsoft Excel'e eklenti olarak eklenip Twitter, Facebook, Youtube ve Flickr'dan karmaşık sosyal ağları toplamanızı, analiz etmenizi ve görselleştirilmesini sağlamaktadır. (The Social Media Research Foundation)

Kısaca NodeXL sosyal medyadan veri toplayarak ağ görselini ve raporunu oluşturur. Basit filtreleme ve ekran özellikleri ile ağlarda önemli vurguları yapmak için kullanılır. NodeXL de işlemler 5 adımda yürütülmektedir.

1. Çeşitli kaynaklardan veri toplama

2. Veri depolama

3. Veri Analizi

4. Görselleştirme

5. Yayınlama (The Social Media Research Foundation)

Nodexl ile yapılan çalışmaları incelediğimizde; Himelboim ve arkadaşları (2017), Twitter üzerindeki ağları sınıflandırarak bilgi akışına göre modellemişlerdir (Himelboim, Smith, Rainie, Shneiderman, & Espina, 2017).

Pew Araştırma Merkezi ve Sosyal Medya Araştırma Vakfı tarafından yapılan çalışmada binlerce Twitter görüşmesinin özel bir analizi ile Twitter'da gerçekleşen konuşma ve sosyal yapılara altı farklı kalıbın olduğunu ortaya çıkarmışlardır (Marc Smith, 2014).

IEEE Üçüncü Uluslararası Sosyal Bilgi İşlem Konferansında Rodriges ve arkadaşları (2011), Sosyal ağlarda yer alan topluluklarda ki bireylerin hangi kategorilere göre (yaş, cinsiyet, meslek, coğrafi konum vb.) birbirleriyle etkileşim içinde olduğunu incelemişlerdir (Eduarda Mendes Rodrigues, 2011).

Hibeilboim ve arkadaşları (2009), forum sitelerinde yapılan politik tartışmaları sonucu ortaya çıkan sosyal ağlarda tartışmayı başlatan, büyüten ve tartışmanın içeriğini belirleyen ve değiştiren kullanıcılar(katalizör) tespit edilmeye çalışılmıştır (Himelboim I. E., 2009).

Araştırmamızda sosyal medya siteleri içerisinde yer alan Twitter üzerinden yapılacaktır. Twitter 2006 yılında kullanılmaya başlanılmış olup 2018 yılı global digital raporuna göre Twitter'da kullanıcı sayısı 330 milyondur (Kemp, 2018).

Bir sosyal ağ perspektifinden sosyal medya araştırması, odağı bireysel özelliklerden sosyal varlıklar arasındaki ilişkisel bağlara kaydırır. (Bruns & Stieglitz, 2013) Bu bağların koleksiyonları, ortaya çıkan desenler veya ağ motifleri halinde toplanır. Sosyal ağ sitelerinde, kullanıcılar kendileriyle bağlantı kurdukları veya bilgi paylaştıkları zaman diğer kullanıcılarla etkileşim kurarak ağ oluştururlar. Twitter'da sosyal ağlar, kullanıcılardan bahsettiklerinde ve birbirlerine cevap verdiklerinde ve diğer kullanıcılarla kurdukları bağlantılardan oluşur (Hansen, Shnejderman, & Smith, 2011).

2.1. Ağlara ilişkin Ölçümler

Düğümler (Vertices), aracılar, varlıklar veya öğeler olarak da adlandırılan tepe noktaları birçok şeyi temsil edebilir. Genellikle insanları veya çalışma grupları, takımlar, organizasyonlar, kurumlar, eyaletler ve hatta ülkeler gibi sosyal yapıları temsil ederler. Diğer zamanlarda, web sayfaları, anahtar kelime etiketleri veya videolar gibi içerikleri temsil ederler. Fiziksel veya sanal konumları veya olayları bile temsil edebilirler.

Bağlantılar (Edges), bağlar, bağlantılar ve ilişkiler olarak da bilinen kenarlar, ağların yapı taşlarıdır. Bir kenar iki köşeyi birbirine bağlar. Kenarlar, yakınlık, iş birlikleri, akrabalık, dostluk, ticari ortaklıklar, alıntılar, yatırımlar, köprü oluşturma, işlemler ve paylaşılan özellikler gibi birçok farklı ilişkiyi temsil edebilir. Resmi bir statü statüsüne sahipse, katılımcılar tarafından tanınıyorsa veya aralarındaki değiş tokuş veya etkileşimle gözlemlenirse, bir bağın var olduğu söylenebilir. Bir bağ, iki varlık arasındaki herhangi bir ilişki veya bağlantı şeklidir.

Bir diğer ölçüm ise "yoğunluk" (density) dur. Yoğunluk düğümler arasındaki bağlantıların yoğunluğunu veya seyrekliğini göstermektedir. Bir ağın yoğunluğu, mevcut bağlantı sayısının mümkün olan maksimum bağlantı sayısına oranı olarak gösterilmektedir (Wasserman & Faust, 1994).

"Çap" (diameter) ağdaki bütün düğüm çiftlerinin arasındaki en kısa patikaların en uzunudur. Çap ne kadar kısa olursa o ağda bilgi o kadar hızlı yayılır. (Gürsakal, 2009)

Öz döngüler (self-loops) kendi içinde bağlantı olmayan düğümleri gösterir. Kullanıcının ağ içerisinde kendi kendine bağlı olduğunu göstermektedir. Marco Bastos Şubat 2014' te yayınladığı makalesinde "öz döngüler ağların yapısı ile yakından ilişkilidir ve kullanıcılar hastac' leri kendi profilin de yayınlama eğilimindedir." Öz döngülerin ağlarda ki önemine vurgu yapmıştır (Bastos, 2014).

2.1.1 Ağlarda Kümelenme (Clusters ve Modularite (Modularity)

Pek çok ağın ortak bir özelliği, aynı üçüncü düğüme bağlı olan iki düğümün de birbiriyle bağlanma olasılığının daha yüksek olduğu kümeleme veya ağ geçirgenliğidir. Basitçe söylemek gerekirse, belirli bir topluluktan yapılacak rastgele seçimden ise arkadaşlarınızdan ikisinin birbirlerini tanıma olasılığı daha yüksektir. Bir ağ büyüdükçe, kümeler daha büyük bir ağ içindeki düğüm alt kümeleri

daha fazla birbirine bağlanır, yani yoğunlaşır bu kümeler arasındaki bağlantılar daha az yoğundur (Newman & Girvan, 2004).

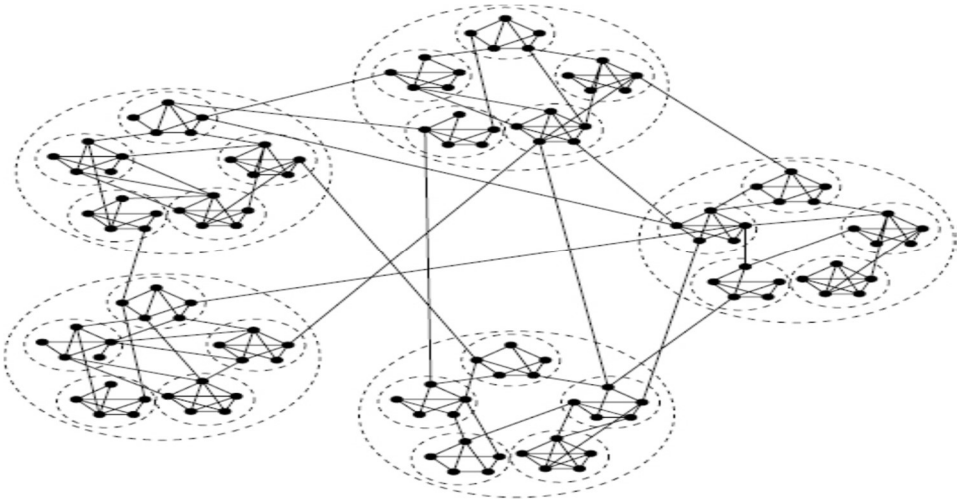
Twitter'da kullanıcılar birbirlerini takip ederek bilgi akışı için yol oluşturmuş olurlar. Bu yöntem ile kullanıcıların ilgi alanlarına göre gruplar oluşmuş olur. Ortaya çıkan bu gruplar bilgi akışının sosyal sınırını (Himelboim, Smith, Rainie, Shneiderman, & Espina, 2017) belirlemiş olur. Oluşan bu kümelerde bilgi serbestçe akarken kümeler arasında ki bilgi akışı, kümeler arasındaki bağlantılarla sağlanmaktadır.

Sosyal medya ağlarında ki kümelenmenin önemini Conver ve arkadaşları (2011), siyasi düşünceleri benzer kullanıcıların retweet kalıplarını ortaya çıkarmışlardır. (M. D. Conover, 2011)

Ağ kümelerinin bilgi akışında ki rolü daha sonra Rodrigez ve arkadaşlarının (2012), çalışması ile de gösterildi (Rodrigez, Leskovec, & Krause, 2012).

Kümelenme katsayısı (Clustering Coefficient)

Kümelenme katsayısı, yoğunluk(density) ölçüsüne benzerdir. Eğer ağda kullanıcılar birbirini tanıyorsa yüksek, bir kümelenme katsayısına sahip olursunuz. Eğer "arkadaşlarınız" (değiştirir) birbirinizi tanımıyorsa, düşük kümelenme katsayısına sahip olursunuz. Kümelenme katsayısı, kullanıcıların başkalarıyla bağlantı kurma biçimlerine ve içinde bulundukları ortamlara bağlı olarak farklı ölçütlere sahip olabilir.



Şekil 4. Kümelenmiş ağ örneği (Kaiser, Görner, & Hilgetag, 2007)

Modülarite (Modularity) Ağın yapısının modülerliği, birbirine bağlı alt kümelere bölünmüş bir ağ olan kümelenme kalitesinin bir ölçüsüdür. (Newman & Girvan, 2004) Modülerlik,

kümelenmiş ve birleştirilmiş yapılara sahip ağlar arasında ayrım yapan kümelerin birbirinden ayrılma derecesini (0'dan 1'e kadar bir değer aralığı) yakalar. Ayrıca, iki çok farklı türde yoğun ağların ayırt edilmesine yardımcı olur. (Himelboim, Smith, Rainie, Shneiderman, & Espina, 2017) Modülerlik, kümeler arasında bölünmenin ne kadar iyi olduğunu, kümeler içinde birçok bağlantı olduğunu ve sadece birkaçının olduğunu ölçmektedir. Modülerlik değerleri 0 ile 1 arasında değişmektedir. Modülerlik değeri ne kadar yüksekse, kümeler o kadar belirgin veya ayrıktır, yani kümeler birbiriyle daha az bağlantılıdır (Himelboim, Smith, Rainie, Shneiderman, & Espina, 2017).

2.1.2 Ağlarda Merkezileştirme (Centralization)

Herhangi bir ağın merkezileşmesi, en merkezi düğümünün diğer tüm düğümlerin merkeziyetiyle ne kadar merkezi olduğuna dair bir ölçüdür (Freeman L. C., 1979).

Bir ağda bilgi akışı kullanıcılar arasında ki bağlantıların dışında ağ yapısının hiyerarşik veya eşitlikçi olmasıyla ilgilidir. Bir veya birkaç sosyal aktörün sayısız bağlantılarının olduğu bir ağdaki bilgi akışı bu aktörlere bağlı olarak daha hiyerarşik iken, bağlantı sayılarının azalarak daha fazla sosyal aktörlere yayıldığı ağlar daha merkezidir şeklinde tanımlayabiliriz (Himelboim, Smith, Rainie, Shneiderman, & Espina, 2017).

Van Den Bos (2006), Hollanda'daki İranlıların web siteleri arasında köprüler tarafından oluşturulan çevrimiçi topluluk yapısının, bu ağda odaklanan birçok nokta ile düşük merkezileşme gösterdiğini ortaya koymuştur.

Woo-young ve Park (2012), yaptıkları çalışmada ABD yer alan haber bloglarının bulunduğu ağın merkezi bir ağ olduğunu tespit etmiştir (Woo-young & Park, 2012).

Merkezilik bir tür popülerlik ölçütü olarak düşünülebilir. Derece, belirli bir tepe noktasına bağlı toplam kenar sayısının ölçüsüdür. Yönlendirilmiş ağlar için iki derece ölçüsü vardır. Bir köşeye içe dönük bağlantıların sayısı girdi derecesi (Input degree), bir köşeden kaynaklanan ve diğer köşelere dışa dönük bağlantılar ise çıktı derecesi (output degree) olarak bilinmektedir (Gürsakal, 2009).

Arasındalık Merkeziliği (Betweenness Centrality), bir mesafe ölçüsüdür. Yol kavramı ağların araştırılmasında merkezi bir öneme sahiptir. Belki de bir ağdaki herhangi iki kişi hakkında sorulacak en doğal sorulardan biri "Ne kadar uzaktalar? "Bu mesafe basitçe ölçülür: komşusu

olmayan insanlar arasındaki mesafe, en az sayıda komşu-komşu atlama sayısı birinden diğerine ölçülür. Örneğin, sizin komşunuz olmayan, ancak komşularınızın komşusu olan insanlar, sizden 2 uzakta. İki kişi arasındaki en kısa yol "jeodezik mesafe" olarak adlandırılır ve birçok merkezi ölçümde kullanılır (Himelboim I. E., 2009).

Yakınlık merkezliliği (Closeness centrality), bir düğüm ile ağdaki diğer her düğüm arasındaki ortalama mesafeyi ifade etmektedir. Düğümlerin yalnızca mesajlarını yalnızca mevcut bağlantılarına iletebileceğini veya verebileceğini varsayarsak, düşük bir yakınlık merkezi olması, bir kişinin doğrudan bağlı olduğu veya ağdaki çoğu kişiden daha hızlı bir şekilde diğer kişilere ulaşabildiği anlamına gelmektedir. Örneğin; şehirde yaşayan insanların şehir merkezine ortalama mesafesini yakınlık merkezi ölçüsü olarak düşünebiliriz (Denny, 2014).

Özvektör merkeziyeti: Bir aktörün diğer iyi aktörlere bağlanma derecesini ölçer. Özvektör merkeziyetçiliği, daha sofistike bir merkeziyet anlayışıdır. Az sayıda bağlantıya sahip bir kişi, eğer bu az bağlantıların kendisi çok iyi bağlanmışsa, çok yüksek öz vektörlük merkeziyetine sahip olabilir. Öz vektör merkezliliği, bağlantıların değişken bir değere sahip olmasını sağlar, böylece bazı köşelere bağlanmak diğerlerine bağlanmaktan daha fazla yarar sağlar.

3. Uygulama

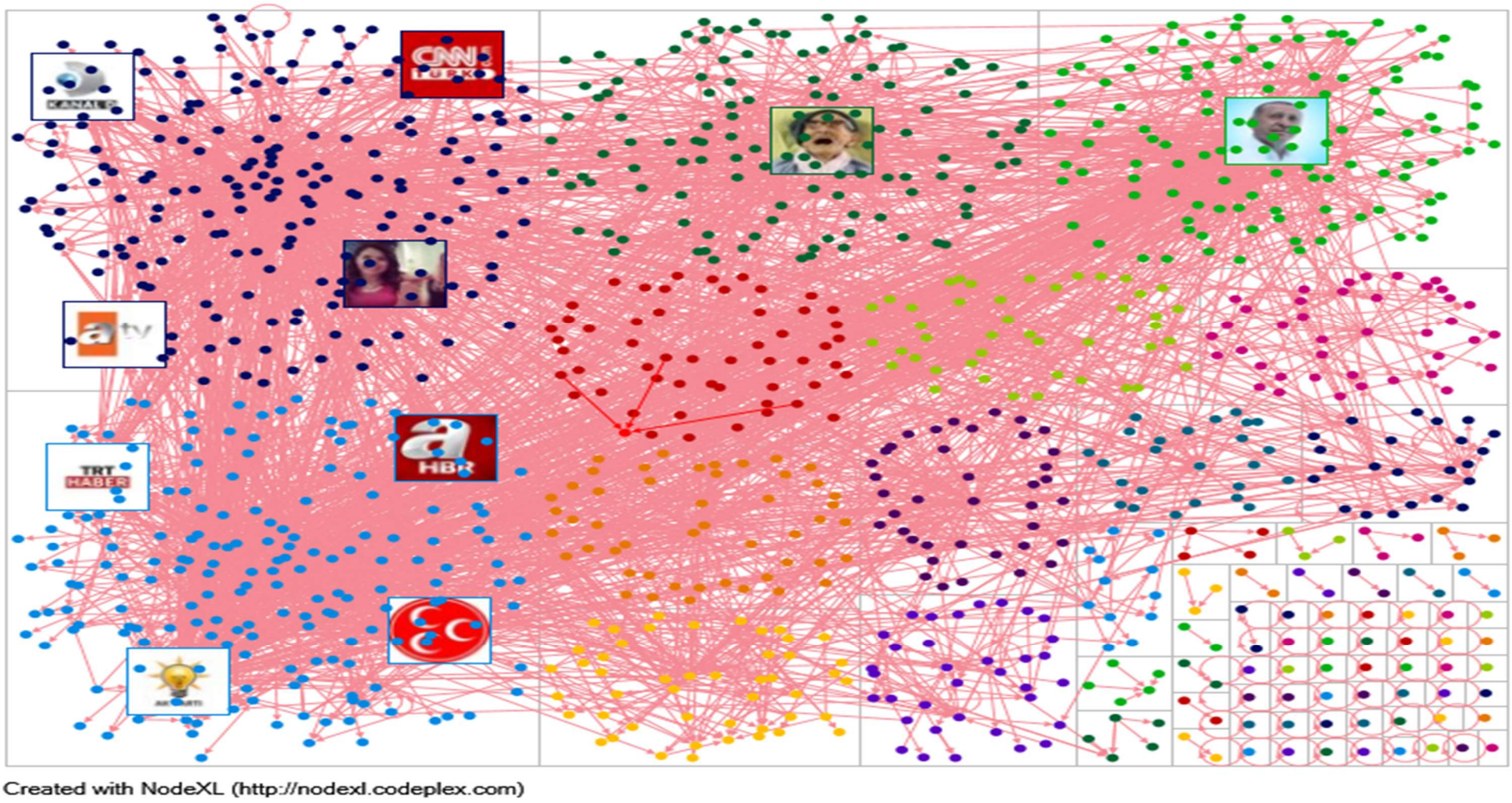
Ağ grafikleri, bir etkinliğin temelini oluşturan sosyal yapıyı anlamak, bir olayla ilgili kilit kişileri tanımlamak, etkinliğin etrafındaki sohbeti haritalamak ve zaman içinde izlemek ve ilgili olayları karşılaştırmak için kullanılabilir (Hansen, Shnejderman, & Smith, 2011).

Son dönemlerde emeklilikte yaşa takılanlar (EYT), gündemde ve sosyal medyayı çok aktif kullanılıyorlar. Bu grubun sosyal yapısını anlamak ve kilit kullanıcıları tanımlamak ve EYT'lilere göre daha eski bir sosyal grup ile karşılaştırmasını yapmak amacıyla;

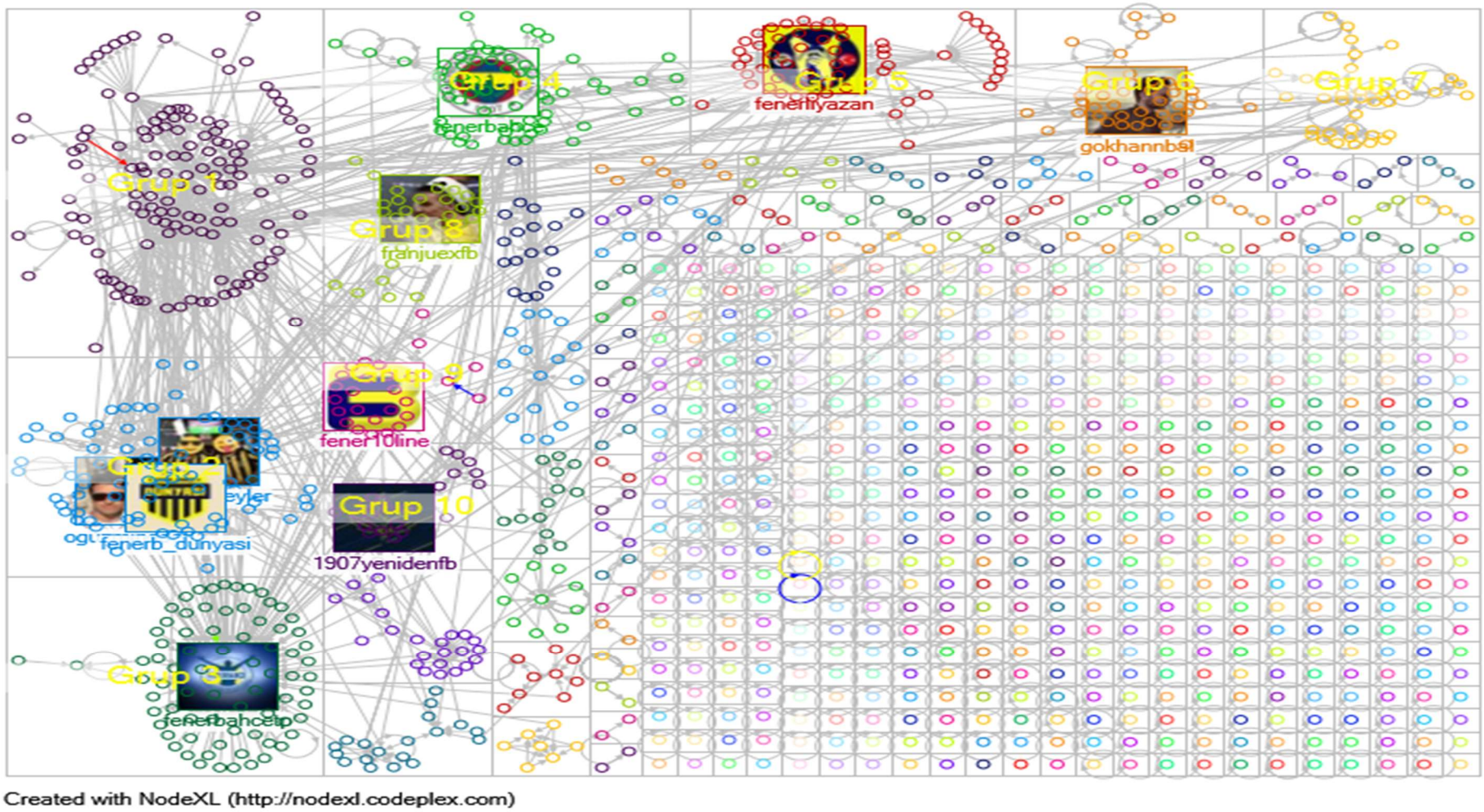
09.02.2019-10.02.2019 tarihleri arasında Twitter'da trend topic olmuş hashtagler arasından #EytHepBirlikteAnkaraya ve #BirlikOlFenerbahçe seçilmiş ve Nodexl aracılığıyla ağ grafikleri çizilmiştir. [1]Şekil 5 ve Şekil 6'daki ağ grafikleri incelendiğinde 2.ağda kümelenmenin daha fazla olduğu ve gruplar arasında ki bağlantıların 1.ağa göre daha az olduğu gözlemlenmiştir.

Nodexl aracılığıyla elde edilen grup istatistiklerine bakıldığında 1.ağda 72, 2.ağda 742, grup olduğunu tespit etmiştir. Çizilen bu iki ağ grafiği de Newman ve Girvan'ın 2004 yılındaki çalışmasında kümelenme hakkında yaptığı tanımı destekler niteliktedir.[1] 2.ağda ki bilgi akışı grupların kendi içinde ki bağlantılarla sağlanırken, 1.ağda ise bu bağlantılar daha etkin kullanıcılar sayesinde gerçekleşmektedir.

Ağlarda ki düğümler genellikle dairesel şekilde gösterilir. Bu çalışmada ağdaki en aktif düğümleri, bir nevi ağlardaki başrolleri, popüler kişileri öne çıkarmak amacıyla, ilk 10 düğümü profil resimleri ile gösterilmiştir. Burada ilk 10'u belirlemek için en fazla bağlantı alan (Input) düğüme göre sıralama yapılarak ağ grafikleri çizilmiştir.



Şekil 5. **#EytHepBirlikteAnkaraya** Kümelenmiş Ağ Grafiği



Şekil 6. **#BirlikOlFenerbahçe** Kümelenmiş Ağ Grafiği

Tablo 1. Grafik Ölçümleri (Graph Metrics)

Grafik Ölçümleri	1.Ağ #EytHepBirlikteAnkaraya	2.Ağ #BirlikOlFenerbahçe
Vertices(düğüm)	1072	1364
Unique Edges (benzersiz kenarlar)	2785	1725
Edges With Duplicates (tekrarlanan kenarlar)	403	342
Total Edges (toplam kenar)	3188	2067
Self-Loops (öz döngü)	129	768
Reciprocated Vertex Pair Ratio (karşılıklı düğüm çiftlerinin oranı)	0,000345066	0,004966887
Reciprocated Edge Ratio (karşılıklı kenar oranı)	0,000689893	0,009884679
Connected Components (birbiriyle bağlı bağlantılar)	57	578
Single-Vertex Connected Components (tek bağlantı)	41	525
Maximum Vertices in a Connected Component (maksimum düğüm sayısı)	994	704
Maximum Edges in a Connected Component (maksimum kenar sayısı)	3125	1338
Maximum Geodesic Distance (Diameter)(çap)	9	10
Average Geodesic Distance (ortalama çap)	3,833291	3,868626
Graph Density (yoğunluk)	0,002478852	0,000652993
Modularity (modularite)	0,42884	0,568192

1.ağımızda toplam 1072, 2. Ağımızda toplam 1364 düğüm bulunmakta olup, 2.ağ daha büyük bir ağdır. 1. Ağdaki kenar sayısı (3188), yani ağ içinde ki bağlantı sayısı 2.ağa (2067) göre daha fazladır. 2.ağda öz döngü sayısı 1.ağa göre daha fazla olup 768 adet bağlantının diğer bağlantılarla ilişkisi olmayıp sadece kendi profillerinde paylaşımda bulunulmuştur.

2.ağda karşılıklı düğüm oranı (0,009) 1.ağa (0,0003) göre büyük olup 2.ağda ki düğümlerin etkileşimi, daha büyüktür. Aynı yorumu karşılıklı kenar oranı içinde yapabiliriz. 2 ağdaki bağlantıların birbirleriyle etkileşimi daha büyüktür. Düğüm ve kenar oranlarının yorumuna göre birbirleriyle bağlı bağlantı sayısında 2.ağda daha fazla olması beklenilmelidir. Tablo 1'i incelediğimizde 2.ağdaki birbirleriyle bağlı bağlantı sayısı (578) daha fazladır.

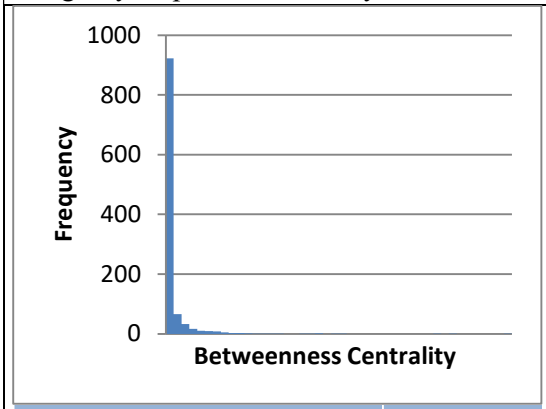
Tek bağlantılara baktığımızda 2.ağdaki (525) bağlantı sayısı daha fazladır. 1.ağdaki Birbirine bağlı maksimum düğüm (994) ve maksimum kenar (3125) sayısı 2.ağa göre daha fazladır.

Maksimum çap 2.ağda (10) daha büyüktür. Ortalama çaplara baktığımızda ise tüm düğümler arasındaki ortalama uzaklıklar (1.ağ 3,83-2.ağ 3,86) birbirlerine yakın değerdedir. Bu değer sıfıra yaklaştıkça ağdaki bilgi daha hızlı yayılır.

1.ağın yoğunluğu (0,0024), 2.ağa (0,0006) göre daha fazla olup bu değer 1'e yaklaştıkça ağdaki bilgi akışı ve etkileşim artar. Modularite değerlerini karşılaştırdığımız da 2. Ağın modularite değeri (0,56) 1'e daha yakın olup, 2.ağdaki kümelenme daha kalitelidir. Bunun anlamı 2.ağda kümelenme daha belirgin, kümeler arasında ki bağlantılar daha azdır. Ağlarda ki gruplar kümelenme işlemi sonucu oluşur. Kümeleme işlemi bağlantılar ne kadar az ise o kadar belirgin olur. Buna göre Twitter'da kullanıcıların birbirlerine çok bağlı olduğu ağlar daha az modüler veya kullanıcıların birbirine daha az bağlı olduğu ağlar daha fazla modülerdir diyebiliriz (Newman & Girvan, 2004).

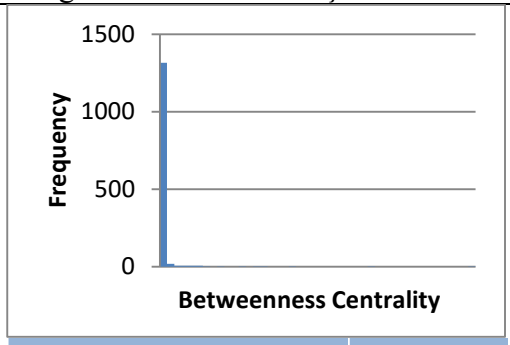
Tablo2. Arasındalık Merkeziliği Değerleri (Betweenness Centrality)

1.Ağ #EytHepBirlikteAnkaraya



Minimum Betweenness Centrality	0,000
Maximum Betweenness Centrality	153965,372
Average Betweenness Centrality	2873,237
Median Betweenness Centrality	58,551

2.Ağ #BirlikOlFenerbahçe



Minimum Betweenness Centrality	0,000
Maximum Betweenness Centrality	195580,697
Average Betweenness Centrality	1045,268
Median Betweenness Centrality	0,000

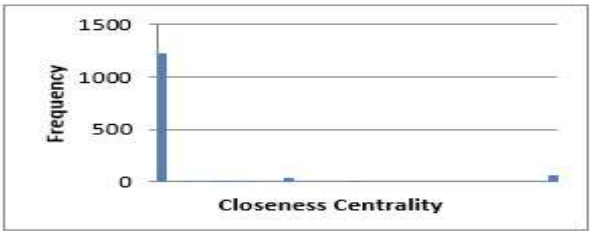
Arasındalık merkeziliği ölçülerine baktığımızda iki düğüm arasındaki en kısa mesafe 2.ağda daha yüksek iken, tüm düğümler arasından iki düğüm arasında ki ortalama mesafe 2. Ağda daha düşüktür. 2.ağda kullanıcılar arasında ki komşuluk bağları daha iyidir. Bu ölçüm değeri sıfıra yaklaştıkça komşuluk bağları artar.

Tablo 3. Yakınlık Merkeziliği Değerleri (Closeness Centrality)

1.Ağ #EytHepBirlikteAnkaraya

Minimum Closeness Centrality	0,000
Maximum Closeness Centrality	1,000
Average Closeness Centrality	0,028
Median Closeness Centrality	0,000

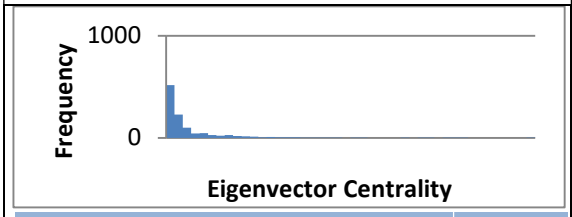
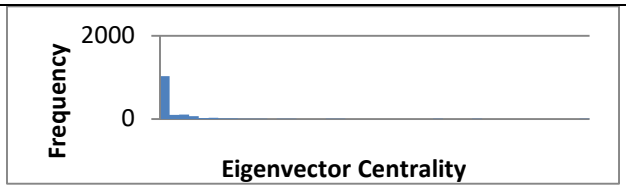
2.Ağ #BirlikOlFenerbahçe



Minimum Closeness Centrality	0,000
Maximum Closeness Centrality	1,000
Average Closeness Centrality	0,062
Median Closeness Centrality	0,000

Yakınlık ölçüsünü bir düğümün diğer tüm düğümlere ortalama mesafesi olarak tanımlamıştık. İki ağda ki ortalama ölçüm değerlerine baktığımızda 1.ağda ortalama değer (0,028) daha düşük olup 1.ağda ki bir kullanıcı 2.ağda ki bir kullanıcıya göre daha hızlı bir şekilde diğer kullanıcılara ulaşabilmektedir. EYT'lilerin sosyal medyada bu kadar etkin olmalarının sebebi de zaten bu olsa gerek.

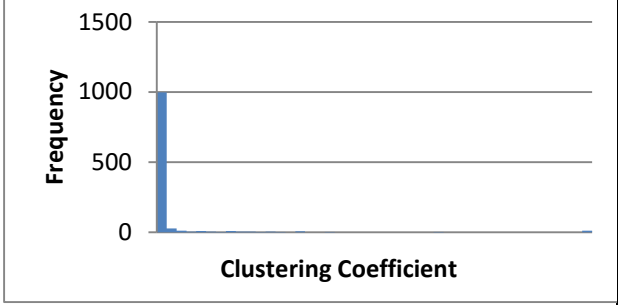
Tablo 4. Öz vektör Merkeziliği Değerleri (Eigenvector Centrality)

1.Ağ #EytHepBirlikteAnkaraya		2.Ağ #BirlikOlFenerbahçe	
			
Minimum Eigenvector Centrality	0,000	Minimum Eigenvector Centrality	0,000
Maximum Eigenvector Centrality	0,015	Maximum Eigenvector Centrality	0,030
Average Eigenvector Centrality	0,001	Average Eigenvector Centrality	0,001
Median Eigenvector Centrality	0,000	Median Eigenvector Centrality	0,000

Özvektör merkeziliği ait ortalama ölçüm değerine baktığımızda ise değerler birbirine eşit. Her iki ağda da düğümlerin birbirine ortalama aynı derece ile bağlanıyor. Her iki ağda da düğümler arasındaki bağlar güçlü.

Yukarıda Merkezilik ölçüm sonuçlarına ilişkin değerlendirmeleri yaptık. 2 ağımızda da derece dağılımına baktığımız kuvvet yasası geçerli gibi. Az sayıda kullanıcının etkinliği ağlarımızda mevcut.

Tablo5. Kümelenme Katsayısı (Clustering Coefficient)

1.Ağ #EytHepBirlikteAnkaraya	2.Ağ #BirlikOlFenerbahçe	
		
	Minimum Clustering Coefficient	0,000
	Maximum Clustering Coefficient	1,000

Minimum Clustering Coefficient	0,000	Average Clustering Coefficient	0,009
Maximum Clustering Coefficient	0,500	Median Clustering Coefficient	0,000
Average Clustering Coefficient	0,010		
Median Clustering Coefficient	0,000		

Ortalama kümelenme katsayısına baktığımızda ise değerler birbirine yakın kümelenme katsayısı sıfırdan uzaklaştıkça kullanıcıların birbirini daha fazla tanıdığı yorumlanabilir. Bu katsayıya göre 1.ağ azda olsa 2.ağa göre birbirini daha fazla tanımaktadır.

Sonuç

Bilgi çok hızlı ve kontrolsüz bir şekilde yayılabiliyor. Sosyal medya, sosyal aktörleri ve bu aktörlerden oluşan toplulukları önemli hale getirdi. Bizimde amacımız sosyal medya da topluluklar aracılığıyla gerçekleştirilen bilgi akışını kümelenmiş ağ grafikleri ile göstermek ve ve sosyal medyada ki etkin grupların oluşturduğu trend topicleri ağ grafikleri ve ölçümleri ile ağları karşılaştırmak ve doğru bir şekilde yorumlayabilmekti. Bunun için ağların görselleştirilmesi ve ölçümleri konusunda Nodexl programı kullanıldı.

İki ağda farklı konulardan oluşan hashtaglerden oluşsa da ağların ölçüm değerlerine baktığımızda birbirine yakın değerler elde edildi. Trend topic olmaları bu ağların benzer sonuçlar göstereceği düşünülse de ağları birbirinden ayıran farklı dinamikler olduğunu tespit edildi. 2.ağda ki grup sayısı 1.ağa göre daha fazla olsa da (1.ağ 72, 2.ağ 742) kümelenme katsayıları birbirine çok yakın değerler (1.ağ, 0,010, 2.ağ 0,009) elde edilmiştir. 1.ağda ki emeklilikte ki yaşa takılanlar, 1.ağda ki Fenerbahçelilere göre daha yeni grup olsalar da aralarındaki sıkı etkileşim nedeniyle birbirlerini daha fazla tanımaktadırlar. Modularite değerlerine göre ise 2.ağda kümelenmenin (0,56), 1.ağa göre daha kaliteli ve belirgin olduğunu bize gösterdi. 1.ağın yoğunluğu 2.ağa göre 4 kat daha fazladır (0,0024-0,006). 1 ağda bağlar daha sıkı ve bilgi akışı 2.ağa göre 4 kat daha hızlıdır.

Sosyal medyada Fenerbahçeliler EYT grubuna göre çok daha eski sosyal topluluk olmasına rağmen yapılan ölçüm sonucunda EYT'lilere ilişkin ağda ki bilgi akışının hızı(yoğunluk), modularite gibi değerlerin daha yüksek olduğu, kümelenme katsayısı, merkezileşme ve çap gibi değerlerinde birbirine çok yakın değerlere sahip olduğu tespit edilmiştir.

Sosyal medya yadsınamaz bir gerçek haline geldi. Bu analiz sosyal medya aracılığı ile tanıtım yapan kuruluşların doğru stratejiler geliştirmesi için kullanışlı olduğu gibi, Sosyal medyada ki başarılarını da bu sayede değerlendirmeye alabilirler. Twitter üzerinden yapılan bu çalışma Nodexl programı sayesinde Facebook, Instagram, Youtube vb. sosyal medya kanallarındaki ağlar içinde uygulanabilir ve analiz edilebilir.

REFERENCES

[1] The Social Media Research Foundation. Smr foundation. Ocak 28, 2019 tarihinde The Social Media Research Foundation: https://www.smrfoundation.org/ adresinden alındı.

[2] Altan, S. (2017, Mart 15). Pazarlamasyon. Ocak 24, 2019 tarihinde www.pazarlamasyon.com: https://pazarlamasyon.com/cristiano-ronaldonun-sosyal-medya-hesaplarindan-bir-yilda-500-milyon-dolar-gelir-elde-edildi/ adresinden alındı

[3] Barabási, A.-L. (2002). Linked: The New Science of Networks. UK: Perseus Publication.

[4] Bastos, M. (2014, Şubat 28). Hastac. Ocak 30, 2019 tarihinde www.hastac.org/: https://www.hastac.org/blogs/herrcafe/2014/02/28/self-loops-and-network-awareness adresinden alındı

[5] Bruns, A., & Stieglitz, S. (2013). Towards More Systematic Twitter Analysis: Metrics for Tweeting Activities. International Journal of Social Research Methodolog, s. 91-108.

[6] Denny, M. (2014, Eylül 26). Social Network Analysis. Unıversity of Massachusetts Amherst, Institute For Social Science Resarch, s. 1-20.

[7] Eduarda Mendes Rodrigues, N. M.-F. (2011). Group-in-a-box Layout for Multi-faceted Analysis of Communities. IEEE Third International Conference on Social Computing (s. 9-11). Boston: IEEE.

[8] Freeman, L. C. (2004). The Development of Social Network Analysis: A Study in The Sociology Of Science. Vancouver: Empirical Press.

[9] Gürsakal, N. (2009). Sosyal Ağ Analizi. Bursa: Dora Yayınları.

[10] Hansen, D., Shnejderman, B., & Smith, M. (2011). Analyzing Social Media Networks With Nodexl: Insights From A Connected World. Burlington MA: Morgan Kaufmann.

[11] Himelboim, I. E. (2009). Discussion Catalysts in Online Political Discussions: Content İmporters And Conversation Starters. Journal of Computer-Mediated Communication, 771-789.

[12] Himelboim, I., Smith, M., Rainie, L., Shneiderman, B., & Espina, C. (2017, Ocak-Mart). Classifying Twitter Topic-Networks Using Social Network Analysis. Social Media + Society, s. 1-13.

[13] John Scott, & Carrington, P. (2011). The SAGE Handbook of Social Network Analysis. London: Sage Publications Ltd.

[14] Kaiser, M., Görner, M., & Hilgetag, C. C. (2007, Mayıs 2). Criticality of Spreading Dynamics in Hierarchical cluster Networks without Inhibition. New Journal of Physics, s. 1-13.

[15] Karagöz, D., & Kozak, N. (2014). Anatolia Turizm Araştırmaları Dergisi'nin Bibliyometrik Analizi: Araştırma Konuları ve Kurumlar Arası İş Birliğinin Sosyal Ağ Analizi ile İncelenmesi. Türk Kütüphaneciliği, 47-61.

[16] Kemp, S. (2018, Ocak 30). DIGITAL IN 2018: WORLD'S INTERNET USERS PASS THE 4 BILLION MARK. Ocak 24, 2019 tarihinde https://wearesocial.com: https://wearesocial.com/blog/2018/01/global-digital-report-2018 adresinden alındı

[17] M. D. Conover, J. R. (2011). Political Polarization on Twitter. Proceedings of the Fifth International AAAI Conference on Weblogs and Social Media (s. 89-96). Bloomington: ICWSM.

[18] Marc Smith, L. R. (2014, Şubat 20). Mapping Twitter Topic Networks: From Polarized Crowds to Community Clusters. USA.

[19] Newman, M. E., & Girvan, M. (2004). Finding And Evaluating Community Structure in Networks. PHYSICAL REVIEW, 1-15.

[20] Rodrigez, M. G., Leskovec, J., & Krause, A. (2012). Inferring Networks of Diffusion and Influence. Transactions on Knowledge Discovery from Data, 21/1-36.

[21] Stelter, B. (2008, Temmuz 7). The New York Times. Ocak 24, 2018 tarihinde www.nytimes.com: https://www.nytimes.com/2008/07/07/technology/07hughes.html adresinden alındı

[22] Wasserman, S., & Faust, K. (1994). Social Network Analysis: Methods and Applications. Cambridge: Cambridge University Press.

[23] Woo-young, C., & Park, H. W. (2012). The Network Structure of the Korean Blogosphere. Journal of Computer-Mediated Communication, 216-230.

Analysis of Earthquake Awareness in Education By Data Mining

Ömer KIVRAK, Filiz ERSÖZ

Department of Industrial Engineering, Karabük University, TURKEY

Abstract: Turkey and the world a lot of earthquakes that have occurred and will continue to come. Due to the earthquake, many people have caused their lives and damage to their shelters. Starting from a young age earthquake in Turkey to create awareness in every age group, observed and loss of life and property damage in the earthquake is one of the lowest levels to minimize the road. Therefore, it is very important to be able to use techniques and systems that can analyze a large number of data sets. The process of converting these raw data into information or meaning can be done by data mining. The aim of this study is to investigate the effects of the education given to the secondary and high school students on the students and the earthquake awareness. The research sample consisted of 14 middle school and 11 high school students, who were randomly selected in Karabük and districts. The questionnaire developed by the researcher was used as the data collection tool. In the scope of the research, 1165 students were given questionnaires before and after 80 minutes of training. In this study, the seismic awareness of the students' education on the students before and after the training was investigated by clustering analysis.

Keywords: Earthquake, Earthquake Awareness, Data mining, Clustering analysis.

REFERENCES

[1] R. Kasap and Ü. Gürlen, "Deprem Magnitüdleri İçin Tekrarlanma Yıllarının Elde Edilmesi: Marmara Bölgesi Örneği," Doğuş Üniversitesi Dergisi, cilt 4, no. 2, pp. 157-166, 2003.

[2] S. Savaş, N. Topaloğlu and M. Yılmaz, "Veri Madenciliği ve Türkiye'deki Uygulama Örnekleri," İstanbul Ticaret Üniversitesi, Fen Bilimleri Dergisi, cilt 11, no. 21, pp. 1-23, 2012.

[3] O. İnan, Veri Madenciliği, KONYA: Selçuk Üniversitesi, Fen Bilimleri Enstitüsü, 2003.

[4] B. Thuarisingham, Web Data Mining And Applications in Business Intelligence And Counter Terrorism, Boca Raton, Fl, Usa, 2003.

[5] Z. Z. Bakır, "İnsani ve Sosyal Araştırma Merkezi," İnsani ve Sosyal Araştırma Merkezi, Http://İnsamer.Com/Tr/Dunyanin-Afet-Tablosu-Ve-Cozum-Onerileri_380.Html 2018.

[6] C. Geray, "Türkiye'de Yıkım (Afet) Olayları Karşısında Önlemler Vv Örgütlenmeler," Amme İdaresi Dergisi, Cilt 7, No. 2, Pp. 91-114, 1977.

[7] H. Yavaş, Doğal Afetler Yönüyle Türkiye'de Belediyelerde Kriz Yönetimi, Ankara: Orion Yayınevi, 2005.

[8] A. Yılmaz, Türk Kamu Yönetiminin Sorun Alanlarından Biri Olarak Afet Yönetimi, Ankara: Pegem Yayıncılık, 2005.

[9] Ö. Kıvrak And S. Kardeş, Kuzey Anadolu Fay Hattı Üzerinde Karabük İli ve Çevresinde Oluşan Depremlerin Veri Madenciliği Kullanılarak İncelenmesi, Karabük, 2014.

[10] F. Ersöz, Veri Madenciliği Teknikleri ve Uygulamaları, Seçkin Kitabevi, 2019.

[11] S. Piramuthu, "Evaluating Feature Selection Methods For Learning in Data Mining Applications," %1 İçinde Thiry-First Hawai International Conference On System Sciences, Kohala Coast Hawaii Usa, 1998.

[12] S. Özekes, "Veri Madenciliği Modelleri ve Uygulama Alanları" İstanbul Ticaret Üniversitesi Feb Bilimleri Dergisi, Cilt 2, No. 3, Pp. 65-82, 2003.

[13] S. Ilgaz, "Eğitimde Farkındalık," [Ttps://Www.Kucukdahi.Com/Blog/Egitimde-Farkindalik. 30 Nisan 2019.

[14] A. M. Bauer, J. A. Quas And W. T. Boyce, "Associations Between Physiological Reactivity And Children's Behavior: Advantages Of A Multisystem Approach," Journal Of Developmental And Behavioral Pediatrics, Cilt 23, P. 102–113, 2002.

[15] R. Mccraty, N. Atkinson, D. Tomasino, J. Goelitz And H. N. Mayrovitz, "The Impact Of An Emotional Self-Management Skills Course On Psychosocial Functioning And Autonomic Recovery To Stress in Middle School Children," Integrative Physiological And Behavioral Science, Cilt 34, No. 4, Pp. 246-268, 1999.

[16] Ö. Kıvrak, F. Ersöz and T. Ersöz, "Karabük İlinde Deprem Farkındalığına Yönelik Deprem Eğitiminin Öğrenciler Üzerindeki Etkisinin Araştırılması," 1. Uluslararası Afet Yönetimi Kongresi, Gümüşhane, 2018.

[17] F. Ersöz, T. Ersöz, Ö. Kıvrak and S. Kardeş, "Kuzey Anadolu Fay Hatti Üzerinde Yer Alan Karabük İli ve Çevresinde Meydana Gelen Depremlerin Veri Madenciliği ve İstatistiksel Yöntemlerle İncelenmesi," Uluslararası Sosyal Araştırmalar Dergisi, cilt 9, no. 47, pp. 353-363, 2016.

[18] F. Aydın, "İlköğretim Sekizinci Sınıf Öğrencilerinin 'Deprem' Algılamaları," Erzincan Üniversitesi. Fen Edebiyat Fakültesi, Türk Dili ve Edebiyatı Bölümü Dergisi, 2010.

[19] K. Ronan and D. Johnston, "Corraletes of Hazard Education Programs for Youth", Risk Analysis, cilt 21, no. 6, 2001.

[20] E. Esgin, Z. Baba, N. Aytaç and E. Turan, "İnternet Tabanlı Kaynakların Doğruluğu ve Güvenilirliği Hakkındaki Farkındalığın İncelenmesi," 5th International Computer & Instructional Technologies Symposium, Fırat University, ELAZIĞ, 2011.

[21] A. Demirci and S. Yıldırım, "İstanbul'da Ortaöğretim Öğrencilerinin Deprem Bilincinin Değerlendirilmesi," Milli Eğitim, cilt 1, no. 207, pp. 89-118, 2015.

[22] R. Shaw and M. Kobayashi, "Role of School in Creating Earthquake-Safer Environment," Disaster Menagement and Educational Facilities, pp. 1-7, 2001.

[23] J. Rogstadius, M. Vukovic, C. A. Teixeira, V. Kostakos, E. Karapanos and J. A. Laredo, "CrisisTracker: Crowdsourced Social Media Curation For Disaster Awareness," IEEE Xplore Digital Library, cilt 5, no. 4, 2013.

[24] K. Ozeki and T. Ojima, "Current Status of Disaster Preparedness of Pharmacies and Differences in Disaster Awareness Based on Pharmacy Size." Disaster Medicine and Public Health Preparedness (2019): 1-5.," Disaster Medicine and Public Health Preparedness, pp. 1-5, 2019.

[25] V. Clerveaux and B. Spence, "The Communication of Disaster Information And Knowledge To Children Using Game Technique: The Disaster Awareness Game (DAG)," International Journal Of Environmental Research, cilt 3, no. 2, pp. 209-222, 2009.

[26] K. Duong, "Disaster Education And Training Of Emergency Nurses in South Australia," Australasian Emergency Nursing Journal, cilt 12, no. 3, pp. 86-92, 2009.

[27] D. Şchiopu, "Applying TwoStep Cluster Analysis for Identifying Bank Customers' Profile," Universittatii Petrol – Gaze din Ploieşti Ştiinte Economice, cilt LXII, no. 3, pp. 66-75, 2010.

[28] J. Han, M. Kamber and J. Pei, "Data Mining Concepts and Techniques," Morgan Kaufmann Publishers Inc, 2012.

[29] İ. Ertuğrul, A. Organ and A. Şavlı, "Veri Madenciliği Uygulamasına İlişkin PAÜ Hastanesinde Hasta Profilinin Belirlenmesi," Pamukkale Üniversitesi Mühendislik Bilimleri Dergisi, cilt 19, no. 2, pp. 97-103, 2013.

[30] S. Giray and F. Esin Gülel, "Avrupa Ülkelerinin İntihar Oranlarına Göre Sınıflandırılması," Süleyman Demirel Üniversitesi Fen-Edebiyat Fakültesi Sosyal Bilimler Dergisi, cilt 31, pp. 235-247, 2014.

[31] R. Blashfield and M. Aldenferder, "The Literature on Cluster Analysis," Multivariate Behavioral Research, cilt 13, pp. 271-295, 1978.

[32] A. Uslu, C. Çetinkaya, E. Özceylan and S. K. İşleyen, "Organize Sanayi Bölgelerinin Hiyerarşik-K-Ortalamalar Yöntemi ile Analizi," Türk Sosyal Bilimler Araştırmaları Dergisi, cilt 2, no. 1, pp. 20-37, 2017.

[33] R. Johnson and D. Wichern, Applied Multivariate Statistical Analysis, New Jersey: Prentice-Hall, 2002.

[34] S. Rundle-Thiele, K. Kubacki, A. Tkaczynski and J. Parkinson, "Using Two-Step Cluster Analysis to Identify Homogeneous Physical Activity Groups," Marketing Intelligence & Planning, cilt 33, no. 4, pp. 522-537, 2015.

[35] A. Tkaczynski, "Segmentation Using Two-Step Cluster Analysis," Segmentation in Social Marketing, p. 109–125, 2017.

[36] S. Savaş and N. Topaloğlu, "Veri Madenciliği Yöntemi ile GSM Şebekelerinin Performans Analizi," Gazi Üniversitesi Mühendislik Mimarlık Fakültesi Dergisi, Cilt 26, No. 4, pp. 741-751, 2011.

[37] Z. Ceylan, S. Gürsev and S. Bbulkan, "İki Aşamalı Kümeleme Analizi ile Bireysel Emeklilik Sektöründe Müşteri Profilinin Değerlendirilmesi," Bilişim Teknolojileri Dergisi, cilt 10, no. 4, pp. 475-485, 2017.

[38] R. Alpar, Uygulamalı Çok Değişkenli İstatistiksel Yöntemler, Ankara: Detay Yayıncılık, 2017.

Tersanelerde Yalın 6 Sigma ve Uygulanabilirliği

Emre GÜVEN, Ufuk CEBECİ

Department of Industrial Engineering, Istanbul Technical University, TURKEY

Abstract: Teknoloji kavramının insanoğlunun hayatına girmesiyle teknolojinin kullanım alanları gün geçtikçe artmıştır. Gelişen teknik beceriler ile üretimin odağı değişmiştir. Zamanla üretimin verimliliği tartışılmaya başlamıştır. Bu tartışmalar sonucunda endüstriyel hayata kalite, optimizasyon, verim gibi yeni kavramlar girmiştir. Üretim süreçleri üzerine yapılan araştırmalarda daha kısa sürede, daha az maliyetli ve daha kaliteli üretim için yeni fikirler ortaya çıkmıştır. İkinci Dünya Savaşı'nın her alanda olduğu gibi üretim yöntemleri üzerinde de etkisi olmuştur. Artan teknoloji arayışları yeni üretim ve kalite anlayışlarına olan ihtiyacı doğurmuştur. Bu çalışmada yalın üretim ve altı sigma kavramları incelenmiş, yöntemler gösterilmiş ve seçilen bir tersanede uygulaması yapılmıştır. Yalın üretim felsefesi, üretim süreci boyunca oluşan artık ve katma değersiz durumları elimine etme üzerine oturtulmuştur. Altı sigma metodoloji ise üretim çıktısındaki hataların azaltılmasını ve üretimin standartlaştırarak hızlanmasını hedeflemektedir. Yapılan uygulama sonucunda teorik olarak süreçlerdeki kayıpların büyük oranda azaldığı ve firmanın iyileştirme çalışmaları doğrultusunda kazanç sağlayacağı görülmüştür.

Keywords: Tersane, Gemi İnşa, Üretim, Yalın, 6 Sigma, Kalite

REFERENCES

[1] Arunagiri, P., & Gnanavelbabu, A. (2014). Identification of Major Lean Production Waste in Automobile Industries Using Weighted Average Method. Procedia Engineering.

[2] Bulut, S. (2012). Beyaz Eşya Yan Sanayi Sektöründe ERP ve Yalın Üretim Olgunluğu Analizi ve Otomotiv Yan Sanayi ile Kıyaslama. Doktora Tezi, İTÜ Fen BİlimleriEnstitüsü.

[3] Doğan, S., & Demiral, Ö. (2008). Yalın Yöntemler ve Altı Sigmayı İçeren Bütünleşik Bir Yaklaşım: Yalın Altı Sigma. Atatürk Üniversitesi İktisadi ve İdari Bilimler Dergisi, 22(1), 343-366.

[4] Eaton, M. (2013). The Lean Practitioner's Handbook.

[5] Ergün, A. K. (2003). Altı Sigma Metodolojisi ve Türkiye'deki Uygulamaları (Doctoral dissertation, Fen Bilimleri Enstitüsü).

[6] Firuzan, A. R., Kuvvetli, Ü., & Gerger, A. (2012). Altı Sigma Metodolojisi ve Otomotiv Sektöründe Bir Uygulama. Journal of Yasar University, 25(7), 4176-4188.

[7] Gecü, B. (2008). İç Lojistik Sistemlerinin Yalın Üretim Bakış Açısıyla Yeniden Tasarlanması ve Otomotiv Sektöründe Örnek Bir Uygulama. Yüksek Lisans Tezi, İTÜ Fen Bilimleri Enstitüsü.

[8] Küçükuysal, K. (2012). Otomotiv Yan Sanayi Sektöründe ERP ve Yalın Üretim Analizi. Doktora Tezi, Fen BİIimleri Enstitüsü, İstanbul Teknik Üniversitesi.

[9] Kaşıkçı, M.K., Cebeci, U. (2018). A Manufacturing Maturity Analysis in The Shipbuilding Industry, ICONASH 2018 İstanbul, Second International Conference on New Approaches in Social Sciences and Humanities,265-268

[10] Siddh, M. M., Soni, G., Gadikar, G., & Jain, R. (2014). Integrating Lean Six Sigma and Supply Chain Approach for Quality and Business Pperformance. 2nd International Conference on Business and Information Management (ICBIM).

[11] Töre, Ö. G. (2000). Yalın Üretim ve Otomotiv Sanayinde Kullanılması. İstanbul Teknik Üniversitesi, Yüksek Lisans Tezi.

[12] Uçan, K. (2014). Otomotiv Yan Sanayisinde Malzeme Besleme Sisteminin Yalın Üretim Yaklaşımıyla Yeniden Tasarlanması ve Bir Uygulama. Yüksek Lisans Tezi, Fen Bilimleri Enstitüsü.

[13] Wahab, A. N., Mukhtar, M., & Sulaiman, R. (2013). A Conceptual Model of Lean Manufacturing Dimensions. Procedia Technology.

Smart and Green Supply Chain Applications in Enterprises

Deniz MERDİN, Filiz ERSÖZ

Department of Industrial Engineering, Karabük University, TURKEY

Abstract: Increasing the flexibility and effecient of enterprises from procurement to sales provides a great competitive advantage for meeting consumer demands. Providing competitive advantage is possible through the effective implementation of innovative technologies of fourth industrial revolution in the all stages of supply chain process. In this context, the technologies related to industry 4.0 were mentioned in the study and the differences between the traditional supply chain and the digital supply chain were determined. In addition, the industry 4.0 applications in the digital and green supply chain are mentioned and the steps that must be followed in the process of transition to the digital supply chain are indicated.

Keywords: Industry 4.0, IoT, Dgital Supply Chain, Green Supply

REFERENCES

[1] B. P. Santos, F. Charrua-Santos, and T. M. Lima, "Industry 4.0: an overview," 2018, no. July.

[2] C. Cimini, G. Pezzotta, R. Pinto, and Sergio Cavalieri, "Service Orientation İn Holonic and Multi-Agent Manufacturing," Stud. Comput. Intell., vol. 640, no. December 2018, 2016.

[3] L. S. Dalenogare, G. B. Benitez, N. F. Ayala, and A. G. Frank, "The Expected Contribution of Industry 4.0 Technologies For İndustrial Performance," Int. J. Prod. Econ., vol. 204, no. August, pp. 383–394, 2018.

[4] J. Smit, "Directorate General For Internal Policies Policy Department A: Economic And Scientific Policy Industry 4.0 Study," 2016.

[5] S. Schrauf and P. Berttram, "Industry 4.0: How Digitization Makes The Supply Chain More Efficient, Agile, And Customerfocused," Strateg. &Technology, p. pg. 1-32, 2016.

[6] R. Burke, A. Mussomeli, S. Laaper, M. Hartigan, and B. Sniderman, "The Smart Factory Responsive, Adaptive, Connected Manufacturing," 2017.

[7] L. Ardito, A. M. Petruzzelli, U. Panniello, and A. C. Garavelli, "Towards Industry 4.0 Mapping Digital Technologies For Supply Chain Management-Marketing İntegration," Bus. Process Manag. J., vol. 29, no. 5, pp. 323–347, 2009.

[8] E. Manavalan and K. Jayakrishna, "A Review of Internet Of Things (Iot) Embedded Sustainable Supply Chain For Industry 4.0 Requirements," Comput. Ind. Eng., vol. 127, no. November 2017, pp. 925–953, 2019.

[9] D. Ivanov, A. Dolgui, and B. Sokolov, "The Impact of Digital Technology And Industry 4.0 On The Ripple Effect And Supply Chain Risk Analytics," Int. J. Prod. Res., vol. 57, no. 3, pp. 829–846, 2019.

[10] J. C. Bendul and H. Blunck, "The Design Space of Production Planning And Control For İndustry 4.0," Comput. Ind., vol. 105, pp. 260–272, 2019.

[11] A. Kattepur, A. Mukherjee, and P. Balamuralidhar, "Verification and Timing Analysis of Industry 4.0 Warehouse Automation Workflows," IEEE Int. Conf. Emerg. Technol. Fact. Autom. ETFA, vol. 2018-September, pp. 1297–1304, 2018.

[12] B. Surajit and A. Telukdarie, "Business Logistics Optimization Using Industry 4.0: Current Status and Opportunities," IEEE Int. Conf. Ind. Eng. Eng. Manag., vol. 2019-December, pp. 1558–1562, 2019.

[13] A. Yıldız, E. U. Ergül, C. Gezegin, and H. Dirik, "Akıllı Depolar için PLC Ünitelerinin Bulanık TOPSIS Yöntemiyle Değerlendirilmesi," in III. Uluslararasi Mesleki ve Teknik Bilimler Kongresi, 2018, October, pp. 65–71.

[14] H. Aydemir, "Sanayi 4.0 ve Türkiye Ekonomisi Açısından Etkileri," Sosyoekonomi, vol. 26, no. 36, pp. 253–261, 2018.

[15] M. H. Calp, E. Bahçekapılı, and M. Berigel, "Endüstri 4.0 Kapsamında Akıllı Fabrikaların İncelenmesi," in 5th International Management Information Systems Conference - MISC 2018, 2018, pp. 16–20.

[16] B. N. Doğru and O. Meçİk, "The Effects Of Industry 4.0 on The Labor Market In Turkey: Firm Predictions," Suleyman Demirel Univ. J. Fac. Econ. Adm. Sci., vol. 23, pp. 1581–1606, 2018.

[17] M. M. Erdoğdu, "Türkiye'nin Dördüncü Sanayi Devrimi Sürecindeki Konumu," In Ii. International Symposium On Economics, Politics And Administration - ISEPA'18, 2018, no. December, pp. 1–10.

[18] K. Bär, Z. N. L. Herbert-Hansen, and W. Khalid, "Considering Industry 4.0 Aspects İn The Supply Chain For An SME," Prod. Eng., vol. 12, no. 6, pp. 747–758, 2018.

[19] D. Ivanov, S. Sethi, A. Dolgui, and B. Sokolov, "A Survey On Control Theory Applications To Operational Systems, Supply Chain Management, And Industry 4.0," Annu. Rev. Control, vol. 46, pp. 134–147, 2018.

[20] B. Wang and J. E. Ha-Brookshire, "Exploration of Digital Competency Requirements within the Fashion Supply Chain with an Anticipation of Industry 4.0," Int. J. Fash. Des. Technol. Educ., vol. 11, no. 3, pp. 333–342, 2018.

[21] A. Phase and N. Mhetre, "Using IoT in Supply Chain Management," Int. J. Eng. Tech., vol. 4, no. 2, pp. 2395–1303, 2018.

[22] M. Shafique, A. Rashid, I. Bajwa, R. Kazmi, M. Khurshid, and W. Tahir, "Effect of IoT Capabilities and Energy Consumption behavior on Green Supply Chain Integration," Appl. Sci., vol. 8, no. 12, p. 2481, 2018.

[23] V. L. da Silva, J. L. Kovaleski, and R. N. Pagani, "Technology Transfer in the Supply Chain Oriented to Industry 4.0: A Literature Review," Technol. Anal. Strateg. Manag., vol. 31, no. 5, pp. 546–562, 2018.

[24] S. Luthra and S. K. Mangla, "Evaluating Challenges to Industry 4.0 Initiatives For Supply Chain Sustainability in Emerging Economies," Process Saf. Environ. Prot., vol. 117, pp. 168–179, 2018.

[25] M. M. Erdogdu and S. Akar, "Dördüncü Sanayi Devrimi Döneminde Sürücüsüz Otonom Araçların Potansiyelleri ve 7 Geleceği: Türkiye Örneği," in Current Debates In Tourism & Development Studies, December 2017, pp. 275–298.

[26] R. Accorsi, M. Bortolini, G. Baruffaldi, F. Pilati, and E. Ferrari, "Internet-of-things Paradigm in Food Supply Chains Control and Management," Procedia Manuf., Vol. 11, June, pp. 889–895, 2017.

[27] O. Elmacı, "How Is Provided Sustainable Competitive Advantage In Business? As The Strategic Planning Tool Bsc Model Proposal in Perspective Industry 4.0," in 3rd International Congress on Social Sciences, China to Ardiatic, 2018, no. October 2016, pp. 349–370.

[28] M. Maslarić, S. Nikoličić, and D. Mirčetić, "Logistics Response to the Industry 4.0: the Physical Internet," Open Eng., vol. 6, no. 1, pp. 511–517, 2016.

Bitcoin Price Prediction by Using Artificial Neural Networks and Time Series

Derya SAĞ, Ufuk CEBECİ

Department of Industrial Engineering, Istanbul Technical University, TURKEY

Abstract: Bitcoin is the most popular cryptocurrency in the market. Satoshi Nakamoto has created the Bitcoin in 2009. Bitcoin movements are a widely discussed topic nowadays. On the other hand, Today machine learning is effectively deployed in an extensive variety of fields from natural language processing to image processing, from medical applications to activity recognition. In this research, we collected and used data about the market value of Bitcoin. Also, Bitcoin is briefly described in this study. Artificial neural networks and time series techniques used to estimate the market value of Bitcoin. Finally, The paper concludes with critical considerations of recent developments and some recommendations for future researches.

Keywords: Machine Learning, Big Data, Data Mining, Bitcoin

Introduction

There are a lot of coin in the cryptocurrency in the market. Bitcoin is still the leader of market. The first transaction actualize between Satoshi Nakamoto and Hal Finney ın 2009. In this research, Bitcoin price forecasting was made by using multilayer perceptron and time series. Sean McNally; Jason Roche ; Simon Caton (2018) are published an article about the prediction of Bitcoin with using RNN and LSTM. The LSTM algorithm achieves the highest classification accuracy [1]. Isaac Madan, Shaurya Saluja, Aojia Zhao (2015) are published an article at Stanford University. Their data set consists of over 25 features relating to the Bitcoin price and payment network over the course of five years, recorded daily. Using this information they were able to predict the sign of the daily price change with an accuracy of 98.7% [2]. Alex Greaves, Benjamin Au (2015) are published an article that performed to analyze the network's influence on overall Bitcoin price. In this paper, They investigated the predictive power of blockchain network-based features on the future price of Bitcoin. they obtained up-down Bitcoin price

movement classification accuracy of roughly 55% [3]. Huisu Jang ; Jaewook Lee (2017) are published an article that An Empirical Study on Modeling and Prediction of Bitcoin Prices With Bayesian Neural Networks Based on Blockchain Information. They measured the effect of BNNs by analyzing the time series of Bitcoin process [4].

Ifigeneia Georgoula(2015) is published an article that title is Using Time-Series and Sentiment Analysis to Detect the Determinants of Bitcoin Prices [5].

Blockchain

Blockchain is a distributed data logging system that provides encrypted transaction tracking. It is not a database because saved data cannot be changed or deleted again. Blockchain allows us to keep records that cannot be changed and manipulated. And what makes this technology so great is that it doesn't need a central authority.

We can summarize this process in 3 main steps.

1. Creation of Transaction

2. Confirmation of Transaction

First, check whether the Bitcoins referenced in the transaction are used before. Secondly, it is checked whether the signature in the transaction is correct. This is found by putting the sender's open address, transaction and signature in a function. If this signature returns to true, this transaction is put into the approved transaction pool. The next step is to add this transaction to the chain.

3. The addition of the block to the chain.

You can think of what we call a block as a text file. This text file contains the block number, the proof of work (POW) number, the proof of work number of the previous block, and finally the approved transactions.

I. Bitcoin

We can say that Bitcoin is simply a non-state currency behind it. The Bitcoin network consists of client computers that are actually connected to this system. Every computer involved in the network communicates with other clients that are close to it, and Peer-to-Peer starts to exchange

data in a way that automatically starts downloading all of the operations that have been done in the network since 2009, so they begin to synchronize with the network. The list of all transactions made throughout history is open to everyone in a transparent manner. It is even possible that you can see some of the web sites currently performing live.

II. Weka

In this reserach, We used to Weka for Bitcoin price prediction. WEKA is one of the packages used in machine language which is one of the important subjects of computer science. Developed on the JAVA language as open source at Waikato University and distributed under GPL license. The name comes from here and consists of the initials of the words Waikato Environment for Knowledge Analysis. We downloaded the deep learning and time series packages for Bitcoin price prediction.

Multilayer Perceptron

There are tree layer.

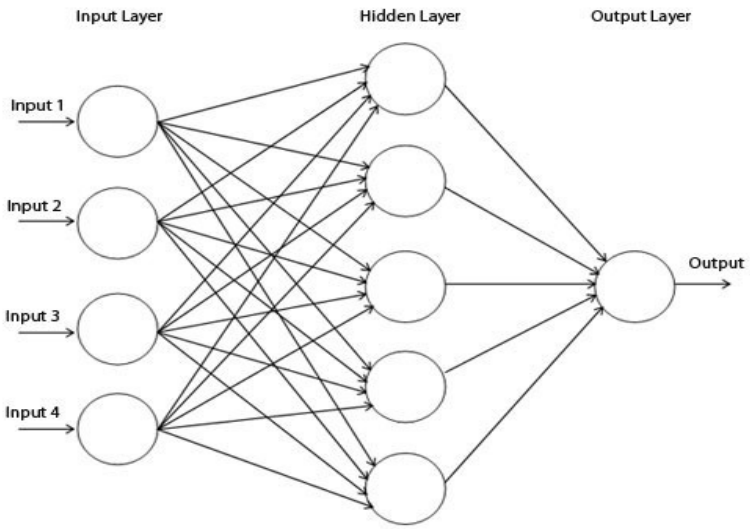


Fig 1. Multilayer Perceptron Graph.

Long Short Term Memory (LSTM)

III. Loss Functions

Mean squared error	$\mathrm{MSE} = \frac{1}{n}\sum_{t=1}^{n} e_t^2$
Root mean squared error	$\mathrm{RMSE} = \sqrt{\frac{1}{n}\sum_{t=1}^{n} e_t^2}$
Mean absolute error	$\mathrm{MAE} = \frac{1}{n}\sum_{t=1}^{n} \lvert e_t \rvert$
Mean absolute percentage error	$\mathrm{MAPE} = \frac{100\%}{n}\sum_{t=1}^{n} \left\lvert \frac{e_t}{y_t} \right\rvert$

Fig 2. Types of Lost Functions.

Activation Functions

The purpose of activation function is to introduce non- linearity into the network. Non-linearity allow us to approximate arbitrarily complex functions.

The sigmoid (logistic) function is a very common choice for feed-forward NNs that need to output only positive values. Although its widespread use, the hyperbolic tangent or ReLU function are generally more convenient. The values of the function are limited to 0 to 1.

One of the most important activation functions is the hyperbolic tangent function. The values of the function are limited to -1 to 1. Its shape is similar to the sigmoid function. There are some advantages over sigmoid function. These include derivatives used in the training of the NN.

Because of the fact that ReLU is a linear, non-saturating function, It has some advantages over other activations functions. On the contrary of the sigmoid or hyperbolic tangent activation functions, ReLU does not saturate to -1, 0 or 1.

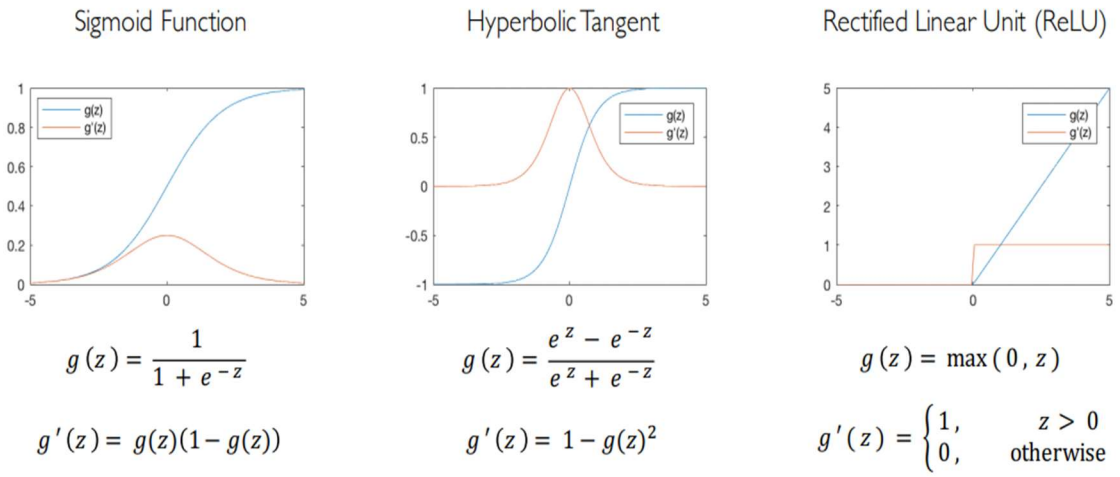


Fig 3. Sigmoid, Hyperbolic Tangent, ReLU and their derivative graphics.

Other Parameters

Stochastic Gradient Descent: Because of the fact that BGD is a slow algorithm, we prefer to use stochastic gradient descent for faster calculation. Because of the fact that GD method only taking a single step for one pass over the training set, it could be a very costly method for a large data set. SGD never converges as BGD does, it is moving around to some close area global minimum.**[31]**

$$\phi = \phi - \eta \, \partial \frac{J(\phi; x^{(i)}; y^{(i)})}{\partial \phi} \qquad (1)$$

Learning Rate: Small learning rate converges slowly and gets stuck in false local minima. Large learning rates can converge more quickly but large learning rates overshoot, become unstable and diverge. This can cause J to increase, rather than decrease monotonically. Instead of large or small learning, we can use adaptive learning rates. The optimal solution for the learning speed is to initially keep the learning rate high and then gradually decrease. When the learning speed is too small at the beginning, it can be attached to the local optimum value, causing the global optimum value not to be reached at all. The learning rate value is generally used as the default value of 0.01, which is reduced to 0.001 after a certain epoch.

Momentum: Momentum is a method that aid to accelerate SGD in the relevant direction and decreased oscillations as can be seen in the figure shown below.

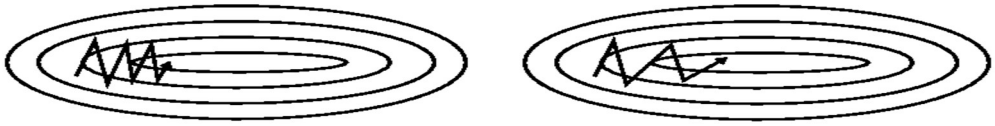


Fig 4. SGD without momentum SGD with momentum.

$$v_t = \gamma v_{t-1} + \mu \nabla_{\phi} J(\phi) \quad (2)$$

γ parameter get value between from 0 to 1. It generally get 0.9 and it calculates how much of the previous gradients being into the calculation.

Adam: Adaptive Moment Estimation is another approach that calculates adaptive learning rates for every parameters.

Application Results

Table 1. Parameters and Their Values For Application.

Parameters	**Value**
Batch-size	64
Learning rate	0.001
Momentum	0.9
Number of epoch	100
Optimization algorithm	Stochastic Gradient Descent
Weight initialization method	XAIVER
Updater for SGD	Adam
Initial weight distribution	Normal distribution

Multi-layer perceptron algorithm prediction are listed in the table for 2 weeks.

Table 2. Prediction And Actual Value Of Bitcoin Price For 2 Weeks.

Date	Prediction of Bitcoin Value	Actual Bitcoin Value
04-03-2018	11360.3894	11.512,60
05-03-2018	11196.8307	11.704,10
06-03-2018	10995.2038	11.500,10
07-03-2018	10794.2281	10.929,50
08-03-2018	10754.9553	10.147,40
09-03-2018	10717.0087	9.466,35
10-03-2018	10544.1645	9.531,32
11-03-2018	10458.7118	9.711,89
12-03-2018	10210.3776	9.937,50
13-03-2018	10074.329	9.470,38
14-03-2018	9914.8471	9.355,85
15-03-2018	9911.4377	8.428,35
16-03-2018	9976.0603	8.585,15
17-03-2018	9862.0574	8.346,53

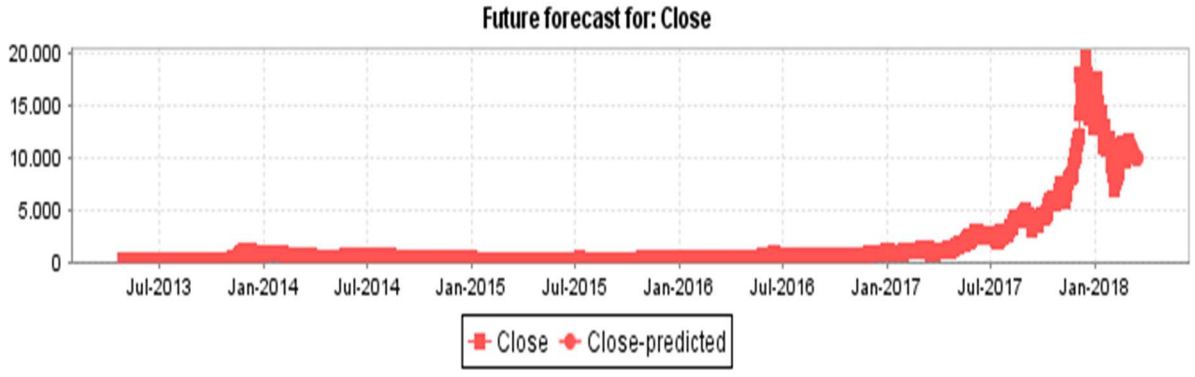


Fig 5. Closing Price of Bitcoin Graph.

Conclusion

In this research, we used 1772 daily closing price of Bitcoin between from 27/04/2013 to 03/03/2018. According to results, Multilayer perceptron has better results than the others.

REFERENCES

[1] S., J., & S. (2018, June 07). Predicting the Price of Bitcoin Using Machine Learning. Retrieved April 12, 2019, from https://ieeexplore.ieee.org/stamp/stamp.jsp?tp=&arnumber=8374483

[2] Madan, I., Saluja, S., & Zhao, A. (2014). Automated Bitcoin Trading via Machine Learning Algorithms.

[3] Greaves, A.& Au,. (2015, December 8). Using the Bitcoin Transaction Graph to Predict the Price of Bitcoin.

[4] Jang, H., & Lee, J. (2017, December 4). An Empirical Study on Modeling and Prediction of Bitcoin Prices With Bayesian Neural Networks Based on Blockchain Information. IEEE Xplore, 6, 5427-5437. doi:10.1109/ACCESS.2017.2779181

[5] Georgoula, Ifigeneia; Pournarakis, Demitrios; Bilanakos, Christos; Sotiropoulos, Dionisios N.; and Giaglis, George M., "Using TimeSeries and Sentiment Analysis to Detect the Determinants of Bitcoin Prices" (2015). MCIS 2015 Proceedings. 20.

[6] Niu, F., Recht, B., Christopher, R., & Wright, S. J. (2011). Hogwild! : A Lock-Free Approach to Parallelizing Stochastic Gradient Descent, 1–22.

Analysis of Presence of Bank Branches According to Settlement in Turkey with Data Mining

Süleyman DÜNDAR, Seliha Seçil BAYRAM

Department of Business Administration, Karabük University, TURKEY

Abstract: Banks are one of the most important actors of the financial system. Banks can perform their services through alternative distribution channels such as branches, ATMs, internet banking and telephone banking. Banks, which are affected by economic developments and who are both financial market actors and profit-making enterprises and employing employment, provide the most efficient services through their branches. The choice of the banks' location is very important in terms of bank success. One of the factors affecting the choice of establishment location of banks is the population of the place. The closest selection criteria are the GDP per capita in the region and the activities of competing banks. It is important to choose the place of operation of the banks as well as the widespread branch network in which the customers will work. In the competitive environment, especially corporate customers work with more than one bank. It is important that the banks where they work for the enterprises that have active or widespread networks in the field have branches together in the same settlement. The aim of this study is; a total of 960 settlements with Turkey's 81 provinces, the districts and towns, bank examination for the presence of the branches of activity and which banks that operate together on the same settlement "data mining" One of the suitable ones association rules "Apriori Algorithm" and is to be determined. For this purpose, the "The Banks Association of Turkey" in the number of audits and the existence of settlements in which they operate according to their location by the banks operating in Turkey "IBM SPSS Modeler" has been analyzed with the program. According to the results obtained, the banks operating together in the same settlement area may determine the bank preferences of the financial managers.

Keywords: Bank Branches, Data Mining, Association Rules, Apriori Algorithm.

REFERENCES

[1] Türkiye Bankalar Birliği https://www.tbb.org.tr/tr/bankacilik/banka-ve-sektor-bilgileri/banka-bilgileri/subeler/65 (erişim tarihi: 01.02.2019)

[2] Türkiye Kalkınma Bankaları Birliği http://www.tkbb.org.tr/banka-genel-bilgileri (erişim tarihi: 01.04.2019) [3] BDDK, (14.02.2019), TBS Temel Göstergeleri Raporu Aralık 2018,

[3] https://www.bddk.org.tr/ContentBddk/dokuman/veri_0014_39.pdf

[4] Yetiz, Filiz, "Bankacılığın Doğuşu ve Türk Bankacılık Sistemi", Niğde Üniversitesi İktisadi ve İdari Bilimler Fakültesi Dergisi, Sayı: 9(2), Nisan 2016.

[5] M.T. Kartal ve S.K. Depren, "Bankacılıkta Şubeleşme Eğilimini Etkileyen Makroekonomik Faktörlerin Belirlenmesi Türk Mevduat Bankaları Üzerine Bir Araştırma - Determination of Affecting Macroeconomic Factors of Branching Trend in Banking," Turkish Studies International Periodical for the Languages, Literature and History of Turkish or Turkic Volume 12/24, p. 97-120, Ekim 2017.

[6] N.T. Çınar, Kuruluş Yeri Seçiminde Bulanık TOPSIS Yöntemi ve Bankacılık Sektöründe Bir Uygulama, KMÜ Sosyal ve Ekonomik Araştırmalar Dergisi, 12(18), 37-45,2010.

[7] A. Başar, Ö. Kabak ve Y. İ. Topçu, "Banka Şubeleri İçin Uygun Yer Seçiminin Belirlenmesine Yönelik Tabu Arama Yaklaşımı: Bir Türk Bankası Uygulaması - Tabu Search Approach For Locating Bank Branches: An Application For A Turkish Bank," Journal Of Industrial Engineering (Turkish Chamber Of Mechanical Engineers), Vol. 26 Issue 3, P2-22. 21p, Temmuz 2015. [8] F. Ersöz, "Veri Madenciliği Teknikleri ve Uygulamaları," 72 Tasarım Dijital Basım ve Yayınevi, Basım sayısı:2, Sayfa sayısı:341, ISBN:978-605-65464-0-2, 2017.

[9] Y.L. Chen, J.M. Chen ve C.W. Tung, "A Data Mining Approach For Retail Knowledge Discovery With Consideration of the Effect of Shelf-Space Adjacency on Sales," Decisions Support Systems, cilt: 42, sayı: 3, "1503-1520, 2006.

[10] R. Agrawal ve R.Srikant, "Fast Algorithms for Mining Association Rules in Large Databases," Proceedings of the 20th International Conference on Very Large Databases (VLDB), Santiago, 487-499,1994.

[11] J. Han ve M. Kamber, "Data Mining: Concepts and Techniques," Morgan Kaufmann Publishers, San Francisco, 2001.

Artificial Bee Colony Algorithm for Container Loading Problem

Tuğrul BAYRAKTAR[1], Filiz ERSÖZ[1], Cemalettin KUBAT[2]

[1]Department of Industrial Engineering, Karabük University,
[2]Department of Industrial Engineering, Sakarya University, TURKEY

Abstract: A container is the one of the main components of transportation systems. Allocating items into limited spaces, is a kind of combinatorial optimization problem and container loading problems is a branch of knapsack problems, in which a set of items are loaded into capacitated domains. Heuristic approaches are mostly applied to solve knapsack problems due to the problem complexity. Artificial Bee Colony Algorithm and Genetic Algorithm are successful for solving object placement issues. The first one can obtain sufficient results as well as the second one. In this study, the performance of Artificial Bee Colony Algorithm, which is applied on CLP rarely, is compared with Genetic Algorithm, which is applied on CLP widely, to see the capability of proposed ABC algorithm for further studies.

Keywords: Artificial Bee Colony Algorithm, Genetic Algorithm, Container Loading Problem.

REFERENCES

[1] Lee, K.Y., and El-Sharkawi, M.A. "Modern Heuristic Optimization Techniques", Wiley-Interscience, 2008.

[2] Che, C. H., Huang, W., Lim, A., and Zhu, W., "The Multiple Container Loading Cost Minimization Problem.", European Journal Of Operational Research, vol. 214(3), pp. 501–511, 2011.

[3] Arora, J. S., Introduction to Optimum Design, 3rd ed., 2012, pp. 643-655.

[4] McCall, J. "Genetic Algorithms For Modelling And Optimization", Journal of Computational and Applied Mathematics, vol. 184(1), pp. 205-222, 2005.

[5] Goncalves, J.F., and Resende, M.G.C., "A Biased Random Key Genetic Algorithm For 2D and 3D Bin Packing Problems.", International Journal Of Production Economics, vol. 145(2), pp. 500-510, 2013.

[6] Karaboga, D., "An Idea Based On Honey Bee Swarm For Numerical Optimization", Erciyes University Engineering Faculty Computer Engineering Department Technical Report-TR06, 2005.

[7] Pham, D.T., Ghanbarzadeh, A., Koc, E., Otri, S., Rahim, S., and Zaidi, M., "The Bees Algorithm - A Novel Tool for Complex Optimization Problems", Technical Note. Manufacturing Engineering Centre, Cardiff University, UK, 2006.

[8] Karaboga, D., and Akay, B., "A Comparative Study Of Artificial Bee Colony Algorithm", Applied Mathematics and Computation, vol. 214, pp. 108-132, 2009.

Image Size Scaling and Feature Transformation Function Application for Image Processing in Machine Learning

Ömer PİŞGİN, Ali Osman ÇIBIKDİKEN

Department of Computer Engineering, Necmettin Erbakan University, TURKEY

Abstract: With the increase in computational power and big data, studies on artificial intelligence are increasing day by day. Especially deep learning applications are seen in almost all areas of our lives. The most successful results of deep learning architectures are in image processing. Different architectural approaches are tried to make image processing fast. Due to the fact that video images consist of large capacity data, it is very important to achieve high performance in these video images. In this study, size reduction function has been proposed that can reduce the size of the high-quality and large-capacity file data and produce results with a high accuracy rate. The results of the proposed method were compared in terms of performance and speed with different architectures in image processing using CNN (Convolutional Neural Network) algorithm. In addition, an application that uses the recommended size reduction function has also been developed using the Python programming language.

REFERENCES

[1] V. L. Quoc, J. Ngiam, A. Coates, A. Lahiri, B. Prochnow, A. Y. Ng, "On Optimization Methods for Deep Learning", Proceedings of the 28th International Conference on Machine Learning (ICML-11). 2011. 265-272, 2011.

[2] M. Courbariaux, Y. Bengio, J. P. David, "Binary Connect: Training Deep Neural Networks With Binary Weights During Propagations", NIPS 28, 2015.

[3] A. Calderón, S. Roa, J. Victorino, "Handwritten Digit Recognition using Convolutional Neural Networks and Gabor filters", Proc. Int. Congr. Comput. Intell., 2013.

[4] P. Yadav, N. Yadav, "Handwriting Recognition System-A Review", Analysis, Vol. 114, pp. 36-40, 2015.

[5] Y. LeCun, L. Bottou, Y. Bengio, and P. Haffner, "Gradient-Based Learning Applied to Document Recognition", Proceedings of the IEEE, 86(11):2278-2324, 1998.

Determination of Weights With Fuzzy AHP in the Job Evaluation Process

Ataberk OLCAY[1], Muharrem DÜGENCİ[1], Mümtaz İPEK[2]

[1]Department of Industrial Engineering, Karabük University
[2]Department of Industrial Engineering, Sakarya University, TURKEY

Abstract: Job evaluation; In order to create input for performance evaluation and wage management in businesses, the skills of job, responsibility, job conditions etc. and formally and systematically. Business valuation deals with the importance of the work and the added value it provides to the business rather than the employee. This study was carried out in the iron and steel industry. Firstly, two-way comparison matrices were formed by Analytical Hierarchy Process which is one of the Multi Criteria Decision Making (MCDM) methods for the main factors and sub-criteria to be used when conducting business evaluation. Then, the paired comparison matrices formed by Fuzzy Analytic Hierarchy Process were reconstructed with triangular fuzzy numbers and the weights of main factors and sub-criteria were calculated by Chang Ground Analysis method.

Keywords: Job Evaluation, Multi Criteria Decision Making, Analytical Hierarchy Process, Fuzzy

REFERENCES

[1] İ. Özdaban, "İş Değerlendirme ve Personel Değerlendirme Üzerine Bir Bulanık Karar Modeli," İstanbul Teknik Üniversitesi, Fen Bilimleri Üniversitesi, İstanbul, 2010.

[2] A. Alkan, "İş Değerlendirme Sürecinde Bulanık Analitik Hiyerarşi Prosesi ile Bütünleşik Bulanık Topsis Uygulaması," Erciyes Üniversitesi, Fen Bilimleri Enstitüsü, Kayseri, 2012.

[3] A. Spyridakos, Y. Siskos, D. Yannacopoulos ve A. Skouris, "Multicriteria Job Evaluation For Large Organizations," European Journal of Operational Research, no. 130, pp. 375-387, 2001.

[4] M. Dağdeviren, D. Akay ve M. Kurt, "İş Değerlendirme, Faktör Derece Puanlarının Belirlenmesinde Hedef Programlama Yönteminin Kullanılması," Gazi Üniv. Müh. Mim. Fak., cilt 19, no. 1, pp. 89-95, 2004.

[5] G. Kayhan, "İnsan Kaynakları Performans Değerlendirilmesinde Bulanık AHP/Bulanık Topsis ile Hibrit Bir Yapının Oluşturulması ve Bir Uygulama," Erciyes Üniversitesi, Fen Bilimleri Enstitüsü, Kayseri, 2010.

[6] B. Das ve A. G. Diaz, "Factor Selection Guidelines For Job Evaluation: A Computerized Statistical Procedure," Computers & İndustrial Engineering, no. 40, pp. 259-272, 2001.

[7] S. Krishnamoorthi ve S. K. Mathew, "Business Analytics And Business Value: A Comparative Case Study," Information & Management, Cilt 55, no. 5, pp. 643-666, 2018.

[8] S. Gupta ve M. Chakraborty, "Job Evaluation in Fuzzy Environment," Fuzzy Sets And Systems, Cilt 100, No. 1-3, pp. 71-77, 1998.

[9] K. Tümay Ateş, "Çok Kriterli Karar Verme Teknikleri ile Teknoloji Geliştirme Bölgelerinde Faaliyet Gösteren Firmaların Performans Sıralaması," Çukurova Üniversitesi, Fen Bilimleri Enstitüsü, Adana, 2018.

[10] B. Erokutan, "Mavi Yakalı Personel Seçiminde Çok Kriterli Karar Verme Yöntemlerinin Kullanılması ve Bir Uygulama," Bilecik Şeyh Edebali Üniversitesi, Sosyal Bilimler Enstitüsü, Bilecik, 2016.

[11] H. Şimşek, "Analitik Hiyerarşi Süreci ve Bulanık Analitik Hiyerarşi Süreci Yöntemlerinin İnsan Kaynaklarının Seçiminde Kullanılması: Güvenlik Sektöründe Bir Uygulama," Süleyman Demirel Üniversitesi, Sosyal Bilimler Enstitüsü, Isparta, 2015.

[12] B. Denizhan, A. Yılmaz Yalçıner ve Ş. Berber, "Analitik Hiyerarşi Proses ve Bulanık Analitik Hiyerarşi Proses Yöntemleri Kullanılarak Yeşil Tedarikçi Seçimi Uygulaması," Nevşehir Bilim ve Teknoloji Dergisi, Cilt 6, No. 1, pp. 63-78, 2017.

[13] F. Sofu, "Bulanık Ortamda Çok Kriterli Karar Verme Yöntemi ile Personel Seçimi: Havacılık Sektöründe Bir Uygulama," İstanbul Ticaret Üniversitesi, Fen Bilimleri Enstitüsü, İstanbul, 2018.

[14] F. Ahmed ve K. Kilic, "Fuzzy Analytic Hierarchy Process: A Performance Analysis Of Various Algorithms," Fuzzy Sets and Systems, cilt 361, no. 1, pp. 110-128, 2018.

[15] S. Arıkan Kargı ve Z. B. Aydın, "Bulanık AHP Yönteminin Yenilenebilir Enerji Alternatiflerinin Seçiminde Kullanılması: Bursa Örneği," Akademik Sosyal Araştırmalar Dergisi, cilt 5, no. 55, pp. 60-74, 2017.

Cost Estimation in the Iron and Steel Industry

Gizem KAPANŞAHİN, Filiz ERSÖZ

Department of Industrial Engineering, Karabük University, TURKEY

Abstract: Sacrifice stands for the production of goods and function, constitutes the costs of enterprises. Cost is also defined as the provision of consumed goods and functions by a production enterprise. Accuracy of enterprise activity analysis is very important in order to make appropriate decisions in enterprises. The consistency of the results ensures correct decision-making; provides right marketing and competitive advantage. Various elements are effective in the process of product costing. The items on the basis of product are examined one by one and the analysis is carried out to obtain the unit costs that reflect the reality. The aim of this study is to investigate the factors affecting the costs and to estimate the cost in the integrated system, with data mining classifying models in the process of billet production, in an A enterprise for the Iron and Steel sector. It is targeted to compare obtained estimation results with the costs presented inside the enterprise.

Keywords: Cost Estimate, Prediction Methods, Artificial Neural Networks

REFERENCES

[1] Özmaden, M. Ş. 'Çok Katlı Yapılarda Bina Tanımlayıcı Özellikleri Kullanılarak Monte Carlo Simülasyon Yöntemi ile Maliyet Tahmini,' Gazi Üniversitesi, Fen Bilimleri Enstitüsü, Ankara, 2010.

[2] 'Muhasebe Uygulama,' [Çevrimiçi]. Available: http://www.muhasebeuygulama.com /maliyet-gider-harcama-nedir.html. [%1 tarihinde erişilmiştir7 Mart 2019].

[3] Çalışkan, A., Uygulamalı Maliyet Muhasebesi, Ankara: Nobel Yayın Dağıtım, 2005.

[4] Bucak, S., 'Otomotiv Sektöründe Yapay Sinir Ağı Kullanarak Maliyet Tahmini,' Sakarya Üniversitesi, Fen Bilimleri Enstitüsü, Sakarya, 2007.

[5] Keçe, F., Ömürbek, V. ve Acar, D. 'Gri Temelli Maliyet Tahmini,' Süleyman Demirel Üniversitesi İktisadi ve İdari Bilimler Fakültesi Dergisi, Cilt 21, No. 2, pp. 453-461, 2016.

[6] Gürsoy, A. 'Yapay Sinir Ağları Yaklaşımıyla Lastik Kalıbı Maliyetlerinin Tahmin Edilmesi', Kocaeli Üniversitesi, Fen Bilimleri Enstitüsü, Kocaeli, 2012.

[7] Karataş, E. 'Yapay Sinir Ağları ile Yazılım Projesi Maliyet Tahmini,' İstanbul Üniversitesi, Fen Bilimleri Enstitüsü, İstanbul, 2011.

[8] Acun, Ö. 'Gri Temelli Maliyet Tahminin Mobilya Üretim Sektöründe Uygulanması,' Burdur, 2018.

[9] Mohamed, A. A. S., 'Improving Cost Calculation in The Iron And Steel Industry In Libya Usıng The Standard Cost Method,' University of Craiova, Romania, 2013.

[10] Asmar, M. E. ve Hanna, A. S. 'New Approach to Developing Conceptual Cost Estimates for Highway Projects,' American Society of Civil Engineers, 2011.

[11] Enshassi, A. 'Factors Affecting the Accuracy of Pre-Tender Cost Estimates in the Gaza Strip,' Journal of Construction in Developing Countries, Malaysia, 2013.

[12] Lotfia, E. 'Cost estimation using ANFIS,' The Engıneerıng Economıst, 2016. 10

[13] Guo, X. ve Liu, S. 'Grey Self-memory Combined Model for Complex Equipment Cost Estimation,' The Journal of Grey System, 2017.

[14] Feldman, P. ve Shtub, A. 'Model For Cost Estimation in a Finite-Capacity Environment,' International Journal of Production Research, 2004.

[15] ¨ggen, A. B., 'Cost Estimates, Cost Overruns, and Project Continuation Decisions,' The Accountıng Revıew, 2016.

[16] Malone, P. K., 'Applying System Readiness Levels to Cost Estimates -A Case Study,' 2016.

[17] Vurşan, H., 'Beton Çelik Çubuklarında Mukavemet Özelliklerinin Yapay Sinir Ağları ve Çoklu Regresyon Yöntemleri ile Tahmini,' Karabük Üniversitesi, Fen Bilimleri Enstitüsü, Karabük, 2017.

[18] Ersöz, F. Veri Madenciliği Teknikleri ve Uygulamaları, Ankara: Dijital Basım Yayınevi, 2016.

[19] Takma, Ç. ve Atil, H. 'Çoklu Doğrusal Regresyon ve Yapay Sinir Ağı Modellerinin Laktasyon Süt Verimlerine Uyum Yeteneklerinin Karşılaştırılması,' Kafkas Üniversitesi Veterinerlik Fakültesi Dergisi, 2012.

[20] Ayyıldız, M. 'Yazılım Projeleri Ölçüm Sonuçları Veritabanının Oluşturulması ve Yeni Yazılım Projelerinin Maliyet Tahmininde Kullanımı,' Yıldız Teknik Üniversitesi, Fen Bilimleri Enstitüsü, İstanbul, 2007.

[21] Okutkan, C. 'Borsa İstanbul Şirketlerinin Hisse Senedi Getirilerinin Yapay Sinir Ağları ve Çoklu Regresyon Yöntemleri Kullanarak Analizi,' Kocaeli Üniversitesi, Fen Bilimleri Enstitüsü, Kocaeli, 2014.

[22] Gökalp, E. 'Bir Konfeksiyon Fabrikası için Maliyet Tahmini ve Üretim Planlamaya Aktiviteleri için Bir Karar Destek Sistemi', Orta Doğu Teknik Üniversitesi, Ankara, 2010.

[23] Ökmen, Ö. 'İnşaat Projelerimim Belirsizlik Altında Faaliyet Şebeke Çizelgelemesi ve Erken Maliyet Tahmini: Risk Analizine Dayalı Bir Modelleme Yaklaşımı,' Gaziantep Üniversitesi, 2008.

[24] Yükseltan, E. 'Türkiye Elektrik Enerjisi Tüketiminin Zaman Serileri ile Analizi,' Kadir Has Üniversitesi, Fen Bilimleri Enstitüsü, 2016.

[25] Gülçiçek, Ü. 'Yapı Parametrelerinin Değişimi ile Yaklaşık Kaba İnşaat Maliyet Tahmini,' Sakarya Üniversitesi, Fen Bilimleri Enstitüsü, 2011.

[26] Kurt, B. 'Tahmin Yöntemleri ile Uçak Verilerinin İşlenmesi,' Erciyes Üniversitesi, Fen Bilimleri Enstitüsü, Kayseri, 2018.

Determination of Socio-Economic Factors Affecting Forest Fires (A Case Study of Forest Regional Directorate of Antalya)

Ufuk COŞGUN

Forest Faculty, Forest Policy and Management Department, Karabük University, TURKEY

Abstract: In general, it is seen that the fire data is presented as the amount of the area burned and the number of fires in time. Flammable material loads in forests, behavior patterns and models of flammable materials according to climatic conditions etc. works are carried out rapidly. All these studies aim to manage the process after the fire. There is a need to work on measures to be taken in order to prevent the occurrence of fire. For this purpose; socio-economic factors that cause fires should be determined in regions where forest fire is common. The elimination of these factors will minimize the occurrence of forest fires. In Turkey; considering that 89% of the forest fires are caused by human beings, the importance of socio-economic studies in these regions is increasing. In the studies to determine the socio-economic factors that cause forest fires; In some period, some studies such as multiple regression, correlation, factor analysis etc. were conducted among some socio-economic data determined according to the conditions of the region and / or fire numbers in the given period. In all these studies a time section / part was taken into consideration. Materials and Methods; First time in Turkey with this work; the relationship between the number of forest areas and forest fire counts and the socio-economic variables determined by considering a certain period of time was analyzed together with time and space. In the forest fires in the years 1980-1990-2000 in Antalya Forest Regional Directorate for twelve governmental forest enterprises, the relationship between the number of forest areas and fire numbers and 25 socio-economic variables were determined. These data were analyzed by panel data analysis or Time Series Cross Section Regression (TSCSREG) analysis method.

Variables with no effect in analysis and at the same time the variables / criteria that were derived from each other were taken into consideration and these criteria were eliminated by multiple linkage analysis (multiple linear analysis, multiple collinearity analysis) and reduced

to 12 variables. Analysis of burned forest areas and selected socio-economic variables; It was tested by Fuller and Battese Methods in the scope of TSCSREG Analysis.

Conclusion: In the analysis of the amount of burned forest areas and selected socio-economic variables; A strong relationship between forest fires and selected socio-economic variables shows that the value of R2 is 90.18%. The value of R2 is 70.19%. It shows that there is a relationship between the numbers of fires and the socio-economic variables selected.

Keywords- Forest Fires, Panel Data Analysis, Socio-Economic Factors, Multiple Collinearity Analysis

Introduction

Turkey's forests are administered by the Ministry of Agriculture and Forestry. The General Directorate of Forestry is the largest unit, with a remit to protect, develop and manage the forest areas with 21.3 million hectares area of Turkey. This unit works with regional organizations via 27 "Regional Directorates of Forest" at the national level. During the period 1937–2010, 86.769 forest fires were recorded in Turkey, destroying an area of 1.617.701 hectares.

The forests of Turkey constitute a part of the Mediterranean Forests, and Turkey's forest belt in the Aegean and Mediterranean Regions shows many ecological similarities with Mediterranean forests [1]. On the other hand, Turkey's geographical location and physical structure shows some differences from the other Mediterranean basin countries. It constitutes a bridge between Europe and Asia and due to the geological structure, many different ecosystems appear, resulting in high biological richness with more than ten thousand plant taxa [2]. Not only ecologic, biologic, geologic and physical differences but also socio-economic conditions in Turkey affecting the forestry studies have also differences from the European countries

Similarly, with the countries with Mediterranean type ecosystems [3], at the southern and western parts of Turkey covered by Mediterranean type ecosystems, fire is the one of the most important issues of forest management [4]. As a result of this importance, big amount of the budget dedicated to the forestry studies are reserved to the studies on forest fires [5].

In the context of that, an effective fire management is required and for this it is needed to know the components of fire in detail like climate, socio-economic conditions, vegetation etc. As the studies on fire in Turkey is observed, it is seen that they were mostly based on fire statistics [6] [7], [8], [9], [10]; [11], [12], [13], [14], [15], fire behavior, vegetation dynamics [16], [1], [17], [18], [19], [20], [21]. On the other hand, the studies on socio-economic conditions, which is strictly connected with forest fires especially with fire occurrence, are very few [22]; [23]. A study on the regional status of fires in the region of the Eastern Black Sea region in the field of Bilgili [14].

Nevertheless, the assessment of the situation in the world and Turkey is seen that some of the fire detection data out of the area and number of fires burning [15]. In the case of Antalya Regional Directorate of Forestry, it is seen that the distribution of forest fires by years and seasons is examined [24].

Knowledge on socio-economic factors are especially important for fire prevention studies. Especially, the fact that 89% of forest fires in Turkey are human induced like negligence, realized intentionally, carelessly, accidentally, indicates the importance of socio-economic conditions [25].

In this context, it was aimed to define the effects of socio-economic factors on forest fires in this work. For this goal the Antalya province was chosen as the study area. Antalya province takes the first place in terms of the coverage of burnt area with the annual coverage of 2633 ha between the periods of 2000-2009. Additionally, the biggest forest fire with the coverage of about 15000 ha in the history of Republic of Turkey occurred in Antalya, Serik and Taşağıl forest enterprises in 2008.

The literature review was mainly focused on gaining the knowledge from the countries that confront the same problem and these are the other southern European countries such as Spain Greece and Italy.

Forest fires in the Mediterranean Europe are mostly related to human activities. More than 90% of fires are originated from either deliberate or involuntary causes. Socio-economic changes occurring in Europe in the last decades (e.g., abandonment of agricultural lands, depopulation of rural areas, changes in agriculture and forestry policies, etc.) have driven

landscape transformations affecting fire risk levels through processes like e.g., increase of unmanaged lands, dead and live biomass accumulation, new uses of the forest and natural lands. In this work we analyzed and attempted modelling the influence of socio-economic factors and their change overtime on forest fire occurrence in the Mediterranean Europe (EU-Med) [26].

Despite Mediterranean countries being strongly sensitive to fire risk, few researches have focused on long-term fire risk trends, especially related to socioeconomic development. The present study offers an integrated time-series analysis along a long temporal series (1961–2017) exploring forest fires in Italy and the relative socioeconomic and demographic Dynamics.

The results of the related study in Italy are as follows "Number of fire events, total burnt area and average fire size were studied between 1961 and 2017 in Italy with the aim to identify homogeneous time periods with similar wildfire frequency and severity and correlate them with the background socioeconomic context. Fire attributes had a diverging behavior over time: the number of fires was the highest in the 1970s and the early 1980s; total burnt area was relatively more constant over time with a peak in the 1980s; and, finally, average fire size decreased quite homogeneously from the peak observed in the 1960s and early 1970s. The number of fires and average fire size were significantly influenced by the value of the same variable one year before. Investigating long-term historical outlines of forest fires, a mixed approach based on time-series statistical analysis, multivariate techniques and regressive models intended to define changes in fire regimes and socioeconomic development. In fact, the comparative valuation of the socioeconomic aspects and wildfire trends can reveal a key step to recognizing mitigation and preventive possibilities. Through a multivariate analysis, a substantial difference in the socioeconomic profile can emerge by decade, evidencing a (more or less) rapid socioeconomic development in relation to the evolution of forest fires in Italy [27].

The results of the study in Spain are as follows "The majority of wildfires (95%) in Spain are caused by human activities. However, much wildfire research has focused on the biological and physical aspects of wildfire, with comparatively less attention given to the importance of socio-economic factors. With recent changes in human activity and settlement patterns in many parts of Spain, potentially contributing to the increases in wildfire occurrence recently observed, the need to consider human activity in models of wildfire risk for this region are apparent. A method from Bayesian statistics used; the weights of evidence (WofE) model, to

examine the causal factors of wildfires in the south west of the Madrid region for two differently defined wildfire seasons. The results show that spatial patterns of wildfire ignition are strongly associated with human access to the natural landscape, with proximity to urban areas and roads found to be the most important causal factors [28].

The results of the collaboration of the Spanish and Italian researchers are as follows; "Three 6-year time periods were considered (1988-1993, 1998-2003, 2004-2009). Fire data was extracted from the European Fire database of the European Forest Fire Information System (EFFIS). The analysis was performed in the most fire-affected area of Europe, the European Mediterranean region covered by Portugal, Spain, France, Italy and Greece. Fire data were analyzed according to their main fire cause (accident-negligence and deliberate). Both, fires and socio-economic variables, which represent anthropogenic factors related to fire activity, were mapped on a 10 km x10 km grid. Models of forest fire density were derived separately for each period and per cause category using ordinal regression statistical methods. The best predictors by period and fire cause category were assessed and differences between time periods analyzed. The result show that the variable wildland-urban interface (WUI) is related to higher ratings of fire density. This result was consistent for all periods and for the two type of fire cause analyzed. The overall fit of the models was 40- 50%.for accident-negligence and 50% deliberated caused fires. Despite the relatively homogeneous socioeconomic characteristics of the Mediterranean Europe, differences are found at regional level and by fire cause category" [26].

Again, a recent research on the forest fires events in Spain is as follows. During the period 1999-2008 an annual average of close to 8,600 forest fires burned about 40,000 ha in Galicia. Most fires are human caused (99%), approximately 82% are set intentionally and 5% are either ignited accidentally or through negligence ([29]. However only a limited number of researches has specifically evaluated how the human presence in this territory increase the risk of fire ignition ([30]; [31]. This contrasts with the increasing literature on empirical assessments of the influence of socioeconomic aspects on forest fire risks, using variables such as population density, land cover changes associated with agriculture abandonment, distance to road or the density of human settlements [32]; [33].

The general situation in Greece about forest fire events is as follows; Especially in Greece, where the problem of wildfires is going to be analyzed, the statistics of burnt forest area is alarming. Throughout the period1970-1997, the average burned area had increased from 11,000 ha/year, which was before1970, to 30,000-80,000 ha/year. The figures in the last three years increased to about 100,000-220,000 ha/year [34]. The customary land use that shaped the Mediterranean ecosystems through the centuries had contributed to the wildfire hazard limitation. Grazing of livestock, collection of firewood for domestic uses, resin collection from pine forests, were traditional practices that keeping down the biomass accumulation. Furthermore, watching over the neighbouring forests and firefighting were two actions in which villagers were traditionally engaged. The changes in land use and demography resulted in a significant increase in fire hazard [35]. The expansion of the cities and all kinds of productive infrastructure over these ecosystems caused severe deforestation over the years. As it was mentioned previously, the number and area burned per year have significantly increased in the last decades due to rural emigration from the mountainous areas and to the change of the socio-economics conditions [36]. Consequently, fire shave become more frequent, and more destructive to the ecosystems than used to be in the past. It is worth to mention that in Greece during the period from 1967-1975 there were an average number of 558 fires per year, which has reached to 1,841 per year in the last decade. Contrary to that, recreational use of nature and tourism, have lately brought people near countryside but with different lifestyles and support urbanization, often increasing fire risks by negligence. These new categories of settlements, they do not live off the land, using controlled burning to diminish the progressively uncontrollable fuels [37]. Summer homes spread over coastlines and tourism replaces traditional industries; political unrest during national elections, are some paradigms of these developments. Another aspect of justifying the hypothesis that recent fire activity is not naturally induced are the years 1981, 1985, 1988 where the total burned area reached to large numbers (table 3). In those years the Greek general elections took place and according to [38] periods of political unrest in Greece, mainly the periods before the elections, emphasize the rate of fires [39].

Material and Method

In this work, fire numbers and burnt areas of the forest enterprises of Forest Regional Directory of Antalya between the years of 1980 and 2010, socioeconomic data of 1980's, 1990's and 2000's and fire crimes data of the same periods were analyzed and the correlations between those variables were observed. Fire numbers and burnt areas data were obtained by Forest Fire Fighting Department of the Forest Regional Directory of Antalya [5]. Forest crime data were captured by the forest enterprises chosen as the study areas and fighting department with forest hazards of Forest Regional Directory of Antalya [40]. Socioeconomic data were obtained by the Statistic Institution of Turkey [41], Agriculture Directory of Antalya City [42] and the Agriculture Directory of the towns of Antalya. In all these works, data observation method was used.

Forest Regional Directory of Antalya consists of thirteen (13) forest enterprises. Each of these enterprises deals with the forestry affairs of the towns of Antalya. But, only in one town, Manavgat town, two forest enterprises known as Manavgat and Taşağıl forest enterprises appears. The second enterprise in Manavgat town was found in 2000. Due to that, socioeconomic data was only obtained for the town, but forestry related data like forest crimes, fire numbers and burnt areas were obtained for two forest enterprises.

The correlation of fire data (fire numbers and burnt area) with socioeconomic data and fire crime data were carried out by using Panel Data Analysis (TSCSREG- time series cross section regression-TSCSREG) [43], [44].

However, in the previous studies in Turkey carried out by using multi regression, correlation and factor analysis [22], [23]. In this analysis, in addition to the fire data, total 25 variables including 13 social, 5 economic and 7 forest crime data were used. These additional variables were analyzed under two groups (Table 1). Some variables were eliminated by using multiple linear regressions and multiple collinearity analysis since they don't have any effects and derived from each other's (Table 2).

In this way, the socioeconomic variable effective on the burnt areas (Y1) and fire numbers (Y2) between the years of 1980-2010 were defined. The correlation between these variables were determined by using Pearson Correlation Analysis.

Table 1. Dependent and Independent (Socio Economic Variables/Factors) Variables

Y1	**FIRE NUMBER**
Y2	**BURNT AREA**
F1	The Populations Working in Agriculture
F2	The Populations Working in Industry
F3	The Populations Working in Construction
F4	The Populations Working in Service
F5	Activities not adequately defined
F6	%Unemployment Rate
F7	Agrarian Holders Older Than 55
F8	total age dependency ratio
F9	Literate
F10	literate ratio
F11	Illiterate
F12	illiterate ratio
F13	completing primary school
F14	completing primary school ratio
F15	completing junior high school & higher education
F16	completing junior high school & higher education ratio
F17	Population
F18	population density
F19	Illegal cutting
F20	Transport
F21	Illegal keeping
F22	Illegal consumption
F23	Illegal opening and settlement
F24	Occupation
F25	Grazing

Table 2. Dependent and Independent (Socio Economic Variables/Factors) Variables

Y1	**FIRE NUMBER**
Y2	**BURNT AREA**
F1	The Populations Working in Agriculture
F2	The Populations Working in Industry
F3	The Populations Working in Construction
F4	The Populations Working in Service
F5	%Unemployment Rate
F6	total age dependency ratio
F7	Literate
F8	completing primary school
F9	Population
F10	Illegal cutting
F11	Illegal opening and settlement
F12	Illegal Grazing

The TSCSREG (**T**ime **S**eries **C**ross **S**ection **Reg**ression) procedure analyzes a class of linear econometric models that commonly arise when time series and cross-sectional data are combined. The TSCSREG procedure deals with panel data sets that consist of time series observations on each of several cross-sectional units. Such models can be viewed as two-way designs with covariates

$$y_{it} = \sum_{k=1}^{K} X_{itk}\beta_k + u_{it} \quad i = 1,..,N; \quad t = 1,..,T$$

where N is the number of cross sections, T is the length of the time series for each cross section, and K is the number of exogenous or independent variables. The performance of any estimation procedure for the model regression parameters depends on the statistical characteristics of the error components in the model. The TSCSREG procedure estimates the regression parameters in the preceding model under several common error structures. The

error structures and the corresponding methods the TSCSREG procedure uses to analyze them are as follows:

- One and two-way fixed and random effects models. If the specification is dependent only on the cross section to which the observation belongs, such a model is referred to as a model with one-way effects. A specification that depends on both the cross section and the time series to which the observation belongs is called a model with two-way effects.

- Therefore, the specifications for the one-way model are

$$u_{it} = \nu_i + \epsilon_{it}$$

and the specifications for the two-way model are

$$u_{it} = \nu_i + e_t + \epsilon_{it}$$

where ϵ_{it} is a classical error term with zero mean and a homoscedastic covariance matrix.

- Apart from the possible one-way or two-way nature of the effect, the other dimension of difference between the possible specifications is that of the nature of the cross-sectional or time-series effect. The models are referred to as fixed effects models if the effects are nonrandom and as random effects models otherwise.

- first-order autoregressive model with contemporaneous correlation

$$u_{it} = \rho_i u_{i,t-1} + \epsilon_{it}$$

- The Parks method is used to estimate this model. This model assumes a first-order autoregressive error structure with contemporaneous correlation between cross sections. The covariance matrix is estimated by a two-stage procedure leading to the estimation of model regression parameters by GLS.

- Mixed variance-component moving average error process

$u_{it} = a_i + b_t + e_{it}$

$$e_{it} = \alpha_0 \epsilon_t + \alpha_1 \epsilon_{t-1} + \ldots + \alpha_m \epsilon_{t-m}$$

- The Da Silva method is used to estimate this model. The Da Silva method estimates the regression parameters using a two-step GLS-type estimator.

The TSCSREG procedure analyzes panel data sets that consist of multiple time series observations on each of several individuals or cross-sectional units. The input data set must be in time series cross-sectional form. " Working with Time Series Data," for a discussion of how time series related by a cross-sectional dimension are stored in SAS data sets. The TSCSREG procedure requires that the time series for each cross section have the same number of observations and cover the same time range.

Fuller method is most commonly used in analysis Usually you cannot explicitly specify all the explanatory variables that affect the dependent variable. The omitted or unobservable variables are summarized in the error disturbances. The TSCSREG procedure used with the Fuller-Battese method adds the individual and time-specific random effects to the error disturbances, and the parameters are efficiently estimated using the GLS method. The variance components model used by the Fuller-Battese method is

$$y_{it} = \sum_{k=1}^{K} X_{itk}\beta_k + v_i + e_t + \epsilon_{it} \quad i = 1,\ldots,N; \quad t = 1,\ldots,T$$

The following statements fit this model. Since the Fuller-Battese is the default method, no options are required.

Results and Discussion

Development of efficient forest fire policies requires an understanding of the underlying reasons behind forest fire ignitions. Globally, there is a close relationship between forest fires and human activities, i.e., fires understood as human events due to negligence (e.g.,agricultural burning escapes), and deliberate actions (e.g., pyromania, revenge, land use change attempts). Wildfire occurrence even for human-ignited fires has also been shown to be dependent on biophysical variables [45].

Relations between burning forest areas and socio-economic factors

Analysis of selected socio-economic variables with the amount of forest areas was tested by TSCSREG analysis and Fuller and Battese Method (Table 3, 4, 5, 6 and 7). There is a strong

correlation between the amount of forest areas burned in forest fires and the socio-economic variables selected and shows that the value of R2 is 90.18% (Table 4). When the relationship between the amount of forest areas and the socio-economic factors selected; F3 (The Populations Working in Construction), F4 (The Populations Working in Service), F5 (%Unemployment Rate), F6 (Total Age Dependency Ratio) and F9 (Population) factors appear to be significant (Table 7).

Table 3. Dependent and Independent Variables of Model Description

Model Description	
Estimation Method	Fuller
Number of Cross Sections	12
Time Series Length	3

Table 4. Dependent and Independent Variables of Fit Statistics

Fit Statistics			
SSE	143.905	DFE	23
MSE	6.2568	Root MSE	2.5014
R-Square	0.9018		

Table 5. Dependent and Independent Variables of Variance Components

Variance Component Estimates	
Variance Component for Cross Sections	15.4931
Variance Component for Time Series	1.10147
Variance Component for Error	5.69724

Table 6. Hausman Test

Hausman Test for Random Effects		
DF	m Value	Pr > m
12	11.66	0.4734

Table 7. Dependent and Independent Variables of Parameter Estimates

Parameter Estimates					
Variable	DF	Estimate	Standard Error	t Value	Pr > |t|
Intercept	1	-33.236	5.3823	-6.18	<.0001
F1	1	0.00075	0.00043	1.74	0.0959
F2	1	0.00864	0.00451	1.92	0.0680
F3	**1**	**0.00992**	**0.00273**	**3.64**	**0.0014**
F4	**1**	**-0.0049**	**0.0014**	**-3.52**	**0.0018**
F5	**1**	**1.21868**	**0.3812**	**3.2**	**0.0040**
D6	**1**	**0.4061**	**0.0687**	**5.91**	**<.0001**
F7	1	0.00031	0.0004	0.77	0.4469
F8	1	-0.0013	0.00076	-1.68	0.1066
F9	**1**	**0.00045**	**5.2E-05**	**8.75**	**<.0001**
F10	1	0.0229	0.0345	0.66	0.5139
F11	1	-0.0331	0.0596	-0.56	0.5841
F12	1	0.00099	0.0276	0.04	0.9716

Relations Between Socio-economic Factors and Numbers of Fire

Analysis of socio-economic variables selected by fire numbers was tested by TSCSREG analysis and Fuller and Battese Method (Tables 8, 9, 10, 11, and 12). There is a relationship between the number of fires and the socio-economic variables selected and shows that the value of R2 is 70.19% (Table 10). When the relationship between the number of fire and selected socio-economic factors; F2 (The Populations Working in Industry), F5 (% Unemployment Rate), F9 (Population), F10 (Illegal cutting), F12 (Illegal Grazing) factors (significant) factors appear to be significant (Table 12).

Table 8. Dependent and Independent Variables of Model Description

Model Description	
Estimation Method	Fuller
Number of Cross Sections	12
Time Series Length	3

Table 9. Dependent and Independent Variables of Fit Statistics

Fit Statistics			
SSE	154364.9674	DFE	23
MSE	6711.5203	Root MSE	81.9239
R-Square	0.7019		

Table 10. Dependent and Independent Variables of Variance Components

Variance Component Estimates	
Variance Component for Cross Sections	572.9875
Variance Component for Time Series	448.1269
Variance Component for Error	6598.732

Table 11. Hausman Test

Hausman Test for Random Effects		
DF	m Value	Pr > m
12	14.08	0.2955

Burning forest areas with quantities; There was a high significant correlation between "The Populations Working in Construction", "The Populations Working in Service", "%Unemployment Rate", "Total age dependency ratio" and Population" socio-economic factors.

There is a high significant correlation between the number of fire and "The Populations Working in İndustry"," %Unemployment Rate", "Population", "Illegal cutting" and "Illegal Grazing socio-economic factors. it is necessary to carry out similar studies and analyzes in other areas characterized by intense forest fires in Turkey. As a result of these analyzes; Policies should be developed for socio-economic variables of high importance.

Table 12. Dependent and Independent Variables of Parameter Estimates

Parameter Estimates					
Variable	DF	Estimate	Standard Error	t Value	Pr > |t|
Intercept	1	121.5446	128	0.95	0.3524
F1	1	-0.01312	0.0112	-1.17	0.2534
F2	**1**	**0.262899**	**0.1048**	**2.51**	**0.0196**
F3	1	0.008581	0.0617	0.14	0.8907

F4	1	-0.04681	0.0313	-1.49	0.1489
F5	**1**	**-18.6333**	**8.7657**	**-2.13**	**0.0445**
F6	1	-0.56673	1.7407	-0.33	0.7477
F7	1	-0.00106	0.0095	-0.11	0.9122
F8	1	-0.01003	0.0175	-0.57	0.5721
F9	**1**	**0.002732**	**0.00106**	**2.57**	**0.0172**
F10	**1**	**2.102254**	**0.7993**	**2.63**	**0.0150**
F11	1	-0.57851	1.3064	-0.44	0.6620
F12	**1**	**-2.0978**	**0.7227**	**-2.9**	**0.0080**

Conclusion

Rapid deterioration in nature, significant developments in climate change, increase the importance of forest fires among Mediterranean countries. Due to the unique social and economic structures of the countries, their evaluations about forest fires are different. When the forest fires are examined; burning field and in the statistics made based on the number of fires is observed that there are differences between European countries and Turkey [5]. The causes of the forest fires in Turkey is 85-90% of human origin [23]. But; more resources and time are devoted to work after the fire. Since 85-90% of the forest fires are of human origin, more precautionary measures should be taken into consideration. In respect of measures to be taken before forest fires; especially the socio-economic factors that cause forest fires are of great importance. It is important to develop training and awareness-raising activities for civil society organizations, other public institutions and institutions and institutional infrastructures in terms of sensitivity to forest fires.

Turkey is the most vulnerable region in terms of forest fires in Antalya Regional Forestry Directorate were examined in the study. Selected socio-economic factors were evaluated statistically by taking into consideration the time sections. In order to determine the socio-

economic factors causing forest fires, panel data analysis was first evaluated in this study. In terms of forest fire numbers and burning areas, two different points of view were analyzed.

Examining the relationship between burning forest areas and socio-economic factors;

- F3 (The Populations Working in Construction),
- F4 (The Populations Working in Service),
- F5 (%Unemployment Rate),
- F6 (Total Age Dependency Ratio) ve
- F9 (Population)

factors were found to be statistically significant differences between the burning forest areas. F3, F4 and F5 factors are completely related to employment. Again, F9 factor is a factor related to the population in rural areas. Factors such as the increase in the population and the decrease of employment are the leading factors leading to the emergence of economic problems. Employment will increase if employment decreases, the unemployment rate (F5)is significantly as a factor. Decreasing the share of the population in the employment in the region will lead to different ways of benefiting from other fields. As a result of increasing burning forest areas; a new additional employment area for the removal of damaged forest products in these areas. Therefore, a statistical relationship between these factors and burning forest areas is quite consistent.

Examining the relationship between the numbers of fire and socio-economic factors

- F2 (The Populations Working in Industry),
- F5 (% Unemployment Rate),
- F9 (Population),
- F10 (Illegal cutting),
- F12 (Illegal Grazing)

There was a statistically significant difference between these factors and fire numbers. Among the socio-economic factors, F2, F5 and F9 factors (which are related to employment and population) have been statistically significant factor on fire numbers. Another group of factors is the illegal utilization of forests (illegal grazing and illegal cutting).98% of forest area in Turkey is a State Forest. Therefore, there are serious legal regulations against illegal exploitation from forests. Unlawful beneficiaries from forest areas can develop negative behaviors towards forest areas in response to criminal practices, such as deliberately burning forests. In the forestry work, those who are employed are unemployed or the burning of forests as a result of the punishment of illegal beneficiaries is the realities encountered for the socio-economic conditions of our country. There are two factors that are found as burning forest areas and fire numbers, which are common and have statistically significant differences on both variables. These; F5 (% Unemployment Rate) and F9 (Population). Population growth and unemployment should be seen as the main factors. As a result of the analysis; The socio-economic factors that occur in two main groups, especially in terms of fire numbers, are the factors that should be examined in more detail on the basis of forest villages in the region.

REFERENCES

[1] Kavgacı, A., Carni, A., Başaran, S., Başaran, M.A., Kosir, P., Marinsek, A., Silc, U., 010. Long-Term Post-Fire Succesion of Pinus Brutia Forest in The East Mediterranean. International Journal of Wildland Fire (basım aşamasında).

[2] Güner, A. Ozhatay, N. Ekim, T. and H.C. Baser. 2000 (eds). Flora of Turkey and the East Aegean Islands, Supplement 2. Edinburg University Press.

[3] Anonym, 2010-2013; FUME Project. 7th Framework Programme of European Union.

[4] Anonym, 2010; YARDOP Tamini No: 6665, Çevre ve Orman Bakanlığı, Orman Genel Müdürlüğü, Sivil kültür Daire Başkanlığı, Ankara.

[5] Anonym, 2012; Orman Yangınları İle Mücadele 2012 Yılı Eylem Planı, Orman ve Su İşleri Bakanlığı, Orman Genel Müdürlüğü, Ankara.

[6] Baş, R., 1965; Türkiye'de Orman Yangın Problemleri ve Bazı Klimatik Faktörlerin Ynagınlara Etkileri Üzerine Araştırmalar, OGM, Yayın No: 421/20, Ankara.

[7] Çanakcıoğlu, H., 1985; Orman Koruma, İ.Ü. Orman Fakültesi, İ.Ü. Yayın No: 3315, Orm. Fak. Yayın No: 376, İstanbul.

[8] Çanakcıoğlu, H., 1988; Orman Yangın İstatistikleri, Türkiye Ormanlarını Yangından Koruma Semineri, Tarım Orman ve Köyişleri Bakanlığı, Orman Genel Müdürlüğü, Orman Koruma ve Yangınla Mücadele Daire Başkanlığı, Yayın No: 29, Seri No: 672, Ankara.

[9] Küçükosmanoğlu, A., 1987; Türkiye Ormanlarında Çıkan Yangınların Sınıflandırılması ile Büyük Yangınların Çıkma ve Gelişme Nedenleri, OGM Yayın No: 662, Seri No: 28, Ankara.

[10] Torlakçık, Ş., 1988; İzmir Orman Bölge Müdürlüğü'nde Orman Yangınları ve Yangından Korunma, Türkiye Ormanlarını Yangından Koruma Semineri, Tarım Orman ve Köyişleri Bakanlığı, Orman Genel Müdürlüğü, Orman Koruma ve Yangınla Mücadele Daire Başkanlığı, Yayın No: 29, Seri No: 672, Ankara.

[11] Kayatürk, D., 1988; Çanakkale Orman Bölge Müdürlüğü'nde Orman Yangınları ve Yangından Korunma, Türkiye Ormanlarını Yangından Koruma Semineri, Tarım Orman ve Köyişleri Bakanlığı, Orman Genel Müdürlüğü, Orman Koruma ve Yangınla Mücadele Daire Başkanlığı, Yayın No: 29, Seri No: 672, Ankara.

[12] Toklu, Y., 1988; Antalya Orman Bölge Müdürlüğü'nde Orman Yangınları ve Yangından Korunma, Türkiye Ormanlarını Yangından Koruma Semineri, Tarım Orman ve Köyişleri Bakanlığı, Orman Genel Müdürlüğü, Orman Koruma ve Yangınla Mücadele Daire Başkanlığı, Yayın No: 29, Seri No: 672, Ankara.

[13] Akaydın, S., 1988; Muğla Orman Bölge Müdürlüğü'nde Orman Yangınları ve Yangından Korunma, Türkiye Ormanlarını Yangından Koruma Semineri, Tarım Orman ve Köyişleri Bakanlığı, Orman Genel Müdürlüğü, Orman Koruma ve Yangınla Mücadele Daire Başkanlığı, Yayın No: 29, Seri No: 672, Ankara.

[14] Bilgili, E., Dinç Durmaz, B., Baysal, İ., Sağlam, B., Küçük, Ö., 2010; Doğu Karadeniz Ormanlarında Orman Yangınları, III. Ulusal Karadeniz Ormancılık Kongresi, Cilt: III, Sayfa: 1280-1290, Artvin.

[15] Ertuğrul, M., 2005; Orman Yangınlarının Dünyadaki ve Türkiye'deki Durumu, ZKÜ, Bartın Orman Fakültesi Dergisi, Bartın.

[16] Kavgacı, A., Silc, U., Carni, A., Başaran, S., Başaran, M.A., Kosir, P., Marinsek, A., Zelnik I., 2009. Serik ve Manavgat Bölgelerindeki Kızılçam Ormanlarında Yangın Sonrasında Vejetasyon

Yapısında Meydana Gelen Değişim. 1. Orman Yangınlarıyla Mücadele Sempozyumu Tebliğleri, 7-10 Ocak 2009, Antalya, Orman Genel Müdürlğü Yayını, s. 295-302.

[17] Tavşanoğlu, Ç., 2010. Seed Production and Fruit Parasitism in *Cistus Salviifolius* L. (Cistaceae) Along A Post-Fire Successional Gradient. J. An. Vet. Adv. 9(7): 1120-1127.

[18] Tavşanoğlu, Ç., Gürkan, B., 2004. Akdeniz Havzasında Bitkilerin Kuraklık ve Yangına Uyumları. Ot Sistematik Botanik Dergisi, 11(1):119-132.

[19] Tavşanoğlu, Ç., Gürkan, B., 2005. Post-Fire Dynamics of *Cistus* Spp. in a *Pinus brutia* forest. Turkish J. Bot., 29:337-343.

[20] Tavşanoğlu, Ç., Gürkan, B., 2009. Post-Fire Regeneration of a *Pinus Brutia* (Pinaceae) Forest In Marmaris National Park, Turkey. Int. J. Bot., 5:107-111.

[21] Tavşanoğlu, Ç., Kaynaş, B.Y., Gürkan, B., 2002. Plant Species Diversity in A Post-Fire Successional Gradient in Marmaris National Park, Turkey. In: Viegas, X.V. (Ed.) Proceedings of the IV. International Conference on Forest Fire Research – 2002 Wildland Fire Safety Summit, 18-23 November, Luso, Coimbra, Portugal.

[22] Öktik, N., 2001; Orman Yangınlarının Sosyo-Ekonomik ve Kültürel Nedenleri, Muğla Üniversitesi Yayın No: 16, Muğla.

[23] Coşgun, U., Yolcu, H., İ., Tolunay, A., Orhan, K., H., 2010; Antalya Orman Bölge Müdürlüğünde Orman Yangınlarına Neden Olan Sosyo-Ekonomik Faktörlerin Belirlenmesi, Batı Akdeniz Ormancılık Araştırma Müdürlüğü, Teknik Bülten No: 40, Antalya.

[24] Coşgun, U., Kavgacı, A., Güngöroğlu, C., 2012; Türkiye'deki Orman Yangın Rejim Verilerinin Değerlendirilmesi Antalya Orman Bölge Müdürlüğü Örneği, Batı Akdeniz Ormancılık Araştırma Müdürlüğü Dergisi, Antalya.

[25] Coşgun, U., González-Cabán, A., 2019; Factors Explaining Forest Fires in the Serik and Taşağıl Forest Provinces (SW Anatolia), Proceedings of Fifth International Symposium on Fire Economics, Planning and Policy: Ecosystem Services and Wildfirees, USDA, Genral Technical Report PNW-GTR-'61 (English). February, pages; 145,165.

[26] Vilar L., Camia A., Ayanz, J., S., M., 2014; Modeling Socio-Economic Drivers of Forest Fires In The Mediterranean Europe, Chapter 7, Advances in Forest Fire Research, Coimbra University press.

[27] Carlucci, M., Zambon, I., Colantoni, A., Salvati, L., 2019; Socioeconomic Development, Demographic Dynamics And Forest Fires in Italy, 1961–2017: A Time-Series Analysis, Sustainable 11(5), 1305 Case Report.

[28] Calcerrada R., R., Novillo, C., J., Millington J., D., A., Jimenez, I., G., 2008 GIS Analysis of Spatial Patterns of Human-Caused Wildfire Ignition Risk In The SW Of Madrid (Central Spain), Landscape Ecol., DOI. 10.1007/s 10980-008-9190-2, Spain.

[29] Chas-Amil M.L., Touza J., Prestemon J.P., 2010. *Spatial Distribution of Human-Caused Forest Fires in Galicia (NW Spain).* In G. Perona and C. A. Brebbia (eds.) Modelling, Monitoring and Management of Forest Fires. WIT Press, 247-258.

[30] Martínez, J., C. Vega-Garcia, E. Chuvieco. 2009. *Human-Caused Wildfire Risk Rating for Prevention Planning in Spain*. Journal of Environmental Management, 90, 1241-1252.

[31] Padilla, M. and Vega-Garcia, C., 2011. *On the Comparative Importance of Fire Danger Rating Indices and Their Integration with Spatial and Temporal Variables For Predicting Daily Human-Caused Fire Occurrences in Spain*. International Journal of Wildland Fire, 20(1), 46-58.

[32] Moreira, F., Viedma, O., Arianoutsou, M., Curt, T., Koutsias, N., Rigolot, E., Barbati, A., Corona, P., Vaz, P., Xanthopoulos, G., Mouillot, F. and Bilgili, E., 2011. Landscape Wildfire Interactions in Southern Europe: Implications for Landscape Management. Journal of Environmental Management, 10, 2389-402.

[33] Narayanaraj, G. and Wimberly M.C., 2012. *Influences of Forest Roads on The Spatialpatterns Of Human- And Lightning-Caused Wilfire Ignitions*. Applied Geography, 32, 878-888.

[34] Efthymiou, P., 2000. *Mediterranean Forest Fires 2000-A Terrible and Complex Situation*. EFI News No. 2. Volume 8. Finland.

[35] Xanthopoulos, G., 1995. *Lectures Notes*. Seminar on Forest Fire Prevention and Security.Educational Society. Athens.

[36] Papanastasis, V.P., 1999. *Land Degradation Caused by Overgrazing and Wildfires and Management Strategies to Prevent and Mitigate Their Effects*. Faculty of Forestry and NaturalEnvironment, Aristotle University. Thessalonica. Greece.

[37] Pyne, S. 1997. *World Fire. The culture of fire on Earth.* Weyerhaeuser Environmental Books.University of Washington Press. Seattle and London.

[38] Papastavrou, A., 1992. *Social, Economic and Cultural Aspects of Forest and Wild land Fires inGreece.* Seminar on Forest Fire Prevention, Land Use and People. Joint Committee on Forest Technology, Management and Training.

[39] Maheras G., 2002 Forests fires in Greece. The Analysis of The Phenomenon Affecting Both Natural Andhuman Environment. The Role of Sustainable Development in Controlling Fire Effects. M Sc. Thesis, LUMES.

[40] Anonym, 2014; Antalya Orman Bölge Müdürlüğü, Koruma Şube Müdürlüğü, 1978-2014 Orman Yangınları Defteri, 2014, Antalya.

[41] Anonym, 2013/a; Türkiye İstatistik Kurumu, İl ve İlçeler Düzeyinde Nüfus ve İstihdam verileri, Ankara.

[42] Anonym, 2013/b; Tarım İl Müdürlüğü Tarımsal Alan Verileri, Antalya.

[43] Qi, Y., Smith, B.L., and J. Guo 2007. *Freeway Accident Likelihood Prediction Using a Panel Data Analysis Approach*, Journal of Transportation Engineering, March 2007, 149-156.

[44] Meserve, S.A., Pemstein, D., and W.T. Bernhard 2009. *Political Ambition and Legislative Behavior in the European Parliament*, The Journal of Politics, 71, 3, 1015–1032.

[45] Chas-Amil M., J., Touza J., Prestemon J., P., McClean, C., J.,2012 Natural and Social Factors Influencing Forest Fire Occurrence At A Local Spatial Scale, Modeling Fire Behavior And Risk (Project09SEC011201PR).

Group Acceptance Sampling Plans Based on Time Truncated Life Tests For Compound Weibull-Exponential Distribution

Canan HAMURKAROĞLU[1], Ayten YİĞİTER[2]

[1]Department of Actuary and Risk Management, Karabük University
[2]Department of Statistics, Hacettepe University, TURKEY

Abstract: Acceptance sampling plans are a decision-making process on the basis of a randomly selected sampling from a party, where it is not possible to completely scan the products for reasons such as time and cost being limited or the formation of damaged products during the inspection. For some products, the life span (time from beginning to failure) may be an important quality characteristic. In this case, the quality control adequacy of the products can be checked with an acceptance sampling plan based on the truncated life test with a censored scheme for the lifetime of the products. Acceptance sampling plans based on life test of product life in industry are called reliability plans. In this study, group acceptance sampling plans based on life tests were studied under the type-I censored scheme for the compound Weibull-Exponential distribution. Optimum sample size, optimum number of groups and acceptance number were obtained.

Keywords: Acceptance Sampling Plan, Truncated Survival Test, Producer Risk, Consumer Risk, Compound Weibull Exponential Distribution

Stochastic Approach to a Buffer Stock Problem

Zülfiye HANALİOGLU[1], Tahir HANALİOGLU[2]

[1]Department of Actuary and Risk Management, Karabük University,
[2]Department of Industrial Engineering, TOBB University, TURKEY

Abstract: In this study, a buffer stock between two machines which are working at the same speed, is considered. It is assumed that the stock level between two machines alters in the interval $\mathbf{0}, \mathbf{2}\boldsymbol{a}$. In the case that only the first machine is broken, the stock level can decrease to zero if repairing time of the first machine is extended. This causes the second machine is required to be halted. After fixing the first machine, the system begins to work again. If the second machine is broken and the first machine is working, then the process proceeds until the stock level reaches the maximum level $\mathbf{2}\boldsymbol{a}$ and then the first machine will be halt compulsory. In order to re-work the system, the repairing of the second machine must be completed. Under these assumptions, the buffer stock level will be stochastically fluctuated in the interval $\mathbf{0}, \mathbf{2}\boldsymbol{a}$. Thus, the buffer stock between two machines can be expressed by a stochastic process $\boldsymbol{Y}\ \boldsymbol{t}$ and it is observed that this process is a random walk with two barriers. Some problems of queuing theory, stock control, reliability, insurance models and risk management can be expressed by random walk and its modifications. In literature, there are several considerable studies (e.g., Aliyev and Khaniyev, Feller, Janseen and Leeuwarden, etc.). In this study, the stationary characteristics of the random walk process $\boldsymbol{Y}\ \boldsymbol{t}$ which represent the buffer stock level, are investigated. Especially, for all moments of the ergodic distribution of the process $\boldsymbol{Y}\ \boldsymbol{t}$, the exact expressions are obtained under the assumption that the random walk $\boldsymbol{Y}\ \boldsymbol{t}$ generated by bilateral exponential distributed summands. Moreover, the exact and approximate expressions for variance, standard deviation, coefficient of variation, skewness and kurtosis of the process, are obtained. It is also observed that the ergodic distribution of the standardized stochastic process $\boldsymbol{W}\ \boldsymbol{t} = \boldsymbol{Y}\ \boldsymbol{t}\ /\boldsymbol{a}$ weakly converges to a triangular distribution in the interval $\mathbf{0}, \mathbf{2}$.

Keywords: Buffer Stock, Random Walk, Ergodic Distribution, Stationary Characteristics.

REFERENCES

[1] Aliyev R. T. and Khaniyev T. A., 2017, "On the Limiting Behavior of the Characteristic Function of the Ergodic Distribution of the SemiMarkov Walk with Two Boundaries" Mathematical Notes, 102 (4), 444-454.

[2] Borovkov A.A.,1984, "Asymptotic Methods in Queuing Theory", J. Wiley, New York.

[3] Chang J.T. and Peres Y., 1997, "Ladder Heights, Gaussian Random Walks And The Riemann zeta function", Annals of Probability, 25, 787–802.

[4] Feller W., 1971, An Introduction to Probability Theory and Its Applications II, John Wiley.

[5] Gihman I.I. and Skorohod A.V., 1975, Theory of Stochastic Processes II, Springer –Verlag.

[6] Janssen A.J.E.M. and Leeuwarden J.S.H., 2007, "On Lerch's Transcendent And The Gaussian Random Walk", Annals Of Applied Probability, 17, 421–439.

[7] Khaniev T. A., Unver I. and Maden S., 2001, "On The Semi – Markovian Random Walk With Two Reflecting Barriers", Stochastic Analysis And Applications, 19(5), 799–819.

[8] Khaniev T. A. and Kucuk Z., 2004, "Asymptotic Expansions For The Moments of The Gaussian Random Walk With Two Barriers", Statistics And Probability Letters, 69(1), 91 – 103.

[9] Lotov V.I., 1996, "Some Boundary Crossing Problems For Gaussian Random Walks", The Annals of Probability, 24(4), 2154–2171.

[10] Spitzer F., 1964, Principles of Random Walk, Princeton, N. J., D. Van Nostrand.

Prediction of Air Permeability of Denim Fabrics Using Articifial Neural Networks

Esra AKGÜL, Emel AYDOĞAN, Yılmaz DELİCE

Department of Industrial Engineering, Erciyes University, TURKEY

Abstract: Denim is a popular fabric among all of the age groups because of its good usage performance and ability to provide convenience in adapting to changing trends in fashion. Apart from the fashion and general performance properties, thermo-physiological comfort properties such as air permeability are important for denim users. Fabric comfort depends on lots of factor such as fabric structure and the types of fibers. Air permeability is a one of the comfort properties of fabric is affected by many parameters of the fabric. A determination of the relationship between the fabric parameters and the air permeability is highly complex and difficult. For this reason, Artificial Neural Network model which has effective performance in very complex problems was used. In the present study, an artificial neural network has been used to predict air permeability amongst different denim production parameters. Finally, by comparison with the experimental results, the efficacy of the proposed model was verified.

Keywords: Artificial Neural Network, Denim Fabric, Optimization.

REFERENCES

[1] Milenkovic L, Skundric P, Sokolovic R & Nikolic T, The Scientific Journal Facta Universitatis, 1(4), 1999, pp. 101. [2] Oğlakcioğlu, N., & Marmarali, A., "Thermal comfort properties of some knitted structures", Fibres & Textiles in Eastern Europe, 15(5-6), 2007, pp. 64-65.

[3] Saville B.P., Physical testing of textiles, The Textile Institute, Woodhead publishing limited, Cambridge-England, 2003. [4] Çay, A., Vassiliadis, S., Rangoussi, M., & Tarakçıoğlu, I. "Prediction of the air permeability of woven fabrics using neural networks", International Journal of Clothing Science and Technology, 19(1), 2007, pp. 18-35.

[5] TS 391 EN ISO 9237, Tekstil-Kumaşlarda Hava Geçirgenliğinin Tayini, Türk Standartları Enstitüsü, Ankara, 1999.

Numerical Investigation of Cutting Forces in Turning of C23000 Brass Alloy

Mehmet Erdi KORKMAZ

Department of Mechanical Engineering, Karabük University, TURKEY

Abstract: Brass alloys are characterized by excellent workability, high thermal and electric conductivity, corrosion resistance as well as exceptional antibacterial properties and are therefore widely used in various industries, such as electric and electronics, automotive and sanitary industry. However, power consumption should be eliminated for cleaner production in terms of sustainable machining. Therefore, this study aims modelling of cutting forces in hard turning of C23000 brass based on finite element method. The cutting parameters are chosen as cutting speed, depth of cut and feed rate with three levels. The average of 4.66% difference is achieved between experimental and simulated feed forces while 4.39% difference for main cutting forces. The finite element modelling of cutting forces is quite compatible with the experimental results and it can be performed by high accuracy without excessive machining experiments of high machinability materials.

Keywords: Brass Alloys, Finite Element Method, Main Cutting Force, Feed Force

Introduction

Brass alloys are characterized by excellent workability, high thermal and electric conductivity, corrosion resistance as well as exceptional antibacterial properties and are therefore widely used in various industries, such as electric and electronics, automotive and sanitary industry. To enhance their machinability, lead is commonly added to brass alloys, leading to excellent chip breakage, low tool wear, and high applicable cutting parameters. Main applications are in electric and electronics, automotive, and sanitary industry. Since the amount of cutting operations when manufacturing brass components is high, different alloying elements enhancing the machinability are usually added to brass [1].

The most essential element in this context is lead (Pb), improving the machinability referring to chip breakage, tool wear, cutting forces, and applicable range of cutting parameters. The machinability of these low-leaded brasses is significantly worse compared to leaded free-cutting brass. Depending on their chemical composition and microstructure, different machinability problems arise.

The cutting forces are a result of extreme conditions at the toolworkpiece interface. The interaction can be directly related to the tool wear and in worst cases to the failure of the tool. Consequently, the tool wear and cutting forces are related to each other. Thus, it is necessary to carry out the optimization of cutting process to evaluate the optimal values of cutting parameters to determine the performance and useful life of the cutting tool [2].

Surinder et al. [3] investigated the cutting forces (tangential and feed force) in turning of unidirectional glass fiber reinforced plastics (UD-GFRP). The process parameters of cutting tool (nose radius, rake angle, cutting speed, feed rate, depth of cut and cutting environment) were investigated using Taguchi robust design methodology. The relative significance of parameters was studied using ANOVA. The tangential force was found to decrease with decrease in tool nose radius, feed rate and depth of cut and increase with the cutting speed. Cascona et al. [4] developed mechanistic model for prediction of cutting forces in turning of non-axis-symmetric parts. This study presents a mechanistic model for predicting the orthogonal turning forces (in 3 directions), torque and power consumption along the machining path of non-axis-symmetric parts. Dorlin et al. [5] studied the geometrical modeling of toolworkpiece interaction and its effects on the cutting forces during turning. The analysis focused on convex contact radius between the machined part and the tool. The experiments were based on cylindrical and face turning of Ti6Al4V titanium alloy. It was observed that the contact radius had significant effect on the cutting forces and the cutting forces increase with the increase in the radius. Xie et al. [6] studied cutting force and cutting temperature during the turning of titanium alloy using micro-grooved tool under dry conditions. The objective of the study was to estimate the influence of shape and size of micro groove on the temperature and force in dry turning. The micro-grooved tool decreases cutting temperature by 103 _C, while as the shear angle increases with decreasing micro-groove depth. Philip et al. [7] studied the effects of cutting speed and feed rate on tool wear, surface roughness and cutting force on

nitrogen alloyed duplex stainless steel in a dry turning process, using Taguchi method. The results revealed that the feed had the most significant influence on the cutting forces. The cutting speed was found to be the most significant parameter affecting the tool wear. Shear force, ploughing force and particle fracture force were considered by Sikder et al. [8] to estimate the cutting force during the machining of metal matrix composites (MMCs). The chip formation force, ploughing force and fracture force were obtained by Johnson-Cook model, slip line filed theory and Griffith's theory respectively. The results showed good agreement between the predicted and experimental values of the cutting forces.

Literature researches showed that there is no study performed on finite element modelling on power consumption for C23000 brass alloy. Therefore, in this work, FE modelling was performed to determine the cutting forces in turning of the material based on cutting parameters which are cutting speed, feed rate and depth of cut.

Material and Method

The aim of this study was to investigate the effects of cutting parameters on cutting forces in turning of C23000 brass with HSS tools by using finite element method. For this purpose, the finite element analysis which are based on experiments by Hanief et al. [2] were performed with Third Wave Advantedge software with 2D orthogonal turning instead of 3D turning due to low calculation time [9].

Work Piece Material

The workpiece has hardening of 120 Bhn by means of heat treatment by quenching in a vacuum atmosphere. The Johnson–Cook model [10] widely-used material model for machining simulations is given in Eq. 1. This material model is particularly suited to model high strain rate deformation of metals. It is generally used in adiabatic transient dynamic analysis. The hardening is a particular type of isotropic hardening in which the yield stress σ0 is assumed as [11]:

$$\sigma^0 = (A + B(\varepsilon^p)^n)\left(1 + C\ log\left(\frac{\dot{\varepsilon}^p}{\dot{\varepsilon}_0}\right)\right)\left(1 - \left(\frac{T - T_r}{T_m - T_r}\right)^m\right) \quad (1)$$

In equation (1), material parameters obtained from mechanical tests that are A, B, C, n and m are yield stress below room temperature, strain hardening, strain rate constant, strain hardening constant and thermal softening constant, respectively. The other parameters ε^p, $\dot{\varepsilon}^p$, $\dot{\varepsilon}_0$, T_r, T_m and T are equivalent plastic strain, plastic strain rate, reference strain rate, room temperature, melting temperature and reference temperature, respectively. Also, $\dot{\varepsilon}_0$ and C are usually measured at or below the reference temperature. The Johnson-Cook parameters and other material parameters for AISI 52100 steel was given in Table 1 and Table 2, respectively from finite element software.

Tablo 1. Johnson–Cook parameters for the C23000 brass alloy material

A	B(MPa)	C	n	m	Tr	Tm	$\dot{\varepsilon}_0$
232	931	0.0682	0.93	0.4682	27	1083	1

Tablo 2. The other material parameters for the C23000 brass alloy material

Density (g/cm3)	Poisson ratio	Young's Modulus (GPa)	Thermal Conductivity (W/m.K)	Specific Heat (J/kg°C)	Thermal Expansion (10-6 °C)
8.4	0.3	103	116	477	20.5

Cutting Tool

HSS inserts have been used as cutting tools in 3D orthogonal analyses. The tool has a rake angle (γ)= -6° and clearance angle (α)= 0° with the edge radius (r) of 0.02 mm because edge radius is used instead of tool nose radius in 2D orthogonal turning (Fig. 1). The material parameters for HSS cutting tool was given in Table 3.

Tablo 3. The material parameters for HSS tool

Density (g/cm3)	Poisson ratio	Young's Modulus (GPa)	Thermal Conductivity (W/m.K)	Specific Heat (J/kg°C)	Thermal Expansion (10-6 °C)
7.8	0.3	207	24	420	11.9

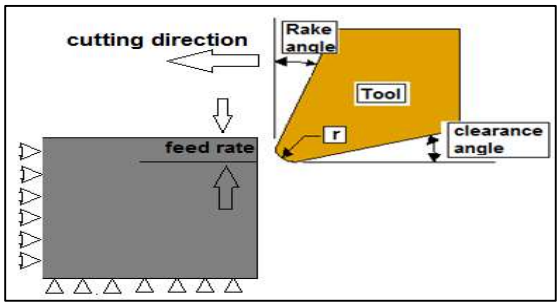


Fig. 1 2D model of cutting tool

Finite Element Simulations

The finite element analyses of cutting forces were performed with Advantedge software depending on finite element method. Advantedge uses an Arbitrary Lagrangian solver and it has adaptive remeshing function to provide more accurate results although it takes more time. The first stage of simulation is to determine the workpiece length (3 mm) and height (1 mm) with workpiece material. The second stage is determining the tool parameters (rake angle, clearance angle and edge radius) with tool material. The final stage is to enter required simulation parameters such as feed rate, depth of cut, length of cut and cutting speed after meshing parameters and coefficient of friction is adjusted.

The interface between tool and work piece was modelled with a standard Coulomb friction which is assumed as 0.6. The meshing parameters were used as 0.1 mm and 0.02 mm for maximum and minimum element size, respectively. After these assumptions, a verification simulation was done with parameters used by Table 1-3. The finite element model was then verified due to the difference between cutting forces in experiment and simulation is less than %5. The cutting parameters chosen for these finite element analyses were given in Table 4.

Tablo 4. Cutting Parameters

Levels	1	2	3
Feed rate	0.12		
Cutting speed	840	1000	1280
Depth of cut	0.10	0.13	0.16
Coefficient of	0.5		

The workpiece was cut off 2mm in analyses of cutting forces. After the cutting process is finished, both the chip and tool are removed, and the workpiece is allowed to thermo-mechanically relax. The simulation model and cutting scheme were shown in Fig. 2.

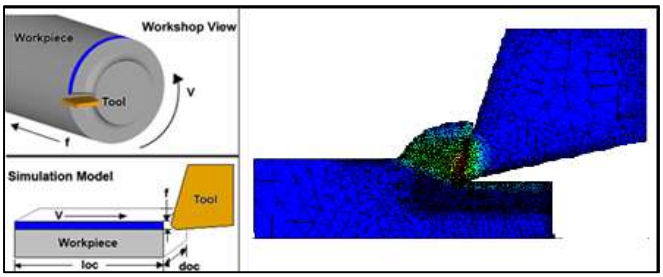


Fig. 2 2D Simulation Model

Results And Discussion

Only 2D simulation was performed in this study instead of 3D simulation due to low calculation of time. As shown from Fig 3-4, the influences of cutting parameters (depth of cut and cutting speed) on cutting forces were assessed by means of the figures. In general, Fig. 3-4 display different tendency. When both methods were compared, increase in depth of cut generally led to decrease and then increase in feed force. This situation may be referred to increasing ploughing effect with increasing depth of cut.

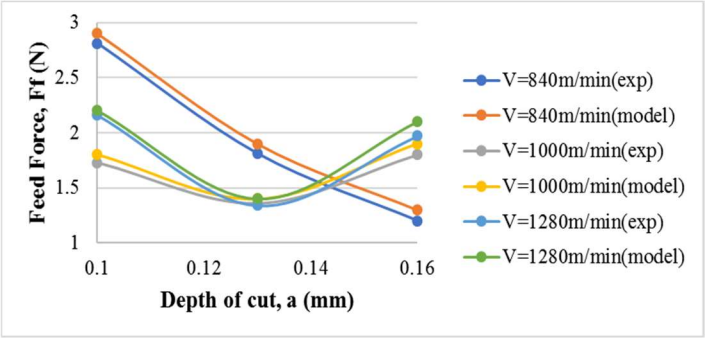


Fig. 3. The Variations of The Feed Force

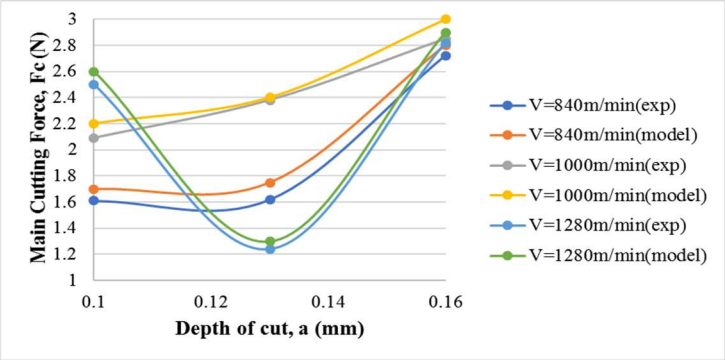


Fig. 4. The Variations of The Main Cutting Force

Numerical Ff values was determined as 4.66% different from that of the experimental results. This result may arise from the data in literature of the Johnson-Cook model for workpiece and coefficient of friction in tool-workpiece interface. It was referred in literature [12,13] that the metallurgical structure and the chemical composition of the standard manufactured material may be different. Mechanical and physical properties that generate material models should be determined according to the related experimental workpiece in cutting simulations. Figure 3 also showed that the cutting speed has great importance on feed force. The numerical feed force decreased about 34% and 32% while depth of cut increased from 0.1 to 0.13 mm and from 0.13 to 0.16 mm, respectively by kept feed rate and cutting speed constant (Figure 3). Decreasing and increasing ratio of numerical feed force was about 38% and 22% while cutting speed increased from 840 to 1000 m/min and from 1000 to 1280 m/min, respectively at constant depth of cut and feed rate. The lowest numerical feed force was obtained as 1.3 N by the feed rate of 0.12 mm/rev, depth of cut of 0.16 mm and cutting speed of 840 m/min.

It was determined that both experimental and numerical Fc values show similar tendency (seen Fig. 4) with a percentage of 4.39% difference. The main cutting force values generally increased with increasing depth of cut while This situation is consistent with literature [14]. As can be derived from Fig. 4, the numerical main cutting force raised about 3% and 60% while the depth of cut increased from from 0.1 to 0.13 mm and from 0.13 to 0.16 mm, respectively at constant feed rate and cutting speed. Increasing ratio of numerical main cutting force was about 29% and 18% while the cutting speed increased from 840 to 1000 m/min and from 1000 to 1280 m/min, respectively by constant depth of cut and feed rate. The lowest

numerical main cutting force was acquired as 1.3 N with the feed rate of 0.12 mm/rev, depth of cut of 0.13 mm and cutting speed of 1280 m/min.

Conclusion

In this study, the effects of cutting parameters on cutting forces have been analyzed with finite element method in hard machining of C23000 brass alloy workpieces with 120 Bhn hardness using coated carbide tools. 2D orthogonal cutting has been used instead of 3D cutting due to low calculation of time in finite element analyses. The following conclusions are drawn from this study:

Firstly, the finite element analysis was resulted in a numerical model and the validation of model was approved by comparing experimental cutting forces, and power consumption. It was found that there is an average of 4.66%, and 4.39% deviation between experimental and simulation results of feed force and main cutting force respectively. It can be concluded that optimizing the depth of cut, cutting speed and feed rate can result in reduced cutting forces. Consequently, 2D finite element model can be recommended in machining operations for optimizing the cutting forces without having the need to perform trial experiments on brass alloys or their components. Therefore, sustainable machining can be acquired in manufacturing industry involving easy-to-machine materials through cutting simulations by finite element method.

REFERENCES

[1] N.M. Vaxevanidis, N.A. Fountas, A. Koutsomichalis, J.D. Kechagias, Experimental investigation of machinability parameters in turning of CuZn39Pb3 brass alloy, Procedia Structural Integrity, Volume 10, (2018), Pages 333-341.

[2] M. Hanief, M.F. Wani, M.S. Charoo, Modeling And Prediction of Cutting Forces During The Turning Of Red Brass (C23000) Using ANN Andregression Analysis, Engineering Science And Technology, an International Journal 20 (2017) 1220–1226.

[3] Surinder Kumar, Meenu Gupta, P.S. Satsangi, Multiple-Response Optimization of Cutting Forces İn Turning Of UD-GFRP Composite Using Distance-Based Pareto Genetic Algorithm Approach, Eng. Sci. Technol. 18 (2015) 1–16.

[4] C. Itxaso, J.A. Sarasuaa, Mechanistic Model For Prediction of Cutting Forces İn Turning Of Non-Axisymmetric Parts, Procedia CIRP 31 (2015) 435–440.

[5] T. Dorlin, F. Guillaume, J.P. Costes, Analysis and Modeling of The Contact Radius Effect On The Cutting Forces İn Cylindrical And Face Turning Of Ti6al4v Titanium Alloy, Procedia CIRP 31 (2015) 185–190.

[6] J. Xie, M.J. Luo, K.K. Wu, L.F. Yang, D.H. Li, Experimental Study on Cutting Temperature And Cutting Force in Dry Turning Of Titanium Alloy Using Microgrooved Tool, Int. J. Mach. Tools Manuf. 73 (2013) 25–36.

[7] D. Philip, P. Selvaraj, P. Chandramohan, M. Mohanraj, Optimization Of Surface Roughness, Cutting Force And Tool Wear Of Nitrogen Alloyed Duplex Stainless Steel in A Dry Turning Process Using Taguchi Method, Measurement 49 (2014) 205–215.

[8] S. Sikder, H.A. Kishawy, Analytical Model For Force Prediction When Machining Metal Matrix Composite, Int. J. Mech. Sci. 59 (2012) 95–103.

[9] John, M.R.S.; Shrivastava, K.; Banerjee, N.; Madhukar, P.D.; Vinayagam, B.K.: Finite Element Method-Based Machining Simulation for Analyzing Surface Roughness During Turning Operation with HSS and Carbide Insert Tool, Arab. J. Sci. Eng. 38, 1615–1623 (2013)

[10] M.E. Korkmaz, N. Yaşar, M. Günay, Finite Element Modeling of Residual Stresses and Cutting Temperature in Hard Turning, International Conference on Engineering and Natural Sciences, Bosnia and Herzegovina, 2016.

[11] A. Dorogoy and D. Rittel, Determination of the Johnson–Cook Material Parameters Using the SCS Specimen, Experimental Mechanics 49, 881–885, 2009.

[12] Bil, H.; Kılıç, S.E.; Tekkaya, A.E.: A Comparison of Orthogonal Cutting Data From Experiments With Three Different Finite Element Models. Int. J. Mach. Tool. Manuf. 44(9), 933-944 (2004)

[13] Özel, T.: The Influence of Friction Models on Finite Element Simulations of Machining. Int. J. Mach. Tool. Manuf. 46, 518-530 (2006)

[14] Günay, M.; Korkmaz, M.E.; Yaşar, N.: Finite Element Modeling of Tool Stresses on Ceramic Tools in Hard Turning. Mechanika. 23(3), 432-440 (2017).

Evaluation of Critical Factors in Industry 4.0 Transition Processes by R'WOT Analysis

Gizem ACAR[1], İlker ÖNALAN[2]

[1]Department of Industrial Engineering Karabük University,

[2]Department of Mechatronics Engineering, Karabük University, TURKEY

Abstract: Industry 4.0 was first announced to the public by the German Federal Government at the 2011 Hannover Fair. The concept of Industry 4.0 is taken into consideration especially in our country, considering the rapidly developing countries. In this study the work of leading organizations in Turkey is being examined and the approaches to Turkey's Industry 4.0 are reported. In the study, SWOT Analyzes are compiled from the results of civil society organizations reports, academic publications, public institutions report, consulting companies and related books. As a result of this review, a single SWOT analysis consisting of 9 items in each item is presented with the text mining method. The SWOT Analysis obtained is linearly scored between 1 and 9 points given from the sources using the R'WOT Analysis method. According to the results of the R'WOT Analysis, our country's approach to Industry 4.0 is considered as an opportunity of 31%.

Keywords: Industry 4.0, SWOT Analysis, R'WOT Analysis, Ranking Technique, Linear Combination Technique

REFERENCES

[1] Adıgüzel, Ö. Y. (2016). Endüstri 4.0 ve Nesnelerin İnterneti.

[2] SIEMENS. (2016). Endüstri 4.0 Olma Yolunda: Dijital Fabrikalar.

[3] Aktan, Ç. C. (2003). Değişim Çağında Yönetim. İstanbul Ticaret Odası.

[4] Aktan, Ç. C. (1999, Şubat). 2000'li Yıllarda Yeni Yönetim Teknikleri 2. Türkiye Genç İş adamları Derneği.

[5] Wheelen, T. L., & Hunger, J. D. (1992). Strategic Management and Business, Policy, Addison-Wesley Publishing Company, Fourth Edition. New York.

[6] Torlak, Ö., Altunışık, R., & Özdemir, Ş. (2002). Modern Pazarlama. İstanbul: Değişim Yayınları.

[7] Yılmaz, E. (2006). R'wot Tekniği; Arıcılık Sektöründekatılımcı Yaklaşım ile Örnek Bir Uygulaması. Çevre ve Orman Bakanlığı, Tarsus. 17.

[8] Dönmez, Y., & Gökyer, E.&Aşkın, K.F. (2015) Safranbolu Yörük Köyü ve Yakın Çevresinin Ekoturizm Potansiyelinin R'WOT Analizi ile Değerlendirilmesi.

[9] Schomoldt, D. L., Peterson, D. L., & Smith, R. L. (1995). The Analytic Hierarchy Process and Participatory Decision Making. 129-1.

Determination of Variables That Affect the Satisfaction Levels of Visiting Tourists By Logistics Regression Analysis

Fatma ATEŞ[1], Nuray TÜRKER[2], Filiz ERSÖZ[3]

[1]Sustainable Management and Planning of Natural Resources Department

[2]Department of Gastronomy and Culinary Arts

[3]Department of Industrial Engineering, Karabük University, TURKEY

Abstract: Tourism in the natural world has an important place. Tourism revenues have a significant share in the development and development of countries. The economic benefit to be achieved; it is possible to maximize the natural, historical and cultural beauties. In this study, which was conducted to determine the satisfaction of tourists, survey was applied to 169 tourists. As a result of analysis; It was concluded that variables such as age, gender, nationality and income did not affect tourist satisfaction, variables related to purchased products, shops and sellers had a significant relationship between tourist satisfaction.

Keywords: Tourism Satisfaction Level, Logistic Regression Analysis

Giriş

Turizm, bugün yabancı para girdisini arttıran ve istihdam sağlama nitelikleriyle ülke ekonomisine katkısı olan, milletlerarası kültürel ve sosyal iletişimi sağlayıcı ve milli ekonomiye faydalı, milletlerarası iletişimi geliştirici etkileriyle küresel barışın muhafaza edilmesine önemli katkısı olan bir faaliyettir.

Turizmi; insanların yaşam alanlarının dışına çıkıp, farklı mekanlara gidip farklı faaliyetlere katılmak isteğiyle, burada konaklayarak gerçekleştirdikleri geçici hareketlilik olarak tanımlayabiliriz. Gerçekleştirilen bu faaliyetler; konaklama süresine, gerçekleştirme amacı ve şekline, ulaşım çeşidine, maliyetine, yolcu sayısına göre çeşitli kategorilere ayrılmıştır 20. Yüzyılın başlarından sonraki dönemlerde hızlı bir ilerleme kaydetmiş ve bütün dünyaya

yayılmıştır. Turizm küresel ekonomiye yön veren bir sektör olarak karşımıza çıkmakta ve her geçen gün önemi artmaktadır.

Dünya genelinde, sürekli değişen tüketici ihtiyaçlarına bağlı olarak farklılık gösteren turizm hareketleri ve turistlerin görüşleri, destinasyonların gelişmişlik seviyelerini de etkilemektedir. Turizm faaliyetlerine katılan bireyler, tatillerini geçirecekleri destinasyonu seçerken, o yerin doğal kaynaklarına, somut ya da somut olmayan kültürel miras varlıklarına, altyapı ve üstyapı olanaklarına, bölge veya yörede doğrudan ya da dolaylı olarak turizm sektörü içerisinde yer alan paydaşların birbirleriyle olan ilişkilerine ve sunulan hizmetlerin kalitesi gibi çeşitli unsurlara değişik seviyelerde önem vermektedirler. Potansiyel turistlerin, tatil öncesi önem verdikleri bu unsurlardan, tatilleri sonrası memnun kalma seviyeleri, gelecek yıllarda tekrar söz konusu destinasyona gitmelerini ya da destinasyonun gönüllü birer pazarlamacısı haline gelmelerinde en kilit noktalardan birisidir [14].

Sürdürülebilirlik prensiplerine bağlı politikalar hayata geçirilemediği takdirde kültürel turizm, bölgelerin doğal ve beşeri turizm varlıkları üzerinde telafisi olmayan etkiler bırakacaktır. Kültür turizmini geliştirmek ve gelecek nesillere de miras bırakmak için atılacak adımlar, mahalli ve folklorik değerlerin yok olmasını engelleyecektir.

Ülkemiz coğrafi konum olarak eski dünya karalarının bitişme noktasında bulunur. Bu durumda karalar arasındaki etkileşimlerin ülkemiz topraklarında gerçekleşmesi olağandır. Aynı zamanda Türkiye, medeniyetler arasında cereyan eden etkileşiminle birçok medeniyete ev sahipliği yapmıştır. Çeşitli savaşlara, göçlere ev sahipliği etmiş halen de bulunduğu coğrafyanın handikaplarını kültürel ve ekonomik olarak hissetmektedir. Doğal ve kültürel zenginlikleriyle ciddi anlamda turizm potansiyeline sahip bir ülkedir. Sahip olduğumuz zenginliklerin sürdürülebilir kulllanımını sağlamak bu potansiyeli etkin ve verimli kullanmanın yegâne koşuludur.

Türk ekonomisinin de vazgeçilmez temel taşlarından birisi olan turizm, bugünkü dış ticaret açığına, enflasyona ve işsizliğe çare arayan hükümetlerin önemle üzerinde durduğu bir konudur [9]. Kılıç ve Pelit (2004); Seddighi ve Theocharous (2002) atfen turizm faaliyetlerinin bir bölgede ya da destinasyonda gelişebilmesinde çeşitli unsurlar etkili olmadığını, bölgenin doğal yapısı, sosyal yapısı, alt ve üst yapısı, ulaşımı, bölgedeki yerel yönetimlerin tavrı, halkın

turizme ve turistlere bakış açısı, bölgede faaliyet gösteren işletmelerin tutumları bölge turizminin gelişmişliğini etkileyen önemli unsurlar olduğunu ifade etmiştir [8].

UNESCO Dünya Miras Listesi'nde ''En iyi korunmuş 20 şehir'' arasında bulunan Safranbolu, gerek Osmanlı-Türk mimarisiyle inşa edilmiş meskenleri gerekse; hamam, han ve çeşmelerin meydana getirdiği kültürel atmosferiyle yerli yabancı çok sayıda turiste ev sahipliği yapmaktadır. Kültür turizmi açısından ülkemizin en fazla ziyaret edilen şehirlerinden bir tanesidir. Burada yaşayanlar insanlarında önemli bir kısmı geçimini turizmden elde etmektedir. Kentte birçok hediyelik eşya dükkânı, yeme içme yerleri ve eğlence merkezleri bulunur. Dönemsel olarak maddi getirisinde değişikler olmakla birlikte özellikle yabancı turistlerin sık ziyaret ettiği aylarda esnaflar tarafından önemli miktarda kazanç elde edilmektedir. Belediye tarafından düzenlenen festivaller ve etkinliklerinde turizm gelirleri üzerindeki payı büyüktür.

Literatür Araştırması

Müşteri Memnuniyetini Oluşturan Faktörlerin Müşteri Sadakatine Etkisinin Lojistik Regresyon Analizi ile İncelenmesi adlı [18] çalışmada, İzmir ve Afyonkarahisar'daki beş yıldızlı termal otel işletmelerinde konaklayan müşterilerin memnuniyetlerini oluşturan faktörlerin sadakatleri üzerindeki etkisinin lojistik regresyon analizi ile belirlenmesidir. Bu kapsamda İzmir ve Afyonkarahisar'daki beş yıldızlı termal otel işletmelerinde konaklayan 423 müşteriye anket uygulaması yapılarak veriler elde edilmiştir. Verilerin analizi sonucu, müşteri memnuniyeti ile müşteri sadakati arasındaki pozitif yönlü kuvvetli bir ilişki olduğu tespit edilmiştir. Uygulanan lojistik regresyon analizi sonuçlarına göre ise, odalar bölümü hizmetlerinden memnun olan müşterilerin memnun olmayan müşterilere göre sadakat davranışlarının 3,5 kat, yiyecek içecek bölümü hizmetlerinden memnun olan müşterilerin memnun olmayan müşterilere göre 6,7 kat ve genel değerlendirme sonucu memnun olan müşterilerin memnun olmayanlara göre 3,4 kat daha fazla olduğu tespit edilmiştir [18].

E-hizmet kalitesinin incelendiği çalışmada İnternet perakendeciliğinde algılanan hizmet kalitesinin (e-hizmet kalitesi) müşteri memnuniyeti (e-memnuniyet) üzerindeki etkisini tespit etmek ve bu etkinin farklı sektörlere göre nasıl değiştiğini ortaya koymak amaçlanmaktadır. Bu kapsamda iki farklı sektör (hazır giyim ve kitap) seçilmiş ve sektördeki en büyük iki rakip markanın müşterileri araştırma kapsamına alınmıştır. İnternet üzerinden anket yöntemi ile 590

kişiden elde edilen verilere regresyon analizi uygulanmıştır. Araştırmada e-hizmet kalitesi algısının e-memnuniyet üzerinde anlamlı bir etkiye sahip olduğu bulunmuştur. e-hizmet kalite algısında önemli unsurlardan olan "gizlilik" ve "teknik" boyutlarının, e-memnuniyeti açıklamakta anlamlı bir etkisinin olmadığı, "etkinlik", "işlem gerçekleştirme", "müşteri hizmetleri", "tasarım", ''eğlence" boyutlarının etkilerinin ise anlamlı olduğu bulunmuştur. En önemli etkinin ise "işlem gerçekleştirme" boyutunda olduğu, en az etkinin ise "tasarım" boyutunda olduğu saptanmıştır. Ayrıca e-hizmet kalitesinin e-memnuniyet üzerindeki etkilerinde sektörlere göre kısmi farklılıkların olduğu tespit edilmiştir [6].

Müşteri memnuniyetine ilişkin çalışmada, algılanan fiyat, algılanan değer ve algılanan faydanın tekrar satın alma eğilimi üzerinde doğrudan ve dolaylı etkileri araştırılmıştır. Çalışmanın verileri şehirlerarası yolcu taşıyan bir firmanın 182 müşterisinden elde edilmiştir. Değişkenler arasındaki ilişkiyi bulabilmek için yapısal eşitlik modeli kullanılmıştır [20].

Müşteri memnuniyetine ilişkin bir diğer çalışmada çeşitli demografik faktörlerden nasıl etkilendiğini analiz etmek amacıyla yapılmıştır. Araştırma, Ankara ilinde bulunan bir kamu hastanesinden sağlık hizmeti alan hastalar üzerinde yapılmış olup tanımlayıcı niteliktedir. Araştırma kapsamındaki hastaların müşteri memnuniyetini (hasta memnuniyeti) ölçmek için T.C. Sağlık Bakanlığı'nın resmi internet sitesinde yer alan memnuniyet anketleri yapılmıştır. 148 adet anket formu değerlendirilmiştir [7].

Tramvay Yolcu Memnuniyetinin Lojistik Regresyon Analiziyle Ölçülmesi: Estram Örneği adlı çalışmasında; toplu taşıma araçlarından biri olan tramvaya yönelik yolcu memnuniyeti, Eskişehir tramvay sistemi (Estram) örneğinde, Binominal Lojistik Regresyon Analizi ile incelenmektedir. İki üniversiteye sahip olan Eskişehir'de öğrenci nüfusun fazla olması ve tramvay için önemli bir yolcu kitlesi olacakları düşüncesiyle çalışma, her iki üniversiteden basit tesadüfî örnekleme yoluyla seçilen 300 öğrenci üzerinde gerçekleştirilmiştir. Öğrencilerin memnuniyetleri ile ilgili binominal düzeyde, gizil bir değişken kullanılmıştır. Uygulanan Binominal Lojistik Regresyon Analizi sonucunda; öğrencilerin Estram'dan memnuniyetleri üzerinde modele alınan tüm bağımsız değişkenlerin negatif etkileri olduğu belirlenmiştir [12].

Müşteri Memnuniyetinin tahmini için yapılan çalışmada Yapay Sinir Ağları, Lojistik Regresyon ve Ayırma Analizinin performansları karşılaştırılmıştır. Veriler Uşak'taki kamu hastanelerinde 2007 yılında yapılan hasta memnuniyetini ölçmeyi amaçlayan bir anket uygulanarak elde edilmiştir ve 364 hastayı kapsamaktadır. Sonuçlar Yapay Sinir Ağlarının diğer yöntemlere göre müşteri memnuniyetini daha iyi tahmin ettiğini göstermiştir [21].

Turist memnuniyetine ilişkin yapılan çalışmada, Alanya ilçesinin Türk Turizmi içerisindeki önemini belirtmek, dış turizmin bölgede meydana getirdiği ekonomik etkileri açıklamak ve konaklama işletmelerinin turist memnuniyeti üzerindeki etkisi incelenmiştir. Araştırmada yerli ve yabancı turistlere 27 sorudan oluşan bir anket uygulanmıştır. Toplam 104 anket değerlendirmeye alınmış ve SPSS 17 programında güvenilirlikleri kontrol edilip, faktör analizi, korelasyon analizi, regresyon analizi yöntemleri ile incelenmiştir [11].

Marina işletmelerinde yapılan bir diğer çalışmada, ilişkisel pazarlama uygulamalarının tekrar satın alma niyeti, tavsiye etme niyeti ve yönetimden memnuniyet düzeyi üzerinde bir etkisinin bulunup bulunmadığı belirlemek ve marina yöneticilerine bu konuda yol gösterecek katkı ve önerilerde bulunulmuştur. Bu amacı gerçekleştirmek üzere Antalya Bölgesi' nde faaliyet gösteren bir marina işletmesinin 78 müşterisine yüz yüze anket uygulaması gerçekleştirilmiş ve veriler SPSS paket programıyla analiz edilmiştir [3].

Materyal ve Yöntem

Safranbolu'yu ziyaret eden turistlerin memnuniyet düzeylerini belirlemek amacıyla Çinli ve Tayvanlı olmak üzere 169 yabancı turiste anket uygulanmıştır. Bu amaca yönelik memnuniyetin hangi değişkenlere bağlı olarak ortaya çıktığını görebilmek için lojistik regresyon analizinden yararlanılmıştır.

Lojistik regresyon; cevap değişkenin (y) kategorik olarak, ikili (binary, dishotomous) ve çoklu (multinominal) kategorilerde gözlendiği durumlarda açıklayıcı değişkenlerle (xi, =1,2,…,k) sebep sonuç ilişkisini belirlemede yararlanılan bir yöntemdir. Cevap değişkenin (y) değişimi üzerinde etkili olan açıklayıcı değişkenlerin/risk faktörlerinin (x,) etki büyüklüklerini (Odds Radio) belirlemeyi sağlayan bir yöntemdir. Ayrıca Lojistik regresyon, açıklayıcı değişkenlere/risk faktörlerine göre cevap değişkenin beklenen değerlerinin olasılık olarak elde edildiği sınıflama ve atama işlemi yapmaya yardımcı olan bir yöntemdir [17].

Lojistik Regresyon (LR) Analizi; sürekli ya da kesikli ya da her ikisinin bir arada bulunduğu sürekli ve kesikli değişkenler setinden yola çıkarak kesikli bir sonucun yordanmasını sağlar. Sosyal bilimlerde sıklıkla kullanılan çok yönlü frekans analizi ve çoklu regresyon analizlerinden farklı olarak LR analizi, bağımlı değişkenin kesikli olabilmesine imkân tanımaktadır. Bu bakımdan sonuçların kategorik olduğu durumlar için bu analizin yaygın kullanımı söz konusudur.

Çok yönlü frekans analizinde yordayıcıların kesikli olması gerekirken LR analizinde böyle bir zorunluluk bulunmamakta, yordayıcılar sürekli olabildiği gibi kesikli de olabilmekte ya da hem sürekli hem kesikli yordayıcılar bir arada bulunabilmektedir. Çoklu regresyon analizi ile karşılaştırıldığında ise LR analizi negatif kestirim olasılıkları üretmemektedir. Ayrıca LR analizi, çoklu regresyon analizi ve diskriminant analizinden farklı olarak yordayıcı değişkenlerin dağılımlarına yönelik varsayımlara sahip değildir. Bu bakımdan LR analizi, bahsedilen diğer yöntemlere göre çok daha esnektir [23].

Verilerin analiz edilmesinde multinominal lojistik regresyon analizinden yararlanılmıştır. Çalışmada kullanılan verilerin analiz edilmesinde IBM SPSS 17.0 paket programı kullanılmıştır.

Lojistik Regresyon Analizi, bağımlı değişkeni iki veya ikiden çok kategoriye sahip olan bir denklemde, bağımsız değişken veya değişkenler ile bağımlı değişken arasındaki ilişkiyi ifade etmekte kullanılan bir yöntemdir. Ancak burada Regresyon Analizinden farklı olarak bağımlı değişkenin kategorilere sahip olması sebebiyle bağımsız değişkenin bağımlı değişken üzerindeki etkisi olasılık olarak ifade edilir [19]. Lojistik regresyon analizinin temel odağı, bireylerin hangi grubun üyesi olduğunu kestirmede bir regresyon denklemi oluşturmaktır. Bu çalışmada, iki kategorili (ikilem/dichotomous/binary) bağımlı değişken olarak ifade edilen belirli gruplara üye olma durumunu en iyi açıklayan bağımsız değişkenler kombinasyonunu belirlemeye yönelik ikili lojistik regresyon analizinin (binary logistic regression analysis) temel kavram ve süreçlerini açıklamak amaçlanmaktadır [10].

Herhangi bir regresyon modelinde E(Y/x) ifadesi koşullu ortalama değerini göstermektedir. Bu ifade, bağımsız değişken (x) verildiğinde bağımlı değişkenin (Y) ortalama değerini göstermektedir [13].

$E(Y/x)= \beta_0 + \beta_1 X$

Lojistik regresyonda ise x verildiğinde Y'nin koşullu dağılım ortalaması $\pi(x) = E(Y/x)$ olarak gösterilir. Lojistik regresyon modelinin özel durumu;

$$E(Y/x) = \pi(x) = \frac{e\beta 0+\beta}{1 + e\beta 0+\beta 1x}$$

şeklindedir. Yukarıdaki eşitliğe lojit dönüşüm uygulandığında aşağıdaki eşitlik elde edilir [13].

$$g(x) = \ln \left[\frac{\pi(x)}{1 - \pi(x)}\right] = \beta_0 + \beta_1 X$$

Logit değeri ile bağımsız ve bağımlı değişkenler arasında doğrusal bir ilişki vardır. Bu değer -∞ ile +∞ arasında olabilir [4].

İkili lojistik regresyon modelinde olayın gerçekleşme ve gerçekleşmeme durumu 0 ve 1 olmak üzere iki durum söz konusudur. Olayın gerçekleşme olasılığının, gerçekleşeme olasılığına oranı odds oranı olarak tanımlanmaktadır. Odds oranı 0 ile +∞ arasında değer alabilmektedir [1].

$$\text{Odds Oranı} = \frac{\pi(x)}{1-\pi(x)} = e\, \beta_0 + \beta_1 X$$

Lojistik regresyon analizinde regresyon katsayılarının tahmininde en küçük kareler yöntemi yerine, en çok olabilirlik yöntemi kullanılmaktadır. En çok olabilirlik yönteminde bir olayın olma olasılığının maksimum olması istenir [2].

Verilere basit ya da çoklu regresyon analizlerinin uygulanabilmesi için değişkenlerin bazı koşullara (varsayımlara) uyması gerekir. Bu koşulların sağlanamadığı veri setlerine basit ya da çoklu regresyon analizleri uygulanamaz. Lojistik regresyon analizi ise Normal Dağılım Varsayımı, süreklilik varsayımı ön koşulu gerektirmeyen bir regresyon yöntemidir [17].

Uygulama

Bu bölümde, kullanılan veriler, çalışmanın amacı ve lojistik regresyon analizi sonucunda elde edilen bulgular yorumlanmaya çalışılmıştır. Bireyin eğitim durumunun, istihdam durumunun, yaşının, sağlık ve hastalık durumunun maddi yoksunluğu anlamlı bir şekilde etkileyip etkilemediği belirlenmek istenmiştir. Bunların yanı sıra konutun oda sayısının, ısıtma

sisteminin maddi yoksunluğu belirlemede anlamlı bir etkisi olup olmadığı tespit edilmeye çalışılmıştır. Bu varsayımlar ışığında Lojistik regresyon analizi kurularak sonuçlar değerlendirilmiştir.

Araştırmanın temel hipotezi: ‘’Bireyin milliyeti, cinsiyeti, geliri, yaşı, ürünün kalitesi, ürünün çeşitliliği, satıcının sunduğu hizmet ve indirimli fiyat değişkenleri memnuniyet düzeyini anlamlı bir şekilde etkileyeceği’’dir.

Alt hipotezler ise;

H_1: Milliyet değişkeni memnuniyet düzeyini belirlemede anlamlı bir etkiye sahiptir.

H_2: Cinsiyet değişkeni memnuniyet düzeyini belirlemede anlamlı bir etkiye sahiptir.

H_3: Gelir değişkeni memnuniyet düzeyini belirlemede anlamlı bir etkiye sahiptir.

H_4: Yaş değişkeni memnuniyet düzeyini belirlemede anlamlı bir etkiye sahiptir.

H_5: Ürünün kalitesi değişkeni memnuniyet düzeyini belirlemede anlamlı bir etkiye sahiptir.

H_6: Ürünün çeşitliliği değişkeni memnuniyet düzeyini belirlemede anlamlı bir etkiye sahiptir.

H_7: Satıcının sunduğu hizmet değişkeni memnuniyet düzeyini belirlemede anlamlı bir etkiye sahiptir.

H_8: İndirimli fiyat değişkeni memnuniyet düzeyini belirlemede anlamlı bir etkiye sahiptir.

Çalışmada ele alınan değişkenler ve alt kategoriler Tablo 1’de sunulmuştur.

Tablo 1: Araştırmada Kullanılan Değişkenler

Bağımsız Değişkenler	Alt kategoriler
1. Milliyet	1. Çin
	2. Tayvan
2. Cinsiyet	1. Kadın
	2. Erkek
3. Gelir	1. 1000 TL’den az
	2. 1001-1500 TL
	3. 1501-2000 TL
	4. 2001-2500 TL

	5. 2501-3000 TL
	6. 3001-3500 TL
	7. 3501-4000 TL
	8. 4001-4500 TL
	9. 4501-5000 TL
	10. 5000 TL ve üzeri
4. Yaş	1. 20 yaş ve altı
	2. 21-30 yaş
	3. 31-40 yaş
	4. 41-50 yaş
	5. 51-60 yaş
	6. 61 yaş ve üstü
5. Ürün kalitesi	1. Hiç memnun değilim
	2. Memnun değilim
	3. Kısmen memnunum
	4. Memnunum
	5. Çok memnunum
6. Ürün çeşitliliği	1. Hiç memnun değilim
	2. Memnun değilim
	3. Kısmen memnunum
	4. Memnunum
	5. Çok memnunum
7. Satıcının sunduğu hizmet	1. Hiç memnun değilim
	2. Memnun değilim
	3. Kısmen memnunum
	4. Memnunum
	5. Çok memnunum
8. İndirimli fiyat	1. Hiç memnun değilim
	2. Memnun değilim
	3. Kısmen memnunum
	4. Memnunum
	5. Çok memnunum

Ele alınan bağımsız değişkenler yardımıyla memnuniyet düzeyi bağımlı değişkenini belirlemek amacıyla Lojistik Regresyon Modeli kurulmuştur. Öncelikle modelin genel anlamda uygunluğunu gösteren Omnibus test sonuçları sunulmuştur.

Tablo 2: Model Katsayılarının Genel Testi

	Ki-Kare	sd	Sig
Adım 1 Adım	158,636	19	,000
Blok	158,636	19	,000
Model	158,636	19	,000

Bütün değişkenler modele eklendikten sonra model uyum iyiliği için Ki-Kare değerine bakılır. Modelin genel anlamlılığının, yani uyum iyiliğinin istatistiksel olarak anlamlı olduğu görülmektedir ($p<0,01$). Modelin verilere uygunluğunu ve modelin genel uyumunu gösteren R2 değeri Tablo 4'te verilmiştir.

Tablo 3: Hosmer ve Lemeshow Testi

Adım	Ki-Kare	df	Sig
1	13,680	8	,091

Tablo 4: Model Özeti

Adım	-2Log likelihood	Cox & Snell R Square	Nagelkerke R Square
1	39,237	,609	,883

Hosmer ve Lemeshow test sonucuna göre; tahmin edilen lojistik regresyon modelinin verilere uygun olduğu ($p=0.091$) görülmüştür. Cox & Snell R^2 ve Nagelkerke R^2 değerleri, model tarafından bağımlı değişkende açıklanan varyansın büyüklüğünü göstermektedir.

Modelin genel uyumunun iyi olduğu (Cox & Snell $R^2=0,609$; Nagelkerke $R^2=0,883$) belirlenmiştir. Maddi yoksunluğun toplam değişimin %88'i ele alınan bağımsız değişkenler tarafından açıklanmaktadır. İkili Lojistik regresyon analizine ait katsayı tahminleri ve odds oranları Tablo 5'da sunulmuştur.

Tablo 5: Model Tahmin Sonuçları

Bağımsız Değişken	**Kategoriler**	**Score**	**df**	**Sig**
Milliyet		,000	1	,989
Cinsiyet		,000	1	,989
Gelir	1000 TL'den az	4,983	9	,836
	1001-1500 TL	,148	1	,701
	1501-2000 TL	1,136	1	,287
	2001-2500 TL	,487	1	,485
	2501-3000 TL	,996	1	,318
	3001-3500 TL	,117	1	,732
	3501-4000 TL	,376	1	,540

	4001-4500 TL	1,142	1	,285
	4501-5000 TL	1,142	1	,285
	5000 TL ve üstü	,021	1	,885
Yaş	20 yaş ve altı	4,850	5	,434
	21-30 yaş	3,011	1	,083
	31-40 yaş	,323	1	,570
	41-50 yaş	1,493	1	,222
	51-60 yaş	,016	1	,898
	61 yaş ve üstü	,577	1	,448
Ürün Kalitesi		62,334	1	,000
Ürün Çeşitliliği		80,577	1	,000
Satıcının Sunduğu Hizmet		65,587	1	,000
İndirimli Fiyat		46,069	1	,000

$p<0,05$

Turist memnuniyetini etkileyen değişkenlerin analizi yapıldığında Tablo 5'e göre memnuniyeti etkileyen değişkenlerden ürün kalitesi, ürün çeşitliliği, satıcının sunduğu hizmet ve indirimli fiyat ile memnuniyet arasında anlamlı bir ilişki bulunmaktadır. Bu sonuca göre alt hipotezlerden H_5, H6, H7, H_8 'in gerçekleşmiş olduğu söylenebilmektedir.

Tablo 6: Sınıflandırma Tablosu

	Tahmin Edilen		
	Memnun Değil	Memnun	Doğruluk (%)
Memnun Değil	0	46	,0
Memnun	0	123	100,0
Genel %			72,8

Kurulan lojistik regresyon modelinin sınıflandırma tablosu Tablo 6'da verilmiştir. Testin duyarlılık oranı %100 olarak elde edilmiştir. Bu durum gerçek durumda turistlerin memnun olanlarının %100'ünün doğru olarak tahmin edildiğini göstermektedir. Modelin doğru sınıflandırma oranı anket uygulanan turistlerin %72,8'inin memnuniyet durumunu doğru tahmin ettiğini göstermektedir. Modelin sınıflandırma gücünün oldukça iyi olduğu söylenebilir.

Sonuç ve Öneriler

Turistlerin milliyetinin, cinsiyetinin, yaşının, maddi gelirinin, ürün kalitesinin, ürün çeşitliliğinin, satıcının sunduğu hizmetin ve indirimli fiyatın turist memnuniyetlerini anlamlı bir şekilde etkileyip etkilemediği belirlenmek istenmiştir. Bu varsayımlar ışığında lojistik regresyon analizi kurularak sonuçlar değerlendirilmiştir.

Bu çalışmada Safranbolu'yu ziyaret eden turistlerin memnuniyet düzeylerini etkileyen değişkenlerin neler olduğunu ortaya koymak amaçlanmıştır. Yapılan analizler neticesinde turist memnuniyetini etkileyen faktörlerden ürün kalitesi, ürün çeşitliliği, satıcının sunduğu hizmet ve indirimli fiyat değişkenlerinin anlamlı olduğu anlaşılmıştır.

Yerel ekonominin atılımı, herhangi bir bölgeye has değişim süreçlerinin motivasyonudur şeklinde açıklanabilir. Turizm, ekonominin önemli sektörlerinden biridir. Geçmişten günümüze kadarki süreçte Safranbolu kültürel özellikleri, tarihi ve otantik kendine has mimarisiyle her daim yerli ve yabancı turistlerin ziyaret ettiği önemli merkezlerden biri olmuştur. Şehir, 1994 yılında UNESCO Miras Listesine dahil edilerek koruma altına alınmıştır. Uluslararası koruma statüsü kazanarak tanınırlığı da artan Safranbolu'yu görmek için gelen turist sayısında da artış görülmüştür. Safranbolu'nun şehir kimliğinin vurgulanarak turizme katkı sağlanması, iç turizmde hatrı sayılır derecede isim yapmasına vesile olmuştur.

Sürdürülebilir turizm uygulamaları, yerel ekonomi canlandırılabilmekte ve yabancı turist çekiciliğinin yanı sıra iç turizmi de hareketlendirebilmektedir. Bu tür çalışmalarla turist memnuniyetinin hangi faktörlere göre değişiklik gösterdiği belirlenerek tarihi ve kültürel zenginliklerimiz için planlanan politikaların o yönde geliştirmek hem elde edilen turizm gelirleri için hem de yörede yaşayan insanların ve çevrenin kalkındırılması açısından isabetli olacaktır.

KAYNAKÇA

[1] Agresti, A. (1996). An Introduction To Categorical Data Analysis (Vol. 1350), New York: Wiley.

[2] Alpar, R. (2013). Çok Değişkenli İstatistiksel Yöntemler (4. Baskı), Detay Yayıncılık, Ankara.

[3] Arlı, E. (2013) Marina İşletmeciliğinde İlişkisel Pazarlama Uygulamalarının Tekrar Satın Alma Niyeti, Tavsiye Etme Niyeti ve Memnuniyet Üzerindeki Etkisi. Anadolu Üniversitesi Sosyal Bilimler Dergisi.

[4] Bayram, N. (2004). Multinominal Lojistik Regresyon Analizinin İstihdamdaki İşgücüne Uygulanması, İktisat Fakültesi Mecmuası, 54(2): 61- 75.

[5] Bilgin, Y. (2017) Restoran İşletmelerinde Hizmet Kalitesi, Müşteri Memnuniyeti ve Müşteri Sadakatinin Ağızdan Ağıza Pazarlamaya Etkisi. İşletme Araştırmaları Dergisi, Bartın Üniversitesi İktisadi ve İdari Bilimler Fakültesi, Turizm İşletmeciliği Bölümü, Bartın.

[6] Bozbay, Z., Yaman, Y., & Özkan, E. (2016). The Role of Service Quality on Customer Satisfaction in Internet Retailing: A Comparative Study of Apparel and Book Industries. Journal of Transportation and Logistics, 1(1), 19–19.

[7] Cihangiroğlu, N. Uzuntarla, Y. (2016) Müşteri Memnuniyetinin Çeşitli Demografik Özellikler Açısından Analizi: Bir Kamu Hastanesi Örneği.

[8] Çakıcı, A, C. Aksu, M. (2006) Gökçeada'ya Gelen Turistlerin Beklenti ve Tatmin Düzeylerinin Karşılaştırılması. İşletme Fakültesi Dergisi.

[9] Çımat, A., & Bahar, O. (2003). Turizm Sektörünün Türkiye Ekonomisi İçindeki Yeri ve Önemi Üzerine Bir Değerlendirme. Akdeniz İİBF Dergisi, (6), 1–18.

[10] Çokluk. Ö, ''2010'', Lojistik Regresyon Analizi: Kavram ve Uygulama. Ankara Üniversitesi Eğitim Bilimleri Fakültesi, Ölçme ve Değerlendirme Bölümü, Eğitim İstatistiği ve Araştırma Anabilim Dalı, Ankara.

[11] Demir, V. Altındağ, E. (2017) Konaklama İşletmelerinde Turist Memnuniyetinin Değerlendirilmesi: Alanya İlçesi Örneği. Turizm ve Araştırma Dergisi

[12] Girginer, N., & Cankuş, B. (2008). Tramvay Yolcu Memnuniyetinin Lojistik Regresyon Analiziyle Ölçülmesi: Estram Örne ğ i. Yönetim ve Ekonomi, 15(1), 181–193.

[13] Hosmer, D. W. ve Lemeshow, S. (2000). Applied Logistic Regression., 2nd edn.(Wiley: New York.). NY, USA.

[14] İpar, M. S. Doğan, M. (2013) Destinasyonun Turist Açısından Önem-Memnuniyet Modeli ile Değerlendirilmesi: Edremit Üzerine Bir Uygulama. Adıyaman Üniversitesi Sosyal Bilimleri Enstitüsü Dergisi, Adıyaman.

[15] Karcı, Z. Bayram Arlı, N. (2018), Maddi Yoksunluğu Etkileyen Değişkenlerin Regresyon Analizi ile Belirlenmesi. Süleyman Demirel Üniversitesi, İktisadi ve İdari Bilimler Fakültesi Dergisi, Isparta.

[16] Meydan Uygur, S. Baykan, E. (2007), Kültür Turizmi ve Turizmin Kültürel Varlıklar Üzerindeki Etkileri. Gazi Üniversitesi, Ticaret ve Turizm Eğitim Fakültesi, Ticaret ve Turizm Eğitim Fakültesi Dergisi, Ankara.

[17] Özdamar, K. (2009), Paket Programlar ile İstatistiksel Veri Analizi-1. Kaan Kitabevi, 7.Baskı, s.571. Eskişehir.

[18] Soybalı, H. H., & Cengiz Mutlubaş, I. (2017). Müşteri Memnuniyetini Oluşturan Faktörlerin Müşteri Sadakatine Etkisinin Lojistik Regresyon Analizi ile İncelenmesi (The Investigation of Effect of Customer Satisfaction Factors on Customer Loyalty with Logistics Regression Analysis). Turk Turizm Arastirmalari Dergisi, 1(3), 1–15.

[19] Şentürk. E, ‘’2011’, Mutluluk Düzeyinin Sosyo-Demografik Özelliklerle Lojistik Regresyon Analizi Aracılığıyla İncelenmesi ve Türkiye İçin Bir Uygulama. Marmara Üniversitesi, Sosyal Bilimler Enstitüsü, Ekonometri Anabilim Dalı, İstatistik Bilim Dalı, İstanbul.

[20] Tayyar, N. Bektaş, Ç. (2009) Algılanan Değer ve Müşteri Memnuniyetinin Tekrar Satınalma Eğilimi Üzerine Etkisi.

[21] Tayyar, N. (2010) Müşteri Memnuniyeti Tahmininde Yapay Sinir Ağları, Lojistik Regresyon ve Ayırma Analizinin Performanslarının Karşılaştırılması. Süleyman Demirel Üniversitesi, İktisadi ve İdari Bilimler Fakültesi Dergisi, Isparta.

[22] Yıldız, B. Çiğdem, Ş. (2018) Havayolu Hizmet Kalitesinin Müşteri Memnuniyeti Üzerindeki Etkisinin Yapısal Eşitlik Modeli ile Analizi. Bingöl Üniversitesi Sosyal Bilimler Enstitüsü Dergisi, Bingöl.

[23] http://www.akademikistatistik.com (Erişim: 15.12.2018).

Demand Estimation of Wood Quantity Used in Wood Industries

Sercan ODABAŞ1, Taner ERSÖZ

1Department of Industrial Engineering, 2Department of Actuary and Risk Management, Karabük University, TURKEY

Abstract: With the rapid development of technology, competition in the markets is increasing. In the conditions of rising competition, it is inevitable for companies to make predictions for the future. The accuracy of the forecasts will also facilitate the decision-making of companies in the future. Demand forecasts are very important for every company considering these conditions. Demand forecasts play an active role in tactical and strategic decisions taken on the administrative side of companies to achieve their near or far-term goals, bringing the company's assets closer to optimal and thus ensuring the link between the targets set and the operations implemented. According to the data obtained from the firm of Antik Atölye operating in Düzce, the regression analysis method and weighted moving average method were compared. For both methods, estimation modeling was established and implemented. As a result for the Antik Atölye company operating in the wood industry regression analysis method was more effective than weighted moving average method.

Keywords: Forecasting, Regression Analysis, Ağırlıklı Hareketli Ortalama Yöntemi

Giriş

Ahşap ürünler, dünyada yapı malzemeleri başta olmak üzere kapılarda, yer döşemelerinde, pencerelerde, çatılarda, masa ve koltuk takımlarında yoğun bir şekilde kullanılmaktadır. Aynı zamanda kullanım yerlerine göre iç – dış mekan mobilyaları, hareketli veya sabit, sökülebilir, modüler, fonksiyonel, lamine, modern, klasik mobilyalar olarak gruplara ayrılır. İnsanoğlunun varoluşundan bu yana sürekli hayatın içerisinde olan ahşap oldukça yüksek katma değere sahip olup önemli ölçülerde istihdam sağlamaktadır. Ülkemizde ahşap sektörü üç milyar doların üzerinde katma değer oluşturan bir potansiyele sahiptir. Gelişen teknolojinin sektöre etkisi kaçınılmaz olup söz konusu sektörde birçok teknolojik cihazlar tezgahlar ve yazılımlar

kullanılmaktadır. Bütün veya parçalar halinde yapılan üretim sonucunda görselliği ve dayanıklılığı artırmak için metal (demir ayak ve tutucular vb.) ve kimyasal (epoksi vb.) malzemeler kullanılmaktadır.

Hızla gelişmekte olan ahşap sektörü ülkemizde diğer Avrupa ülkeleri ile kıyaslandığında önemli bir yere sahiptir. Teknoloji yoğun olmasından ziyade emek yoğun olan bir sektör olması ülkemizi Avrupa ülkeleri karşısında avantajlı konuma getirmektedir. Hammadde ve işçiliğin Avrupa'ya kıyasla daha ucuz olması imalat yapılabilmesi kısıtları düşünüldüğünde ihracat dengesine büyük katkı sağlamaktadır.

21. yüzyılın enerji ve çevre yüzyılı olduğu değerlendirildiğinde dayanıklılığı ve sağladığı istihdam ile ahşap sektörü gelişimini devam ettirmektedir. Ahşap üretimin hızla büyümesi ekonomik büyümeyi de olumlu etkiletmektedir. Avrupa'ya olan coğrafi yakınlık ve kolaylıkla ulaşım sağlanan kaliteli hammadde kriterleri doğrultusunda kaliteli işçilik ile mükemmele yakın üretim yapılması her geçen gün ihracat rakamlarımızı geliştirmekte ve ekonomik büyümeye katkı sağlamaktadır.

Ahşap sektörü ihracat rakamlarını 90'lı yıllar baz alındığında 2000 yılı başlarında %16, 2010 yılı başlarında %52, 2017 yılı itibari ile de %75 lik artış göstererek ekonomik büyümeye katkı sağlamıştır.

İmalat sanayi grubu; orman ürünleri sanayisi, ağaç ve mantar ürünleri ile mobilya sanayisinden oluşmaktadır. Ağaç ve mantar ürünleri ara malı üreten sanayiler arasında tanımlanırken, mobilya sanayisi ise tüketim malı üreten sanayiler grubunda anılmaktadır [1]. Orman sanayinin alt grubu olarak görülen ahşap sanayisi atölye tipli ve fabrikasyon halinde iki grup olarak imalat sanayi üretimi içerisinde % 4'lük pay ile yer almaktadır. Ahşap ürünlerine ilişkin talep her geçen gün artmaktadır. 90'ların sonunda ihracat payı 11 milyon dolar olan ahşap sektörü 2000'li yılların başında bu payı 180 – 200 milyon dolara kadar çekmiş 2010 yılında 573 milyon doları yakalamış ve 2018 yılında bu rakam 826 milyon dolarlık paya ulaşmıştır. Türkiye içi piyasada ise 2010 yılının başında ahşap ürüne olan talep 2,6 milyon iken bu rakam 2017 yılı sonunda 5,5 milyona ulaşmıştır.

Talep; söz konusu ürün grubunda tüketici kesimin bir hizmet veya ürünü belirli bir fiyat aralığından almaya hazır bulundukları miktardır (Tekin, 1996). Talep tahmini ise; belirlenen

zaman periyotları içerisinde gelecekte firmaların söz konusu ürün veya hizmeti için oluşacak talebin belirlenmesidir (Acar, 1989). Talepler aynı zamanda satış değerlerini de göstereceği için literatürde 'talep tahmini' yerine 'satış tahmini' de kullanılmaktadır.

Talep tahmini, firma yönetimi için taktik ve stratejik kararların alınmasında, uzun ve kısa dönemli hedeflere ulaşmak amacıyla kullanılabilecek araçların başında yer almaktadır. Talep tahmininin etkinliği şirket fonksiyonlarında optimal kar çizgisine doğru hareket sağlar ve mevcut hedeflerin gereksinimler ile arasında bulunan çatışmayı minimize eder (Bolt, 1994). Talep tahmini çalışmaları her uygulamada olduğu gibi belirli kurallar ve yöntemlerin ışığında yapılır. Fakat seçilen yöntem her ne olursa olsun çıkan sonuçların hepsi kendi özünde doğrudur ancak bu ifade sonuçların talebi %100 olarak açıklayacağı anlamına gelmemektedir. Söz konusu ürün için yapılan talep tahmininde ürünün sahip olduğu özellikler ve seçilen yöntem, tahminlerin tiplerini ve bulunacağı zaman sürelerini etkilemektedir. Ürün için oluşan talepte dalgalanmalar mevcutsa, yapılacak tahmin en az bir dönemi barındırmak zorundadır. Talep için uzun dönemde içinde eğilim söz konusu ise, yapılacak tahminin süresi de daha uzun olmalıdır. Mevsimlere göre talebinde değişiklik meydana gelen ürünlerin talep tahmini yapılırken değişimlerin sebeplerini görebilmek adına mevsimsel tahmin yöntemleri kullanılmalıdır (Acar, 1999).

Tahmin; incelenen olgunun geçmişinden edinilen bilgiler ışığında, söz konusu olgunun geleceği hakkında ön görüde bulunmaktır. İncelenen olgunun geçmişi hakkındaki bilgi farklı yollar ile toplanabilir. Bilgi toplanırken başvurulan yollardan biri ham ve sayısal verileri matematiksel yöntemler yardımıyla yorumlamaktır. İkincisi ise konu hakkında uzman görüşlerini ve fikirlerini birleştirerek analizi gerçekleştirmektir (Taşdemir, 2012).

Bu çalışmada; Düzce ilinde ağaç sektöründe üretim faaliyeti gösteren bir işletmede, gelecek yılların tahmin edilmesinde çoklu doğrusal regresyon analizi ve ağırlıklı hareketli ortalama yöntemleri kullanılmış ve sonuçlar karşılaştırılmıştır.

Materyal ve Metod

Bu çalışmada Düzce ilinde faaliyet gösteren Antik Atölye firmasına ait 2015 ile 2018 yılları arasında gerçekleşmiş olan ahşap kullanım verileri kullanılmıştır. Ahşap üretimi kesme,

kurutma, depolama, şerit, planya, CNC, kalınlık, montaj ve boyama kalemlerinden oluşmaktadır.

Tahmin analizinde geçmiş gözlem değerleri kullanılarak, sürecin oluşmasına katkıda bulunan ilişkiler belirlenir ve bu ilişkilerin geleceği nasıl şekillendireceği tahmin edilmeye çalışılır.

Tahmin metotları; nitel ve nicel tahmin metotları olmak üzere ikiye ayrılır. Nitel tahmin yöntemi; tahminde bulunacak kişinin düşüncelerine bağlı, kişiye özgü bir yöntemdir. Nitel tahmin yöntemleri; matematiksel modellere dayanan yöntemlerdir. Nicel tahminde bulunmak için iki temel yaklaşım kullanılmaktadır; sebep sonuç ilişkisine ve zaman serileri analizine dayalı modeller [2]. Nitel tahmin metotlarına; Delphi ve araştırma teknikleri verilebilir. Nicel tahmin metotları ise zaman serine dayalı yöntemler; Hareketli ortalamalar, Üstel düzeltme, Arima, Arch, Garch, Tar ve YSA verilebilir. Nicel tahmin metotlarında sebep sonuç ilişkisine dayalı yöntemler olarak; basit doğrusal regresyon, doğrusal olmayan regresyon, çoklu doğrusal regresyon ve Yapay sinir ağları (YSA) verilebilir.

Tahminin başarısı önemlidir. Anakütle parametresine 'yakınlık' çeşitli ekonometri tahmin yöntemleri ile bulunmuş tahminlerin örnekteki dağılımların ortalaması ve varyansıyla ölçülür. İyi bir nokta tahmincisinin özellikleri: Sapmasızlık, minimum varyans (Etkinlik), en küçük ortalama hata karesi (OHK), tutarlılık ve yeterliliktir [3].

Regresyon analizi, gelecekte oluşacak bilinmeyen olayları mevcut bilinen veriler ile tahmin edilirken kullanılır. Regresyon analizinde bağımlı değişken ile bağımsız değişken veya değişkenler arasında olan ilişki ve etkileşmeyi doğrusal eğri yöntemi ile bir tahminde bulunur (Köse, 2008). Bağımlı ve bağımsız değişkenler arasında olan ilişkiyi gösteren regresyon denklemi ile bu ilişkide bulunan parametrelerin değerleri bulunabilir. Bağımlı değişken üzerinde etkisi olan bağımsız değişkenler tahmin edilerek, söz konusu bağımlı değişken üzerinde ne tür değişimler olacağı, hangi planların uygulanabileceği, bağımsız değişkenlerden hangisi veya hangilerinin daha önemli ve etkili olduğu ortaya çıkarılabilir. Regresyon analizi ile bağımsız değişkenlerin bağımlı değişken üzerinde yapmış olduğu etkilerin ne derecede olduğu, bağımsız değişkenlerde meydana gelecek olan değişimlerin bağımlı değişkeni nasıl

etkileyeceği hesaplanabilir (Çağlar, 2007). Bununla beraber istatistiksel yöntemlerle bağımsız değişkenlerin bağımlı değişkenler üzerinde yaptıkları etkiler gösterilmektedir (Yoldaş, 2006).

Bağımlı bir değişken ile ona etki eden bağımsız değişkenleri inceleyen regresyon analizinde farklı yöntemler mevcuttur. Çünkü bağımlı ve bağımsız değişkenler arasında bulunan ilişki eğrisel veya doğrusal olabilmektedir (Aksoy, 2008). Regresyon analizi değişkenler arasında bulunan ilişkiyi incelerken şu yöntemleri kullanır;

Regresyon analizi, zaman serileri tahmin modellerinin kullanmış olduğu geçmişte meydana gelen verilerden yola çıkarak gelecekte oluşabilecek olan taleplerin tahmin edilmesi yöntemini kullanmayıp değişkenler arasında bulunan ilişki üzerinden yola çıkarak tahmin yapan bir yöntemdir (Üreten, 2013). Regresyon analizi; bağımsız değişkenler (X1, X2 …Xn) ile bağımlı değişken (Y)'deki değişimi açıklamayı hedefler.

Regresyon modeli iki (ya da daha çok) değişken arasındaki ilişkinin fonksiyonel şeklini göstermekle kalmaz, değişkenlerden birinin değeri bilindiğinde, diğeri hakkında tahmin yapılmasını da sağlar.

İki ya da daha çok değişken arasındaki ilişkinin matematiksel bağıntısı "Regresyon Analizi" ile, ilişkinin yönü ve derecesi ise "Korelasyon Analizi" ile incelenir.

Regresyon analizi bir tahmin (öngörüsel) analizi olup, bağımlı değişkenin bağımsız değişkenler yardımıyla tahmin edilmesini sağlar. Ayrıca bağımlı değişkeni etkileyen en önemli bağımsız değişken/değişkenlerin hangisi olduğunu ortaya çıkarır.

Regresyon analizi;

- Basit doğrusal regresyon analizi,
- Çok değişkenli regresyon analizi,
- Doğrusal olmayan regresyon analizi olarak sınıflandırılabilir.

Aralarında ilişki araştırılması istenen değişkenler sayılabilir veya ölçülebilir nitelikte olabilir. Üzerinde durulan değişkenlerden bağımlı değişken y, bağımsız değişken x ise, $y = f(x)$ şeklindeki fonksiyona regresyon modeli denir

Basit doğrusal regresyon; bağımsız değişken (X) ile bağımlı değişken (Y)'deki değişimi açıklamayı, bağımsız değişkendeki bir birimlik değişimin bağımlı değişken üzerindeki etkisini ölçmeyi amaçlar.

Stokastik (Olasılıklı) bir model olan ve anakütledeki ilişkiyi gösteren basit doğrusal regresyon denklemi aşağıdaki gibi ifade edilir [3]:

$$y = \beta 0 + \beta 1x + \varepsilon$$

Burada;

β0: Doğrunun y-eksenini kestiği yer ve regresyon sabitidir.

β1: Doğrunun eğimi veya regresyon katsayısıdır.

ε: Rastgele (Tesadüfi-Şans) hata değeridir.

Regresyon denklemi kullanılarak, verilen bir x değeri için y'nin tahmini değeri bulunur; ancak x 'in büyüklüğü örnek veri setindeki minimum ve maksimum değerler arasında ise daha iyi tahminler yapılır.

$\hat{y}$: y'nin tahmini değeri olarak tanımladığımızda, örneklem için basit regresyon modeli aşağıdaki gibi gösterilir.

$$\hat{y} = b0 + b1x$$

e terimi örneklemden elde edilen hata ve $\hat{y}$ tahmini bağımlı değişken değeri olmak üzere, hata terimi olan $e = y - \hat{y}$ ile gösterilir. Hata terimi her bir gözlem çiftindeki bağımlı değişkene ilişkin gerçek değer ile modelden tahmin edilen değer arasındaki farktır ve aşağıdaki gibi gösterilir.

$$\varepsilon i = (\beta 0 + \beta 1X) - Yi$$

Regresyon modelinde anakütle hata terimi ε (epsilon) için genel varsayımlar vardır. Basit doğrusal regresyon modeli için bu varsayımların sağlanması gereklidir. Bu varsayımlar aşağıda verilmiştir [3]:

- Regresyon modeli parametrelerine (β 'lara) göre doğrusaldır.

- Bağımsız değişkenin (X'in) değerleri tekrar eden örneklerde sabittir. Bir başka deyişle, X' in stokastik olmadığı varsayılmaktadır.
- Regresyon modelinde hataların beklenen değeri sıfırdır.

Regresyon analizinde parametre tahmini "En Küçük Kareler Metodu" (Least Squares Method) kullanılarak yapılır. Amaç; dağılım grafiğinde (scatter plot) görülen tüm noktalar için, doğruya uzaklıklarının bulunması ve bunların toplamının minimize edilmesidir.

Regresyon analizini yapmaktaki amaç; örneklem bilgilerinden yararlanarak anakütle için tutarlı ve güvenilir tahmin değerleri elde etmek olduğundan, gerçek y değerleri ile tahmini ŷ değerleri arasında farkın olmaması ya da az fark olması beklenir. Bu nedenle β1 ve β2 katsayılarının $(y - \hat{y})$ değerini en küçük yapacak şekilde bulunması gerekir. Bunu sağlayan yöntemlerden birisi de En Küçük Kareler Yöntemi (EKKY)'dir [3]:

Bu sebepten dolayı maksimum fayda ve verim beklenen en iyi regresyon modelini kurabilmek adına farkları minimize edecek a ve b değerlerinin bulunması gerekmektedir.

Basit doğrusal regresyondan farklı olarak çoklu doğrusal regresyon bağımlı değişkenin birden fazla bağımsız değişken ile ilişkili olduğu durumlarda kullanılmaktadır (Yılmaz ve Top, 2009). Çoklu regresyon analizinde birden fazla bağımsız değişkenin bağımlı değişken ile olan ilişkileri incelenmektedir. Regresyon analizine dahil edilen bağımsız değişkenin bağımlı değişkeni ne düzeyde açıklayabildiği çoklu regresyon modellerinde önemli bir noktadır. Regresyon analizinde bağımlı değişken üzerinde etkisi olan bağımsız değişkenlerin katsayıları etki düzeylerini göstermediğinden korelasyon analizi yardımı ile bağımsız değişkenlerin etki düzeyleri bulunur (Karaca, 2015). Çoklu doğrusal regresyon yöntemine ait matematiksel denklem 'Denklem 1'de gösterilmiştir.

Denklem 1 : $Yi = a + b1X1 + b2X2 + \cdots + bnXn$

Burada Yi hesaplanmak istenen bağımlı değişkeni, a regresyon doğrusuna ait başlangıç değerini, b regresyon doğrusuna ait olan eğimi ve Xi bağımsız değişkenleri ifade etmektedir (Aksoy, 2008). 'Denklem1'de b katsayıları bağımsız değişkenlerde oluşacak değişimin bağımlı değişken üzerinde bıraktığı etkiyi göstermektedir. Çoklu doğrusal regresyon fonksiyonuna ulaşma yolunda basit doğrusal regresyon fonksiyonunda olduğu gibi en küçük kareler yönteminden yararlanılır (Yoldaş, 2006).

Doğru denklemleri değişkenler arasında bulunan ilişkiyi ifade etmekte her zaman yeterli olmayabilir. Bu gibi durumlarda eğri denklemleri kullanılır. Bağımlı değişken ile bağımsız değişkenler arasındaki ilişkiyi bulmak için kullanılan eğri denklemlerinde en kolay uygulanabilecek işlem verileri grafik olarak çizmektir. Veriler sonucunda ulaşılan eğriler üssel, hiperbolik ya da parabolik olabilirler. Modellerin belirlenmesinden sonra değişkenler arasında bulunan ilişkiyi en verimli ve en sağlam açıklayan model uygun olarak seçilmelidir (Özsoy, 2006).

Bağımlı değişken Y ve bağımsız değişken X olarak tanımlandığında y = f(x) şeklindeki fonksiyona regresyon modeli denir. f(x) fonksiyonu aşağıda farklı şekiller alabilir (Ersöz, Ersöz, 2019):

Doğrusal: $y = ax + b$

Parabolik: $y = ax\ 2 + b$

Üstsel: $y = ab\ x$, $y = ae\ x$

Geometrik: $y = ax\ b \rightarrow \log y = b \log(ax)$

Hiperbolik: $y = (ax + b)\ -1$

Ağırlıklı hareketli ortalama yöntemi genel olarak basit hareketli ortalama yöntemine benzemektedir. Hesaplanmasında son dönemlere ait verilerin toplanarak aritmetik ortalaması alınan ve eşit değerde ağırlık atanan basit hareketli ortalama yönteminden farklı olarak bu yöntemde hesaplamaya dahil edilen dönemlere farklı farklı ağırlıklar atanabilmektedir (Özer, 2009). Bir diğer ifadeyle, yapılacak tahmine etki düzeyi düşünüldüğünde daha fazla etkiye sahip olacak döneme diğer dönemlere nazaran daha büyük ağırlık, daha az etkiye sahip olacak döneme diğer dönemlere nazaran daha küçük ağırlık atanmaktadır. Bu yöntemde tahmin, belirlenen dönemlerin talepleri atanan ağırlıklar ile çarpılır. Elde edilen çarpım sonuçları toplanır ve daha sonra ağırlıkların toplamına bölünür. Tahmin sonucu elde edilir (Taşdemir, 2012).

Ağırlıklı hareketli ortalama yönteminde dönemlere atanan ağırlık değerleri 0 ile 1 arasında değerler almaktadır ve ağırlık toplamları 1'e eşit olur (İlhan, 2015). Üç aylık dönem baz alınarak yapılan ağırlıklı hareketli ortalama yönteminde son döneme ait ağırlık katsayısı

0,5, son dönemden bir önceki döneme ait ağırlık katsayısı 0,3 ve en uzak döneme ait ağırlık katsayısı 0,2 olarak örneklendirilebilir (Malhotra, 2014). Ağırlıklı hareketli ortalama yönteminde tahmin edilen döneme en yakın dönemin ağırlık katsayısı büyük verilerek hareketli ortalama yönteminde her dönemin eşit ağırlığa sahip olması problemi çözüme ulaştırılır (Demirbaş, 2011). Ağırlıklı hareketli ortalama yöntemine ait matematiksel formül 'Denklem 2' de gösterilmiştir.

Denklem 2 : $St = Wt{-}1At{-}1{+}Wt{-}2At{-}2{+}\cdots{+}Wt{-}nAt{-}n \sum W$

Burada St tahmin değeri olarak ifade edilirken, hareketli ortalamaya dahil edilen dönem sayısı n ile, dönemler içerisinde gerçekleşen değerler A ile ve söz konusu dönemler için belirlenen ağırlık değerleri W ile ifade edilmektedir (Yüksel, 106).

Hareketli ortalama yönteminden farklı olarak ağrılıkların hesaba dahil edilerek istenen veriler adına ortalama içerisindeki pay artırılmaktadır. Tahminlerin son dönemlerde gerçekleşen değişikliklere daha hızlı geri dönüş, cevap vermesi, son dönemlere verilen ağırlıkların büyük olmasına bağlıdır. Bu aşamada en önemli nokta dönemlere verilen ağırlıkların belirlenmesidir. Ağırlıklar deneme yanılma yönteminden faydalanılarak belirlenmektedir (Yılmaz ve Top. 2009).

Bulgular

Bu çalışmada, Düzce ilinde faaliyet gösteren Antik Atölye firması adına gelecekte oluşacak ağaç sarfiyatını hesaplamak için 2015Q1 ve 2018Q12 arasındaki dönemlere ait veriler kullanılarak nüfus artış oranı, kereste fiyatı ve TÜFE değişkenlerine ait aynı dönemdeki verilerin söz konusu atölye adına ağaç sarfiyatına olan etkisi araştırılmıştır. Çalışmada çoklu doğrusal regresyon ve ağırlıklı hareketli ortalama yöntemleri kullanılmış ve SPSS Statistics22 paket programından yararlanarak analizler yapılmıştır. Çalışma sonucunda her iki yöntem adına bulunan sonuçlar kıyaslanmış ve yorumlanmıştır.

Tek bağımsız değişken olduğunda bakılan R değeri ve çalışmada kullanılan gibi birden fazla bağımsız değişken olduğunda değerlendirilen ARS değeri; bağımsız değişkenlerin bağımlı değişkenin varyansının yüzde kaçını açıkladığını söylemektedir. Yani nüfus artış oranı, kereste fiyatı ve Tüfe Antik Atölye adına kullanılan kereste miktarının %67,2 sini açıklamaktadır.

Kurulan modelin anlamlılığı Anova ile araştırılmış ve modelin anlamlı olduğu görülmüştür. Yani bağımsız değişkenlerin en birinin bağımlı değişken üzerinde anlamlı bir etkiye sahip olup olmadığını göstermektedir.

Modelde değişkenler arasında çoklu bağlantı problemi bulunmamaktadır. Bütün haliyle model %5 anlamlılık düzeyinde anlamlıdır. Bağımsız değişkenlerin katsayıları tek tek incelendiğinde %5 anlamlılık düzeyinde sabit terim katsayısı 0,000 ile, nüfus artış oranı katsayısı 0,001 ile, kereste fiyatı katsayısı 0,000 ile, $\alpha = 0{,}05$ değerinden küçük olup anlamlı oldukları anlaşılmaktadır. Fakat bununla beraber Tüfe katsayısı 0,911 ile anlamsız çıkmıştır. Ek olarak burada nüfus artış oranındaki değişim ile bağımlı değişken arasında ters yönlü ilişki vardır. R^2 değeri %67,2 bulunmuş ve bağımsız değişkenlerin bağımlı değişkeni açıklama oranı yeterli görülmüştür.

Bu bağlamda regresyon denklemi;

$Y = 13{,}721 - 3{,}626(Nüfus\ Ar.) + 0{,}014(Keres.\ fiy.) - 0{,}04(Tüfe)$ şeklinde tanımlanabilir. Bu fonksiyonda yer alan sabit değer, nüfus artış oranı ve kereste fiyatı söz konusu atölye için önerilen üretim tahmin fonksiyonunda yer almalıdır. Tüfe değişkeni %5 anlamlılık düzeyinde anlamsız olduğundan önerilen üretim tahmin fonksiyonunda yer almasına gerek yoktur.

Bu çalışmada ayrıca talep tahmini için zaman serisi yöntemlerinden ağırlıklı hareketli ortalama yöntemi kullanılmıştır. Çalışmada kullanılan dönemler adına hesaplama yapılırken üç aylık periyotlar oluşturulmuştur.

$$St = Wt{-}1At{-}1 + Wt{-}2At{-}2 + \cdots + Wt{-}nAt{-}n \sum W$$

Formülü doğrultusunda 48 dönem için farklı tahmin sonuçları elde edilmiştir. Ağırlıklı hareketli ortalama yönteminde her dönem için birbirinden farklı ağırlıklar uygulanmaktadır. Son döneme ait verilerin tahmin yapılırken etkisinin daha büyük olacağı kanası mevcutsa o dönemin ağırlığı daha fazla olur. Bu çalışmada tahmin yapılırken son döneme ait ağırlık katsayısı 0,5, bir önceki döneme ait ağırlık katsayısı 0,3 ve yine ondan bir önceki döneme ait ağırlık katsayısı 0,2 olarak alınmıştır. Çalışmada tahmin edilen değerler ile gerçekleşen değerler kıyaslandığında değerlerin birbirine yakın olduğu görülmektedir. Bu durumda bu tahmin yöntemi gelecek yıllar adına yapılacak üretim planlaması için uygun olacaktır. Gerçekleşen

değerler ile tahmin edilen değerlerin ortalaması alındığında %98'lik bir oran çıktığı ve tahmin yönteminin gayet başarılı olduğu görülmektedir.

Sonuç ve Tartışma

Hızla gelişen teknoloji çağında işletmeler müşterilerine maksimum verimli ve minimum maliyetli hizmetler sunma eğilimindedir. İşletmelerin bu söz konusu kısıtları en iyi şekilde yerine getirebilmesi için ileride nelerle karşılaşacaklarını bilmeleri ya da tahmin etmeleri gerekmektedir. Bu bağlamda işletmeler için talep tahmini çalışmaları oldukça önemli bir noktaya sahiptir.

Bu çalışmada Düzce ilinde faaliyet gösteren Antik Atölye firması adına 2015Q1 ile 2018Q12 dönemleri arasında gerçekleşen veriler ışığında gelecek dönemler adına tahmin yapılabilmesi için bir tahmin denklemi oluşturulmaya çalışılmış ve kullanılan yöntemler adına çıkan sonuçlar kıyaslanarak hangi yöntemin daha etkili olduğu ortaya konmak istenmiştir. Bu amaçla çalışmada çoklu doğrusal regresyon analizi yöntemi ve ağırlıklı hareketli ortalama yöntemleri kullanılmıştır.

Regresyon analizi sonucunda söz konusu firma için üretim tahmini formülü $Y = N\ -0{,}362 + K\ 0{,}566 + T\ -0{,}010$ şeklinde bulunmuştur. Üretim tahmin fonksiyonunda Antik Atölye firması adına nüfus artış oranının ters yönlü anlamlı bir ilişkisi olduğu, kereste fiyatında doğru orantılı anlamlı bir ilişki olduğu görülmüştür. Tüfe değerinin firmanın üretim tahmini adına anlamsız olduğu saptanmıştır.

Ağırlıklı hareketli ortalama yönteminde ise firmadan alınan veriler ile üç aylık periyotlara üzerinden tahminlerde bulunulmuş ve gerçekleşen değerler ile kıyaslanmıştır.

Çalışmada kullanılan iki yöntem karşılaştırıldığında ağaç endüstrisinde faaliyet gösteren firma için ağırlıklı hareketli ortalama yönteminin tahmin değerleri regresyon analizi yönteminin tahmin değerlerinden daha iyi sonuç verdiği gerçekleşen değerler ile tahmin edilen değerlerin farkından ortaya konmuştur.

Literatürde talep tahmini için birçok yöntem bulunmaktadır. Farklı sektörlerde farklı yöntemler daha iyi sonuç vermektedir. Bu çalışmada ağırlıklı hareketli ortalama yöntemi

gerçekleşen değerlere daha yakın sonuçlar vermiştir. Bununla beraber yapay sinir ağları yöntemi ile tahminleme yapılması daha hassas sonuçlar elde etmek için denenebilir.

KAYNAKÇA

[1] H. Cındık, H., Akyüz, C. ve Serin, "Orman Ürünleri Sanayinde Küçük ve Orta Öçekli İşletmeler," in 1. Ulusal Mobilya Kongresi Bildiri Kitabı, 1997, pp. 176–189.

[2] J. S. Montgomery, D.C., Johnson, L.A. and Gardiner, Forecasting and Time Series Analysis. NY: McGraw Hill, 1990.

[3] E. T. Ersöz F., İstatistik – I. Ankara: Seçkin Yayınevi, 2019.

[4] F. Ersöz; IBM SPSS ile İstatistiksel Veri Analizi, Seçkin Kitabevi, 2019

Deep Learning Based Abnormality Detection Application in Enterprise Network Traffic

Emrullah ERGİNAY, M. Ali AKÇAYOL

Department of Computer Engineering, Gazi University, TURKEY

Abstract: In this paper, a deep learning model has been developed to detect whether malware/spyware leaks data to command and control servers and a new dataset has been obtained from real-time environment for test of the model. In addition, effect of the size of the data set and hyperparameters such as the number of layers of the deep neural network on the success rate have been investigated. In this study, real-time data for harmful and normal İnternet traffic have been obtained in the application layer and 100 features have been selected. The developed deep learning model has been applied to 16,000 sample obtained from real-time Internet traffic. From the experimental results, accuracy rates of 90% to 94% were obtained with various number of samples and various number of layers in the deep learning model. It has been seen from the experimental results that increase the number of samples increases the accuracy rate. As well as, it has been seen that as increase the number of layers in the deep neural network the accuracy rate increased first, further increase the hidden layers did not affect the success rate. In this study, more distinctive and important features have been investigated than others in the literature and the results have been tested.

Keywords: Anomaly Detection, Deep Learning, Corporate Networks, Malware

REFERENCES

[1] D. Zhao, I. Traore, B. Sayed, W. Lu, S. Saad, A. Ghorbani and D. Garant, "Botnet detection based on traffic behavior analysis and flow intervals", Computers & Security 39, 2-16, April 2013.

[2] M. Ichino, K. Kawamoto, T. Iwano, M. Hatada and H. Yoshiura, "Evaluating Header Information Features for Malware Infection Detection", Journal of Information Processing, Vol.23, No.5, 603-612, September 2015.

[3] S. Kumar, A. Viinikainen and T. Hamalainen, "Machine Learning Classification Model For Network Based Intrusion Detection System", The 11th International Conference for Internet Technology and Secured Transactions (ICITST-2016), 242 – 249, December 2016.

[4] K. Singh, S. Guntuku, A. Thakur and C. Hota, "Big Data Analytics framework for Peer-to-Peer Botnet detection using Random Forests", Information Sciences 278, 488-497, March 2014.

[5] A. Boukhtouta, S. Mokhov, N. Lakhdari, M. Debbabi and J. Paquet, "Network malware classification comparsion using DPI and flow packet headers", Journal of Computer Virology and Hacking Techniques, 69-100, July 2015.

[6] M. Mimura, Y. Otsubo, T. Hidehiko and T. Hidema, "A Practical Experiment of the HTTP-Based RAT Detection Method in Proxy Server Logs", 12th Asia Joint Conference on Information Security, 31-37, August 2017.

[7] L. Boero, M. Marchese and S. Zappatore, "Support Vector Machine meets Software Defined Networking in IDS domain", 29th International Teletraffic Congress (ITC 29), 25-30, September 2017.

[8] A. Verma and V. Ranga, "Statistical analysis of CIDDS-001 dataset for Network Intrusion Detection Systems using Distance-based Machine Learning", Procedia Computer Science 125, 709-716, 2018.

[9] Z. Chen, Q. Yan, H. Han, S. Wang, L. Peng, L. Wang and B. Yang, "Machine learning based mobile malware detection using highly imbalanced network traffic", Information Sciences, 346-364, April 2018.

[10] A. Lashkari, A. Kadir, H. Gonzalez, K. Mbah and A. Ghorbani, "Towards a Network-Based Framework for Android Malware Detection and Characterization", 15th International Conference on Privacy, Security and Trust, August 2017.

[11] V. Kant, M. Singh and N. Ojha, "An Efficient Flow based Botnet Classification using Convolution Neural Network", International Conference on Intelligent Computing and Control Systems (ICICCS 2017), 941-946, June 2017.

[12] T. Shibahara, T. Yagi, M. Akiyama, D. Chiba and T., "Efficient Dynamic Malware Analysis Based on Network Behavior Using Deep Learning", Global Communications Conference (GLOBECOM), December 2016.

[13] D. Y. Kim and H. Y. Song, "Method of Predicting Human Mobility Patterns Using Deep learning", Neurocomputing 280, 56-64, March 2018.

[14] S. Lin, K. Ying, C. Lee and Z. Lee, "An İntelligent Algorithm With Feature Selection And Decision Rules Applied To Anomaly İntrusion Detection", Applied Soft Computing 12, 3285-3290, October 2012.

[15] K. P. Ferentinos, "Deep learning models for plant disease detection and diagnosis", Computers and Electronics in Agriculture 145, 311-318, February 2018.

[16] Ç. Kaya and O. Yıldız, "Makine Öğrenmesi Teknikleriyle Saldırı Tespiti: Karşılaştırmalı Analiz", Marmara Fen Bilimleri Dergisi 2014, 3:89-104, March 2015.

[17] S. B. Kotsiantis, "Supervised Machine Learning: A Review of Classification Techniques", Informatica 31 (2007), 249-268, July 2007.

[18] A. Kamilaris and F. Prenafeta-Boldu, "Deep Learning in Agriculture: A Survey", Computers and Electronics in Agriculture 147, 70-90, April 2018.

[19] A. A. Diro and N. Chilamkurti, "Distributed Attack Detection Scheme Using Deep Learning Approach For Internet of Things", Future Generation Computer Systems 82, 761-768, May 2018.

[20] J. Gu, Z. Wang, J. Kuen, L. Ma, A. Shahroudy, B. Shuai, T. Liu, X. Wang, G. Wang, J. Cai and T. Chen, "Recent advances in convolutional neural networks", Pattern Recognition 77, 354-377, October 2018.

[21] W. Wang, M. Zhu, X. Zeng, X. Ye and Y. Sheng, "Malware Traffic Classification Using Convolutional Neural Network for Representation Learning", Information Networking (ICOIN), January 2017.

[22] T. Tang, L. Mhamdi, D. McLernon, S. A. R. Zaidi and M. Ghogho, "Deep Learning Approach for Network Intrusion Detection in Software Defined Networking", Wireless Networks and Mobile Communications (WINCOM), October 2016.

[23] H. Chougrad, H. Zouaki and O. Alheyane, "Deep Convolutional Neural Networks for breast cancer screening", Computer Methods and Programs in Biomedicine 157, 19-30, April 2018.

Lean Production in Iron and Steel Production Line

Taner ERSÖZ[1], Filiz ERSÖZ[1], Karanfil SARIZ[2]

[2]Department of Actuary and Risk Management,
[2]Department of Industrial Engineering, Karabük University, TURKEY

Abstract: Rapidly developing technology in the world endangers many companies we are in today. The companies are now obliged to improve themselves in order to compete in both domestic and international markets. In this sense, concepts such as quality, cost and customer satisfaction are becoming more important. Lean production is a production method that is not outdated at this point. In the 1950s, the lean production which was laid the foundation with the Toyota Production Company, is still up to date. Lean production, which adopts zero error and zero inventory principle, is applied in both the manufacturing sector and the service sector. Today, it is still very important in terms of the continuity of lean production which is developed and integrated with new technologies. Lean production is a gateway to the world from the Japanese people. In this thesis," Lean Manufacturing Application in Iron-Steel Production Line Hatt is tried to be explained. Lean production and techniques are examined in detail. Then, in light of this information, the results obtained in iron and steel industry were interpreted.

Keywords: Lean, Just in time, Profit, Cost, Waste, Simulation

REFERENCES

[1] Arslan, S., "Yalın Üretim ve Man Türkiye A.Ş.' de Örnek Bir Yalın Üretim Enstitüsü, Ankara, 3-43 (2008).

[2] Atmaca, M., Terzi, S., "Yalın Üretim Sistemi Açısından Değer Akışı Maliyetlemesinin İncelenmesi", Fakülte Dergisi, 16, 449-466 (2011).

[3] Michalska, J., Szewieczek, D., "The 5S Methodology As a Tool For İmproving The Organisation", Journal of Achievements in Materials and Manufacturing Engineering, 2: 211-214(2007).

[4] Ohno, T., "Toyota Ruhu", Hakan Feyyat/ Canan Feyyat, İstiklal Caddesin Han Geçidi Sokak, Galatasaray/İstanbul,39-197 (2017).

[5] Sarı, Ö., Seri, V. ve Yazgan, H., "Toyota Üretim Sisteminin Özellikleri", SAÜ Fen Bilimleri Enstitüsü Dergisi, 2, 129-134 (1998).

[6] Şeker, A., "Yalın Üretim Sisteminde Kanban, Tek Parça Akışı ve U Tipi Yerleştirme Sistemleri", The Journal of Academic Social Science Studies, 449-470 (2016)

[7] Türkan, Ö., "Üretimde Yalın Dönüşümün Temel Performans Kriterleri", BAÜ Fen Bilimleri Enstitüsü Dergisi, 12, 28-41 (2010).

[8] Yılmaz, E., "Siparişe Göre Üretim Yapan Sistemlerde Yalın Üretim Uygulamaları", Yüksek Lisans Tezi, İstanbul Teknik Üniversitesi Fen Bilimleri Enstitüsü, İstanbul, 11-57 (2012).

[9] İpek, F., "Siparişe Göre Üretimde Malzeme Hazırlama ve Üretim Hattının Beslenmesinin İyileştirilmesi: Yalın Üretim Uygulaması", Yüksek Lisans Tezi, Gazi Üniversitesi Fen Bilimleri Enstitüsü, Ankara, 3-28 (2015).

[9] İraz, R., Zerenler, M., "Japon yönetim Anlayışı ve Şirket Ağları (KEIRETSU) Analizi, Balıkesir Üniversitesi Sosyal Bilimler Dergisi, 758- 773.

[10] Jones, D., Womack, J., "Yalın Düşünceler", Oygur Yamak, Osmangazi Mahallesi/İstanbul, 21-387 (2016).

[11] Jones, D., Womack, J., "Yalın Çözümler", Saadet Özkal, Akçaburgaz Mahallesi Esenyurt/İstanbul, 15-143(2015).

[12] Karaboğa, K., Zerenler, M., "Müşteri Memnuniyetinin Sağlanmasında Hataların Önlenmesine Yönelik Üretim Odaklı Bir Bakış Açısı: PokaYoke Sistemleri", Sosyal Bilimler Enstitüsü Dergisi, 263-275 (2014).

[13] Kaymakçı, Ö., "Bir PTT Şubesinde Yalın Üretim-5S Uygulaması", Yüksek Lisans Tezi, Sakarya Üniversitesi Fen Bilimleri Enstitüsü, Sakarya, 4-57 (2012).

[14] Kılıç, A., "Otomotiv Yan Sanayinde Yalın Üretim Uygulaması", Yüksek Lisans Tezi, İstanbul Üniversitesi Fen Bilimleri Enstitüsü, İstanbul, 3-51 (2016) Rother M., Shook J., "Görmeyi Öğrenmek", Ayşe Soydan, Merkez Mah. Sefir Sokak Sarıyer/İstanbul,1-102 (1999).

A Road Map for The Big Data Implementation in the Judicial System

Zeynep ÇAĞLAR

Sakarya University, TURKEY

Abstract: This research involves Big Data and its definition. Also, it covers general visualization of the problems in Turkish legal system and deal with a real case to create a methodology for solving the problems in filing an indictment process with the help of Big Data. The goal is to show that how can the finishing time of the process of filing indictments significantly reducing when Big Data technology integrated to the process. This has been done by creating a decision support system that can apply to all the cases in the phase of linking evidences and filling the indictment. This research will provide valuable information regarding the methodology of applying Big Data and address the role of Big Data in solving problems of ineffective and complex organizations.

Keywords: Big Data Road Map, Indictment, Methodology.

REFERENCES

[1] Laney, D. (2001) 3D Data Management: Controlling Data Volume, Velocity and Variety. Working paper, Application Delivery Strategies, Meta Group AG.

[2] Lämmel, R. (2008) Google's Map Reduce Programming Model, Revisited. Science of Computer Programming. 70: 1–30

[3] Grover, P. and Kar, A. (2017). Global Journal of Flexible Systems Management. Mumbai: Springer India, pp.203-229.

Supplier Selection Using an Intuitionistic Fuzzy Evaluation System

Ayşenur AKIN1, M. Bahar BAŞKIR[1], Hamza GAMGAM[2]

[1]Department of Statistics, Bartın University,
[2]Department of Statistics, Gazi University, TURKEY

Abstract: Selection and evaluation problems have uncertainties due to concept and its perception discrepancies. Fuzzy set theory is one of the widely used methodology to cope with these uncertainties. There is a growing interest in evaluations using intuitionistic fuzzy sets.

Differently from fuzzy sets, intuitionistic sets include both belonging degrees and nonbelonging degrees. Thus, the evaluation using intuitionistic fuzzy sets gives more realistic results. In this study, an intuitionistic fuzzy set-based evaluation system is proposed for supplier selection problem of a construction company. This system has qualitative- and quantitative evaluation parts. As qualitative part, decision makers of the company evaluate suppliers by the supplier selection criteria: i) quality, ii) price, iii) delivery, iv) productivity, v) service, vi) flexibility. The quantitative part includes supplier scores calculated through the current evaluation system of the company. A supplier-evaluation database was created by the abovementioned parts. The database was structured by α-cut representation of the evaluations. The calculations using intuitionistic fuzzy sets were done for this database, which was occurred by lower and upper bounds. After defuzzifying the database, suppliers were classified using the well-known fuzzy clustering algorithm, fuzzy c-means. The classification using fuzzy clustering algorithm has 95.0% accuracy.

Keywords: Intuitionistic Fuzzy Sets, Supplier Selection, Fuzzy Clustering, α-Cuts, Accuracy.

Environmental Impact Assessment of Two Alternative Wastewater Neutralization Chemicals in Textile Industry Wastewater Treatment Plant

Fatma Şener FİDAN[1], Emel Kızılkaya AYDOĞAN[1], Niğmet UZAL[2]

[1]Dept. of Industrial Engineering, Erciyes University,
[2]Dept. of Civil Engineering, Abdullah Gül University, TURKEY

Abstract: The wastewater treatment in the textile industry is of special importance due to the intensive use of chemicals and dyes. However, wastewater treatment plants have impacts on environmental and these environmental impacts should also be evaluated and minimized. Neutralization is one of the processes with the huge chemical consumption in the wastewater treatment plant. Therefore, the chemical alternatives used in the neutralization process should be compared in terms of their environmental impacts. In this study, the performances of carbon dioxide and sulfuric acid as two alternative chemicals used in the neutralization process applied in a textile factory wastewater treatment plant are compared using the life cycle approach. The neutralization process using carbon dioxide yielded better results in the categories of abiotic depletion, fossil fuels, ozone layer depletion (ODP), fresh aquatic ecotoxicity, marine aquatic ecotoxicity, terrestrial ecotoxicity, photochemical oxidation, acidification, and eutrophication.

Keywords: Wastewater treatment, Sustainability, Life Cycle Assessment, Textile Industry, Neutralization.

REFERENCE

[1] WCED, 1987. World Commission on Environment and Development, Our Common Future. Oxford University Press, New York.

[2] European Commission. Integrated pollution prevention and control (IPPC). Best Available Techniques (BAT) Reference Document for Waste Treatment, 2018, http://eippcb.jrc.ec.europa.eu/reference/

[3] Baumann, H., Tillman, A.M., 2004. The Hitch Hiker's Guide To LCA: An Orientation in Life. Cycle Assessment Methodology and Application. Studentlitteratur, Lund.

[4] ISO, 2006a. ISO 14040 Environmental Management -- Life Cycle Assessment -- Principles and Framework, International Organization for Standardization, Geneva.

[5] ISO, 2006b. ISO 14044: Environmental Management -- Life Cycle Assessment -- Requirements and Guidelines, International Organization for Standardization Internatio, Geneva

[6] Raghuvanshi, S., Bhakar, V., Sowmya, C., Sangwan, K.S., 2017. Wastewatertreatment Plant Life Cycle Assessment: Treatment Process to Reuse of Water. The 24th Cirp Conference on Life Cycle Engineering, Procedia Cirp 61, 761 – 766.

[7] Yildirim, M., Topkaya, B., 2012. Assessing Environmental Impacts of Wastewater Treatment Alternatives for Small-Scale Communities. Clean E Soil, Air, Water 40 (2), 171e178.

[8] Venkatesh, G., Brattebø, H., 2011. Environmental Impact Analysis of Chemicals and Energy Consumption in Wastewater Treatment Plants: Case Study of Oslo, Norway. Water Science and Technology 63 (5), 1018e1031.

[9] Tillman, A., Svingby, M. And Lundstrom, H., Life Cycle Assessment Of Municipal Wastewatersystems, The International Journal Of Life Cycle Assessment, Vol. 3, No. 3, Pp. 145-157, 1998, Http://Dx.Doi.Org/10.1007/Bf02978823

[10] Vidal, N., Poch, M., Martí, E., Rodríguez-Roda, I., 2002. Evaluation of The Environmental Implications to Include Structural Changes in A Wastewater Treatment Plant. Journal of Chemical Technology & Biotechnology, Vol. 77, No. 11, Pp 1206-1211, Http://Dx.Doi.Org/10.1002/Jctb.674

[11] Remy, C., Jekel, M., 2008. Sustainable Wastewater Management: Life Cycle Assessment of Conventional and Source-Separating Urban Sanitation Systems. Water Science and Technology 58 (8), 1555e1562.

[12] Pasqualino, J. C., Meneses, M., Castells, F., 2011. Life Cycle Assessment of Urban Wastewater Reclamation and Reuse Alternatives. J. Ind. Ecol., Vol. 15, No. 1, Pp 49-63, Http://Dx.Doi.Org/10.1111/J.1530-9290.2010.00293.X

[13] Y. Li, X. Luo, X. Huang, D. Wang, W. Zhang, 2013. Life Cycle Assessment of a Municipal Wastewater Treatment Plant: A Case Study in Suzhou, China. J. Clean. Prod., Vol. 57, 221–227.

[14] Opher, T., Friedler, E., 2016. Comparative LCA of Decentralized Wastewater Treatment Alternatives for Non-Potable Urban Reuse. J. Environ. Manag. 182, 464–476.

[15] Alfonsín, C., Hospido, A., Omil, F., Moreira, M.T., Feijoo, G., 2014. Ppcps In Wastewater Update and Calculation of Characterization Factors for Their Inclusion in LCA Studies.J. Clean. Prod. 83, 245–255.

[16] Guest, J.S., Skerlos, S.J., Barnard, J.L., Beck, M.B., Daigger, G.T., Hilger, H., Jackson, S.J., Karvazy, K., Kelly, L., Macpherson, L., Mihelcic, J.R., Pramanik, A., Raskin, L., Van Loosdrecht, M.C.M., Yeh, D., Love, N.G., 2009. A New Planning and Design Paradigm to Achieve Sustainable Resource Recovery from Wastewater. Environmental Science and Technology 43 (16), 6121e6125.

[17] Ecoinvent Centre, 2013. Database Ecoinvent Data V3.0, Swiss Centre for Life Cycle Inventories. URL Www.Eco-Invent.Org

[18] Pre Sustainability, 2017. Pré Consultant, Simapro 8.5.2.0, Https://Simapro.Com/, Geneva.

[19] Guinée J. B., Gorrée M., Heijungs R., Huppes G., Kleijn R., De Koning A., Van Oers L., Wegener Sleeswijk A., Suh S., Udo De Haes H. A., De Bruijn H., Van Duin R., Huijbregts M. A. J., Lindeijer E., Roorda A. A. H. And Weidema B. P., 2001a. Life Cycle Assessment; An Operational Guide to The ISO Standards; Parts 1 And 2. Ministry of Housing, Spatial Planning and Environment (VROM) And Centre of Environmental Science (CML). Den Haag And Leiden, The Netherlands, Retrieved From: Http://Www.Leidenuniv.Nl/Cml/Ssp/Projects/Lca2/Lca2.Html

[20] Guinée J. B., Gorrée M., Heijungs R., Huppes G., Kleijn R., De Koning A., Van Oers L., Wegener Sleeswijk A., Suh S., Udo De Haes H. A., De Bruijn H., Van Duin R., Huijbregts M. A. J., Lindeijer E., Roorda A. A. H. And Weidema B. P., 2001b. Life Cycle Assessment; An Operational Guide to The ISO Standards; Part 3: Scientific Background. Ministry of Housing, Spatial Planning and Environment (VROM) And Centre of Environmental Science (CML). Den Haag And Leiden, The Netherlands, Retrieved From: Http://Www.Leidenuniv.Nl/Cml/Ssp/Projects/Lca2/Lca2.Html.

Empati Çıkar Yöntemiyle Akademik Çalışma Gruplarının İncelenmesi

Ömer Faruk ACAR[1], Burhan SELÇUK[2]

[1]Yüksek İhtisas University, [2]Karabük University, TURKEY

Abstract: Bu çalışmada akademisyenlerin yayınları kullanılarak yayından yazara birliktelik analizi yapılmış ve akademik çalışma grupları oluşturulmaya çalışılmıştır. Yayınların başlık, anahtar kelime, özetlerinde geçen kelimeler morfolojik olarak köklerine ayrılmış, metinsel temizleme yapılmış ve Empati-Çıkar yöntemi kullanılarak benzerlik/yakınlık katsayıları hesaplanmıştır. Oluşan benzerlik/yakınlık matrisi kullanılarak yayında katkısı bulunan yazarlar yakınlıklarına göre çalışma gruplarına dahil edilmeye çalışılmıştır.

Keywords: Semantik Ağ, Graf Yapısı, Veri Madenciliği, Türkçe Morfoloji, Birliktelik Analizi, Karmaşık Ağ, Empati - Çıkar.

Application and Comparison of Biclustering Methods in Detecting Crime Regions

Nazan SARI, Sümeyye Gizem ÇAKAR, Olcay EYDEMİR, İbrahim ÇİL

Department of Industrial Engineering, Sakarya University, TURKEY

Abstract: Crime analysis has importance for the detection of crime regions, the prediction of crimes before processing and the security forces to take necessary measures. By using biclustering methods to detect crime regions, simultaneous clustering of the types of crimes and regions where crime is committed to producing more comprehensive results than traditional clustering methods. In this study, CC and Xmotif algorithms of biclustering methods were applied to the real data set in order to detect the crime regions. "Crimes in Boston" data set was used in real data set application. In order to measure the efficiency of the biclusters, the performance of the algorithms was compared with Chia and Karuturi bicluster score (CCPS). The results were obtained by using Matlab functions and it was observed that results of the CC algorithm were better compared to Xmotif algorithm.

Keywords: Biclustering, CC, Xmotif, Crime Data.

REFERENCES

[1] G. Eason, B. Noble, and I. N. Sneddon, "On Certain Integrals of Lipschitz-Hankel Type Involving Products of Bessel Functions," Phil. Trans. Roy. Soc. London, Vol. A247, Pp. 529–551, April 1955. (References)

[2] J. Clerk Maxwell, A Treatise on Electricity and Magnetism, 3rd Ed., Vol. 2. Oxford: Clarendon, 1892, Pp.68–73.

[3] I. S. Jacobs And C. P. Bean, "Fine Particles, Thin Films and Exchange Anisotropy," In Magnetism, Vol. Iii, G. T. Rado And H. Suhl, Eds. New York: Academic, 1963, Pp. 271–350.

[4] K. Elissa, "Title of Paper If Known," Unpublished.

[5] R. Nicole, “Title of Paper with Only First Word Capitalized,” J. Name Stand. Abbrev., In Press.

[6] Y. Yorozu, M. Hirano, K. Oka, And Y. Tagawa, “Electron Spectroscopy Studies on Magneto-Optical Media and Plastic Substrate Interface,” Ieee Transl. J. Magn. Japan, Vol. 2, Pp. 740–741, August 1987 [Digests 9th Annual Conf. Magnetics Japan, P. 301, 1982].

[7] M. Young, The Technical Writer’s Handbook. Mill Valley, Ca: University Science, 1989.

Kronik Böbrek Hastalığının Makine Öğrenmesi Teknikleri ile Sınıflandırılması

Mustafa İlker ERDURSUN[1], Hasan ERBAY[2], Ömer Faruk AKMEŞE[1], İbrahim DOĞAN[1]

[1]Bilgisayar Teknolojileri, Hitit Üniversitesi, Çorum

[2]Bilgisayar Mühendisliği, Kırıkkale Üniversitesi, Kırıkkale

Abstract: Teknolojinin ilerlemesi ile birlikte birçok veri dijital ortamlarda kayıt altına alınarak büyük veri yığınları ortaya çıkmıştır. Veri madenciliği sayesinde bu büyük veri yığınlarının içinden anlamlı ve yararlı bilgilerin ortaya çıkarılması için çalışmalar yapılmaktadır. Özellikle büyük veri yığınlarını analiz etmede klasik analiz yöntemlerinin yetersiz kalması veri madenciliği yöntemlerinin önemini arttırmıştır. Her dönemde olduğu gibi günümüzün en önemli araştırma alanı olan tıp alanında da sürekli olarak hastalara ait veriler artarak kayıt altına alınmaktadır. Kayıt altına alınan veriler bazen tek başına anlamsız gibi görünürken diğer verilerle birlikte bütünsel olarak analiz edildiğinde gizli kalmış önemli bilgiler elde edilebilmektedir. Bu değerli bilgiler, sağlık sektörünün gelişmesine ve doktorların daha doğru bir şekilde teşhis verebilmesine yardımcı olmaktadır. Bu çalışmada, Kronik Böbrek Hastalığı (KBH) veri seti üzerinde analiz yapılmıştır. Farklı modeller oluşturularak, bu modellerin veri üzerindeki tahmin sonuçları karşılaştırılmış ve bu sonuçlara bağlı olarak veriler üzerinde hangi modelin daha iyi sonuç verdiği belirtilmiştir.

Keywords: Makine Öğrenmesi, Tıbbi Veri Madenciliği, Hastalık Tahmini, Sınıflandırma.

REFERENCES

[1] Meganathan, D., & Marudachalam, N. International Journal of Engineering Sciences &Management Research.

[2] Raghavendra, S., & Indiramma, M. (2015). Classification and Prediction Model Using Hybrid Technique For Medical Datasets. Analysis, 127(5).

[3] Singla, M., & Singh, K. Heart Disease Prediction System Using Data Mining Clustering Techniques.

[4] Joshi, S., Shenoy, D., Rrashmi, P. L., Venugopal, K. R., & Patnaik, L. M. (2010, February). Classification of Alzheimer's Disease And Parkinson's Disease By Using Machine Learning and Neural Network Methods. Machine Learning And Computing (ICMLC), 2010 Second International Conference (Pp. 218-222). IEEE.

[5] Peter, T. J., & Somasundaram, K. (2012, March). An Empirical Study on Prediction of Heart Disease Using Classification Data Mining Techniques. In Advances İn Engineering, Science And Management (ICAESM), 2012 International Conference (Pp. 514-518). IEEE.

[6] Robu, R., & Hora, C. (2012, June). Medical Data Mining With Extended WEKA. In Intelligent Engineering Systems (INES), 2012 IEEE 16th International Conference on (Pp. 347-350). IEEE.

[7] Dewan, A., & Sharma, M. (2015, March). Prediction Of Heart Disease Using A Hybrid Technique İn Data Mining Classification. In Computing For Sustainable Global Development (Indıacom), 2015 2nd International Conference on (Pp. 704-706). IEEE.

[8] Soni, J., Ansari, U., Sharma, D., & Soni, S. (2011). Predictive Data Mining For Medical Diagnosis: An Overview Of Heart Disease Prediction. International Journal Of Computer Applications, 17(8), 43- 48.

[9] Domingos, P., & Pazzani, M. (1997). On The Optimality of The Simple Bayesian Classifier Under Zero-One Loss. Machine Learning, 29(2-3), 103- 130.

[10] Shrivas, A. K., & Yadu, R. K. (2017). An Effective Prediction Factors For Coronary Heart Disease Using Data Mining Based Classification Technique. International Journal on Recent And Innovation Trends İn Computing And Communication, 5(5), 813-816.

[11] Le Cessie, S., & Van Houwelingen, J. C. (1992). Ridge Estimators İn Logistic Regression. Applied Statistics, 191-201.

[12] Tong, S., & Chang, E. (2001, October). Support Vector Machine Active Learning For İmage Retrieval. In Proceedings Of The Ninth ACM İnternational Conference on Multimedia (Pp. 107-118). ACM.

[13] Hearst, M. A., Dumais, S. T., Osuna, E., Platt, J., & Scholkopf, B. (1998). Support Vector Machines. IEEE Intelligent Systems And Their Applications, 13(4), 18-28.

[14] Lin, C. F., & Wang, S. D. (2002). Fuzzy Support Vector Machines. IEEE Transactions on Neural Networks, 13(2), 464-471.

[15] Alam, M. S., & Vuong, S. T. (2013, August). Random Forest Classification For Detecting Android Malware. In Green Computing And Communications (Greencom), 2013 IEEE And Internet Of Things (İthings/Cpscom), IEEE International Conference on And IEEE Cyber, Physical And Social Computing (Pp. 663-669). IEEE.

[16] Pal, M. (2005). Random Forest Classifier For Remote Sensing Classification. International Journal Of Remote Sensing, 26(1), 217-222.

Survey on Dynamic Bayesian Network Software Tools

Hüseyin ÇAMBAŞI[1], Özgür KURU[1], Mehmet Fatih AMASYALI[1], Sofiene TAHAR[2]

[1]Department of Computer Engineering, [2]Department of Electrical and Computer Engineering

[1]Yıldız Technical University, TURKEY, [2]Concordia University, CANADA

Abstract: Bayesian networks are probabilistic graphical representations which are used to build models from data and/or expert opinion. They can be utilized for a wide range of tasks including prediction, anomaly detection, diagnostics, automated insight, reasoning, decision making, etc. Dynamic Bayesian Networks (DBN) are extensions of Bayesian networks with temporal support, which can be used to model systems that dynamically change by the time. Nowadays, DBNs are utilized in a wide range of applications including robotics, data mining, speech recognition, digital forensics, protein sequencing, and bioinformatics. Several software tools exist in the public as well as commercial domains that support modelling and simulation of DBNs. However, these DBN software tools differ in terms of features support, ease of use, documentation, user's community, etc. Therefore, it has become important to establish various metrics for selecting the proper software tools for creating and simulating DBNs, such as cost, licensing, GUI, built-in support for inference algorithms, structural learning, data types, etc. The goal of this survey is the evaluation and comparison of existing software tools for building DBNs based on a set of users centered criteria.

Keywords: Dynamic Bayesian Networks, Modeling, Simulation, Software, Tools.

REFERENCES

[1] N. Friedman, D. Geiger, And M. Goldszmidt, Bayesian Network Classifiers, Machine Learning, 29(2-3):131-163, 1997.

[2] V. Mihajlovicand M. Petkovic, Dynamic Bayesian Networks: A State of The Art, Technical Report, Electrical Engineering, Mathematics and Computer Science, University of Twente, the Netherlands, 2001.

[3] K. Murphy, Software For Graphical Models: A Review. Isba Bulletin (December 2007).

[4] K. Murphy, Software Packages For Graphical Models, Software Packages for Graphical Models. [Online]. Available: Http://People.Cs.Ubc.Ca/ Murphyk/Software/Bnsoft.Html. [Accessed: 18-Apr-2019].

[5] M.A.Mahjoub And K. Kalti, Software Comparison Dealing with Bayesian Networks, Advances in Neural Networks Lncs 6677 Pp.168-177, 2011.

[6] Bayesialab, Bayesia, Https://Www.Bayesialab.Com/, 2019, Accessed April 28, 2019.

[7] Bayesfusion, Bayesfusion Ltd, Https://Www.Bayesfusion.Com/, 2019, Accessed April 28, 2019.

[8] Bayes Server, Bayes Server Ltd, Https://Www.Bayesserver.Com/, 2019, Accessed April 28, 2019.

[9] Netica, Norsys Software Corp, Https://Www.Norsys.Com/Netica.Html, 2019, Accessed April 28, 2019.

[10] Hugin Lite, Hugin Expert A/S, Https://Www.Hugin.Com/, 2019, Accessed April 28, 2019

[11] Bnt, Https://Github.Com/Bayesnet/Bnt, 2018, Accessed April 28, 2019.

[12] Banjo, Https://Users.Cs.Duke.Edu/ Amink/Software/Banjo/, 2010, Accessed April 28, 2019.

[13] Amidst, Http://Www.Amidsttoolbox.Com/, 2018, Accessed April 28, 2019.

[14] Bnlearn, Http://Www.Bnlearn.Com/, 2019, Accessed April 28, 2019.

[15] Infer.Net, Https://Dotnet.Github.İo/İnfer/, 2019, Accessed April 28, 2019.

[16] Mocapy++, Https://Sourceforge.Net/Projects/Mocapy/, 2015, Accessed April 28, 2019.

[17] Libpgm, Https://Pythonhosted.Org/Libpgm/, 2012, Accessed April 28, 2019.

[18] Pgmpy, Http://Pgmpy.Org/, 2016, Accessed April 28, 2019.

Determination of Factors Affecting Employee Productivity

Taner ERSÖZ[1], Filiz ERSÖZ[2], Nurdan KALELİ[2], Burçin ATICI[2]

[1]Department of Actuary and Risk Management, [2]Department of Industrial Engineering

Karabük University, TURKEY

Abstract: Productivity is the relationship between the output produced by a production or service system and the input used to create this output. High efficiency is to produce more with the same amount of resources or more output with the same input. However, it is accepted that efficiency and working life quality are closely linked. The efficiency of the enterprises is important for the society and the country in which they carry out their activities as well as for themselves. While the productivity of the enterprises is reflected in the decrease in the costs, increase in the profit, the efficiency of the enterprises has reflections on the new investments, increased added value, the welfare of the society and employment in terms of society and country. It is also important to measure the efficiency that is so important in order to make it sustainable in enterprises. The productivity measured continuously in enterprises will enable managers to make correct and rational decisions, to see problems on time and to produce solutions. The purpose of this study is to determine the factors affecting the productivity of the employees in the steel industry. The study was examined under five titles. These are: "Demoraphic Information", "Economic Factors", "Physical and Ergonomic Factors", "Psycho-Social Factors" and "Risk Factors". According to the findings of the study, the effect of factors on the employees will be determined and improved findings will be presented to the management of the enterprise.

REFERENCES

[1] Birdal, İlker & Nilgün Aydemir; Yönetim Teorileri, Sistem Yayıncılık, İstanbul, 1992.

[2] Bingöl, Dursun. Personel Yönetimi ve Beşeri İlişkiler. Erzurum: Atatürk Üniversitesi Basımevi, 1990.

[3] Byars, Lloyd L.; Concepts of Strategic Management, 3rd Edition, Harper Collins Publishers, New York, 1992.

[4] Can, Halil. Organizasyon ve Yönetim, Ankara: Siyasal Kitabevi, 1999.

[5] Düren, Zeynep. 2000'li Yıllarda Yönetim, İstanbul: Alfa Basım Yayın Dağıtım A.Ş, 2000

[6] Eren, Erol Yönetim ve Organizasyon, İstanbul: Beta Basım Yayım Dağıtım A.Ş, 1998. İşgören Verimliliğini Etkileyen Faktörlerin İncelenmesine 251 C.13, S.3

[7] Hersey Paul & Kenneth H. BLANCHARD, Management and Organizational Behaivor, Englewood Cliffs: Prentice Hall, 1979.

[8] Luthans, Fred. Organizational Behaivor. U.S.A.: Mc Grow Hill, 1995.

[9] Özdamar, Serpil. "İnsan Gücü Potansiyelimizin En verimli Biçimde Değerlendirilmesi Öncelikli Hedeflerimiz Arasındadır", İşveren Dergisi. 8, 1998.

[10] Özçer, Sema. Verimliliğe Etkileri Açısından Sanayi İşletmelerinde Örgüt Yapıları ve Liderlik Biçimleri, Ankara: MPM Yayınları, 1995.

[11] Preffer, Jeffery Rekabette üstünlüğün Sırrı: İnsan.Çeviren: Sinem Gül, İstanbul: Sabah Kitapçılık San. Ve Tic. A.Ş, 1995.

[12] Sabuncuoğlu, Zeyyat & Melek Tüz, Örgütsel Psikoloji, Bursa: Alfa Basım Yayım, 1998

[14] Sapancalı, Faruk. Çalışanların Güdülenmesinde Kullanılan Özendirici Araçlar, Verimlilik Dergisi, 1993.

[15] Sherman, A. W. Herbert J. Churden & J.R., Personnel Management, Ohio: South Western Publishing Co.

[16] Storey John & Sisson Keith, Managing Human Resources and Industrial Relations, Buckingam: Open University Pres, 1995.

[17] Uğur, Adem "Türkiye'de İşgücü Verimliliğini Etkileyen Sosyo-Kültürel Faktörlerin Önemi", I. Verimlilik Kongresi, 1991.

[18] Özkılıç, Ö. (2014), Risk Değerlendirmesi (Atex Direktifl eri-Patlayıcı OrtamlarBüyük Endüstriyel Kazaların Önlenmesi ve Etkilerinin Azaltılması-Kantitatif Risk Değerlendirme). Ankara: Türkiye İşveren Sendikaları Konfederasyonu Yayınları.

[19] Ocaktan, E. (2014). Meslek Hastalıkları. Çalışan Sağlığı ve Güvenliği Eğitimi ve Çalıştayı. 02-05 Aralık. Ankara.