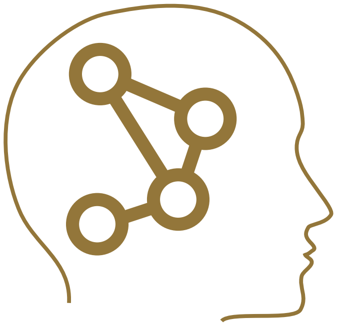


THE JOURNAL OF COGNITIVE SYSTEMS

an international, peer-reviewed
indexed, and open-access periodical

www.dergipark.gov.tr/jcs

VOLUME 05
NUMBER 02
YEAR 2020
ISSN 2548-0650

THE JOURNAL OF
COGNITIVE SYSTEMS

An international, peer-reviewed,
indexed, and open-access periodical

ISSN 2548-0650

OWNER

ISMAIL KOYUNCU, (RECTOR), Istanbul Technical University, Turkey

GENERAL PUBLICATION DIRECTOR

S. SERHAT SEKER, Istanbul Technical University, Turkey

EDITOR-IN-CHIEF

OZER CIFTCIOGLU, Delft University of Technology, Delft, The Netherlands

EDITORIAL BOARD

T. CETIN AKINCI, Istanbul Technical University, Turkey
CEMIL COLAK, Inonu University, Turkey
EVREN DAGLARLI, Istanbul Technical University, Turkey

SCIENTIFIC COMMITTEE

ABRAHAM LOMI (Indonesia)
ALTAN CAKIR (Turkey)
A.TARIK ZENGIN (Turkey)
AYKUT HOCANIN (Turkish Republic of Northen Cyprus)
BELLE R. UPADHYAYA (USA)
BERK USTUNDAG (Turkey)
BURAK BARUTCU (Turkey)
BURCU OZDEMIR (Turkey)
CHANDAN KUMAR CHANDA (India)
DENIZ TURKPENCE (Turkey)
ERKAN KAPLANOGLU (USA)
ESEN YILDIRIM (Turkey)
GOKHAN ERDEMIR (Turkey)
HAKAN TEMELTAS (Turkey)
HASAN DEMIREL (Turkish Republic of Northen Cyprus)
JELENA DIKUN (Lithuania)
KUNIHIKO NABESHIMA (Japan)
MURAT OKATAN (Turkey)
MUSA YILMAZ (Turkey)
NECDET OSAM (Turkish Republic of Northen Cyprus)
OKYAY KAYNAK (Turkey)
OLEKSII TURUTA (Ukraine)
OSMAN NURI UCAN (Turkey)
OMER FARUK ERTUGRUL (Turkey)
RITUPARNA DATTA (South Korea)
SALIH BARIS OZTURK (Turkey)
TANJU SURMELI (Turkey)
UFUK KORKMAZ (Turkey)

AIM & SCOPE

The Journal publishes original papers in the field of artificial intelligence, biomedical, quantum information, big data analysis and statistical areas, which are related with cognitive science and engineering, as well as the applications of the cognitive studies in social areas. Letter to the editor is also encouraged.

THE JOURNAL OF
COGNITIVE SYSTEMS

THE JOURNAL OF COGNITIVE SYSTEMS (JCS) is published bi-annually. Responsibility for the content rests upon the authors and not upon the JCS or the publisher. **Reuse Rights and Reprint Permissions:** Educational or personal use of this material is permitted without fee, provided such use: i) is not made for profit; and ii) includes this notice and a full citation to the original work on the first page of the copy; and iii) does not imply JCS endorsement of any third-party products or services. Authors and their companies are permitted to post their JCS-copyrighted material on their own web servers without permission, provided that the JCS copyright notice and a full citation to the original work appear on the first screen of the posted copy. Permission to reprint/republish this material for commercial, advertising or promotional purposes or for creating new collective works for resale or redistribution must be obtained from JCS by writing to the JCS General Publication Director Prof. Dr. S. Serhat Seker, Faculty of Electrical and Electronics Engineering, Istanbul Technical University, Istanbul, Turkey; or sekers@itu.edu.tr.
Copyright © 2016 JCS. All rights reserved.

CONTACT

Prof. Dr. Özer Ciftcioglu
Editor-in-Chief of The Journal of Cognitive Systems
Delft University of Technology, The Netherlands
Istanbul Technical University, Istanbul, Turkey

E-MAIL

cognitive@itu.edu.tr

WEB PAGE

www.cognitive.itu.edu.tr

URL

www.dergipark.gov.tr/jcs

THE JOURNAL OF COGNITIVE SYSTEMS

VOLUME 05, NUMBER 02

DECEMBER 2020

CONTENTS

ASSESSMENT OF ASSOCIATIVE CLASSIFICATION APPROACH FOR PREDICTING MORTALITY BY HEART FAILURE

Z. Kucukakcali, I. Balikci Cicek, E. Guldogan, and C. Colak

Abstract— *Aim:* This study aims to predict mortality status by heart failure and to determine the related factors by applying the relational classification method, one of the data mining methods, on the open-access heart failure data set.

Materials and Methods: In this study, the associative classification model has been applied to the open-access data set named "Heart Failure Prediction". The performance of the model was evaluated by accuracy, balanced accuracy, sensitivity, selectivity, positive predictive value, negative predictive value, and F1-score.

Results: Accuracy, balanced accuracy, sensitivity, specificity, positive predictive value, negative predictive value, and F1-score values obtained from the model were 0.866, 0.819, 0.688, 0.951, 0.868, 0.865 and 0.767 respectively.

Conclusion: The findings obtained from this study showed that successful results were obtained in the study performed with the associative classification model on the heart failure data set. Also, certain rules regarding the disease to be used in preventive medicine practices were obtained with the proposed model.

Keywords— Heart failure, classification, association rules, relational classification.

1. INTRODUCTION

HEART failure (HF) is a medical syndrome caused by cardiac structural or functional abnormalities, accompanied by typical symptoms and signs that occur due to decreased cardiac output (CO) and/or increased intracardiac pressure at rest, stress, and effort. These symptoms are shortness of breath, pretibial edema, weakness; findings are increased jugular venous pressure, pulmonary rales, and peripheral edema. [1]. HF causes serious mortality and morbidity and poses a serious burden to the healthcare system worldwide. Heart failure is seen in one in 10 people aged 75 and over in developed countries. In the 2025 program of the World Health Organization (WHO), it has been stated that the burden of HF disease is a potential target to be reduced. In a study carried out at our country, the incidence of HF in Turkey was determined to be 2.9%. Besides, it has been shown that CG affects 1.5 million people, and 3 million people are at risk soon [2, 3].

Data mining can be defined simply as the discovery of useful information hidden in data. Data mining enables researchers to make effective and informed decisions with techniques offered by different disciplines such as artificial intelligence, machine learning, statistics, and optimization. It also enables revealing hidden, implicit, beneficial relationships, patterns, relations, or trends that are difficult to reveal with classical methods. Data mining is the search for the relations and rules that will allow us to make predictions about the future of a large amount of data using computer programs [4]. Associative classification is a branch of scientific work, known as data mining. Associative classification combines the association rule and classification, two known methods of data mining, to create a model for predictive purposes. In other words, associative classification is a type of classification approach that is created with a set of rules obtained by the association rule mining to create classification models. One of the important advantages of using a classification based on association rules according to classical classification approaches is that the output of an associative classification algorithm is represented by simple if-then rules, making it easier for the users to understand and interpret it [5].

This study aims to predict mortality status by heart failure and to determine the related factors by applying the associative classification method, one of the data mining methods, on the open-access heart failure data set.

2. MATERIAL AND METHODS

2.1. Dataset

In the study, the associative classification model, which is a data mining method that combines classification and association rules methods, has been applied to an open-access data set named "Heart Failure Prediction" [6].

There are 299 patients in the data set used. 96 (32.1%) of these patients died after a certain period of follow-up. Explanations about the variables and their properties in the data set are given in Table I.

Zeynep KUCUKAKCALI, Inonu University Department of Biostatistics and Medical Informatics, Faculty of Medicine, Malatya, Turkey, (zeynep.tunc@inonu.edu.tr)

İpek BALIKCI CICEK, Inonu University Department of Biostatistics and Medical Informatics, Faculty of Medicine, Malatya, Turkey, (ipek.balikci@inonu.edu.tr)

Emek GULDOGAN, Inonu University Department of Biostatistics and Medical Informatics, Faculty of Medicine, Malatya, Turkey, (emek.guldogan@inonu.edu.tr)

✉ **Cemil COLAK**, Inonu University Department of Biostatistics and Medical Informatics, Faculty of Medicine, Malatya, Turkey, (cemil.colak@inonu.edu.tr)

Manuscript received August 12, 2020; accepted Sep. 29, 2020.
Digital Object Identifier:

TABLE I
EXPLANATIONS ABOUT THE VARIABLES IN THE DATASET AND THEIR PROPERTIES

Variable	Variable Description	Variable Type	Variable Role
age	Decrease of red blood cells or hemoglobin	Quantitative	Predictor
anemia	Level of the CPK enzyme in the blood (mcg/L) (boolean)	Qualitative	Predictor
creatinine_phosphokinase	Level of the CPK enzyme in the blood (mcg/L)	Quantitative	Predictor
diabetes	If the patient has diabetes (boolean)	Qualitative	Predictor
ejection_fraction	Percentage of blood leaving the heart at each contraction (percentage)	Quantitative	Predictor
high_blood_pressure	If the patient has hypertension (boolean)	Qualitative	Predictor
platelets	Platelets in the blood (kiloplatelets/mL)	Quantitative	Predictor
serum_creatinine	Level of serum creatinine in the blood (mg/dL)	Quantitative	Predictor
serum_sodium	Level of serum sodium in the blood (mEq/L)	Quantitative	Predictor
sex	Woman or man (binary)	Qualitative	Predictor
smoking	If the patient smokes or not (boolean)	Qualitative	Predictor
time	Follow-up period (days)	Quantitative	Predictor
death_event	If the patient deceased during the follow-up period (boolean)	Qualitative	Output

3. ASSOCIATIVE CLASSIFICATION

Rules of association are a type of unsupervised data mining that looks for the relationship between records in a data set. Association rules are the process of determining the events or features that occur together. Association rules are often expressed as if it happens, then this happens. Mostly used in descriptive data analysis, data preprocessing, determining discrete values, and finding trends and relationships [7]. Association rules are rules with support and confidence measurements in the form of "IF- precursor expression-, IF- successor expression" [8].

Association rules share many common features with classification. Both use rules to characterize regularities in a dataset. However, these two methods differ greatly in their goals. While classification focuses on prediction, association rules focus on providing information to the user. In particular, it focuses on detecting and characterizing unexpected relationships between data items. [9].

Associative classification is a data mining method that combines classification and association rules methods to make predictions. In other words, an associative classification is an approach that uses rules obtained with association rules to create classification models. Associative classification is a special association rule mining with the target/response/dependent/class variable to the right of the rule obtained. In a rule such as $X \rightarrow Y$, Y must be the target / response / dependent / class variable. One of the principal benefits of using a classification based on association rules according to classical classification approaches is that simple if-then rules represent the output of an associative classification algorithm. This advance makes it easier for the user to understand and interpret the results [10].

3.1. Performance evaluation criteria

The classification matrix for the calculation of performance metrics is given in Table II.

TABLE II
THE METRICS OF MODEL'S CLASSIFICATION PERFORMANCE

		Real		
		Positive	**Negative**	**Total**
Predicted	Positive	True positive (TP)	False negative (FN)	TP+FN
	Negative	False positive (FP)	True negative (TN)	FP+TN
	Total	TP+FP	FN+TN	TP+TN+FP+FN

Accuracy = (TP+TN)/(TP+TN+FP+FN)

Balanced accuracy = [[TP/(TP+FP))] + [TN/(TN+FN)]]/2

Sensitivity = TP/(TP+FP)

Specificity = TN/(TN+FN)

Positive predictive value = TP/(TP+FN)

Negative predictive value =TN/(TN+FP)

F-score = (2*TP)/(2*TP+FP+FN)

4. DATA ANALYSIS

Quantitative data are summarized by median (minimum-maximum), and qualitative variables are given by number and percentage. Normal distribution was evaluated with the Kolmogorov-Smirnov test. In terms of input variables, the existence of a statistically significant difference and the relationship between the categories of the output variable, "who died during follow-up" and "who did not die during follow-up", was examined using the Mann-Whitney U and Pearson Chi-square test. The values of $p<0.05$ were deemed statistically significant. In all analyzes, IBM SPSS Statistics 26.0 for the Windows package program was used.

5. RESULTS

Descriptive statistics related to the target variable examined in this study are presented in Table 3 and Table 4. A statistically significant difference exists between output variable classes in terms of age, ejection_fraction, serum_creatinine, serum_sodium, time variables. (p<0.001)

TABLE III

DESCRIPTIVE STATISTICS FOR QUANTITATIVE INPUT VARIABLES

Variables	death_event		**P* value**
	Survived patients **Median (min-max)**	Dead patients **Median (min-max)**	
age	60(40-90)	65(42-95)	**<0.001**
creatinine_phosphokinase	245(30-5209)	259(237-861)	0,684
ejection_fraction	38(17-80)	30(14-70)	**<0.001**
platelets	263000(25100-850000)	258500(47000-621000)	0,425
serum_creatinine	1(0,5-6,1)	1,3(0,6-9,4)	**<0.001**
serum_sodium	137(113-148)	136(116-146)	**<0.001**
time	172(12-285)	45(4-241)	**<0.001**

*: Mann Whitney U test

TABLE IV

DESCRIPTIVE STATISTICS FOR QUALITATIVE INPUT VARIABLES

Variables		death_event		**P* value**
		Survived patients	Dead patients	
anaemia	absence	120(59.1%)	50(52.1%)	0.252
	presence	83(40.9%)	46(47.9%)	
diabetes	absence	118(58.1%)	56(58.3%)	0.973
	presence	85(41.9%)	40(41.7%)	
high_blood_pressure	absence	137(67.5%)	57(59.4%)	0.170
	presence	66(32.5%)	39(40.6%)	
Gender	woman	71(35.0%)	34(35.4%)	0.941
	man	132(65.0%)	62(64.6%)	
smoking	absence	137(67.5%)	66(68.75%)	**0.827**
	presence	66(32.5%)	30(31.25%)	

***: Pearson's chi-square test**

The classification matrix of the associative classification model used to classify the heart failure dataset in this study is given below in Table V.

TABLE V

CLASSIFICATION MATRIX FOR THE ASSOCIATIVE CLASSIFICATION MODEL

Prediction	**Reference**		
	Survived patients	Dead patients	**Total**
Survived patients	193	30	223
Dead patients	10	66	76
Total	203	99	299

The values for the metrics of the classification performance of the model are given in Table 6. Accuracy, balanced accuracy, sensitivity, specificity, positive predictive value, negative predictive value, and F1-score values obtained from the model were 0.866, 0.819, 0.688, 0.951, 0.868, 0.865 and 0.767 respectively.

TABLE VI

VALUES FOR THE CLASSIFICATION PERFORMANCE METRICS OF THE MODEL

Metric	**Value**
Accuracy	0.866
Balanced accuracy	0.819
Sensitivity	0.688
Specificity	0.951
Positive predictive value	0.868
Negative predictive value	0.865
F1-score	0.767

Table 7 shows the first 5 of the association rules used by the classification algorithm. As expressed in Table 7, when anaemia=0,ejection_fraction=[27.5,80),serum_sodium=[136,148),time=[73.5,285) are considered, the patient's survival probability is 98.7 %.

TABLE VII

Left-hand side rules	**Right-hand side rules**	**Support**	**Confidence**	**Freq.**
{anaemia=0,ejection_fraction=[27.5,80),serum_sodium=[136,148),time=[73.5,285)}	{death_event=0}	0.247	0.987	74
{age=[40,71),diabetes=0,ejection_fraction=[27.5,80),serum_creatinine=[0.5,1.81),serum_sodium=[136,148),time=[73.5,285)}	{death_event=0}	0.214	0.985	64
{age=[40,71),ejection_fraction=[27.5,80),serum_creatinine=[0.5,1.81),sex=1,time=[73.5,285)}	{death_event=0}	0.294	0.978	88
{age=[40,71),anaemia=0,ejection_fraction=[27.5,80),platelets=[1.28e+05,8.5e+05),serum_creatinine=[0.5,1.81),time=[73.5,285)}	{death_event=0}	0.288	0.977	86
{age=[40,71),diabetes=0,ejection_fraction=[27.5,80),platelets=[1.28e+05,8.5e+05),serum_creatinine=[0.5,1.81),time=[73.5,285)}	{death_event=0}	0.261	0.975	78

If age =[40,71),diabetes = 0, ejection_fraction=[27.5,80), serum_creatinine=[0.5,1.81),serum_sodium=[136,148),time=[73.5,285) are considered, the patient's survival probability is 98.5 %.
As age=[40,71),ejection_fraction=[27.5,80),serum_creatinine=[0.5,1.81),sex=1,time=[73.5,285) are considered, the patient's survival probability is 97.8 %. If age=[40,71),anaemia=0,ejection_fraction=[27.5,80),platelets=[1.28e+05,8.5e+05),serum_creatinine=[0.5,1.81),time=[73.5,285) are considered, the patient's survival probability is 97.7 %. age=[40,71),diabetes=0,ejection_fraction=[27.5,80),platelets=[1.28e+05,8.5e+05),serum_creatinine=[0.5,1.81),time=[73.5,285) are considered, the patient's survival probability is 97.5 %.

6. DISCUSSION

Cardiac failure is the final stage of all forms of cardiac disease, a health problem that is increasing in prevalence and incidence, affecting at least 23 million people worldwide. Heart failure is still one of the most common cardiovascular diseases in the world, and similar clinical results are seen in our country. In recent years, the incidence of heart failure has continued to increase all over the world, and death rates are still at very high levels. Advances in the treatment of cardiovascular disorders increase the survival and lifespan of individuals. Therefore, the follow-up and treatment of patients with heart failure is becoming more important and remains an open area for research and new developments [11, 12].

Association rules, one of the descriptive models of data mining, are methods that analyze the coexistence of events. These relationships are based on the coexistence of data elements and express the co-occurrence of events together with certain possibilities. Classification analysis is one of the basic methods of machine learning and is used by a large scientific community. Classification is an estimation process that assigns each observation in the dataset to the predetermined classes under certain rules [13]. Associative classification makes classification by combining two common data mining methods, association rules, and classification methods. In recent years, association rules methods have been successfully used to create correct classifiers in associative classification [5].

In this study, the associative classification model, one of the data mining methods, was applied to the data set named “Heart Failure Prediction”, which is an open-source data set. For this purpose, different factors (explanatory variables) that may be associated with heart failure (dependent variable) were estimated with the relational classification model, and rules were obtained. According to the experimental results, from the performance metrics obtained from the model, accuracy, balanced accuracy, sensitivity, specificity, positive predictive value, negative predictive value, and F1-score values obtained from the model were 0.866, 0.819, 0.688, 0.951, 0.868, 0.865 and 0.767 respectively.

In a study conducted with the same data set, the results were compared using ten different machine learning methods. According to the results of this study, the highest accuracy was obtained as 0.74 with the Random Forest model [6]. In this study, an accuracy of 0.866 was obtained, and rules about the disease were also obtained.

As a result, the associative classification model used produced successful results in the study conducted with the heart failure data set. Besides, certain rules regarding the disease to be used in preventive medicine practices have been obtained with this model.

REFERENCES

[1] P. Ponikowski, A. Voors, S. Anker, H. Bueno, J. Cleland, A. Coats, et al., "Authors/Task Force Members; Document Reviewers (2016) 2016 ESC Guidelines for the diagnosis and treatment of acute and chronic heart failure: The Task Force for the diagnosis and treatment of acute and chronic heart failure of the European Society of Cardiology (ESC). Developed with the special contribution of the Heart Failure Association (HFA) of the ESC," Eur J Heart Fail, vol. 18, pp. 891-975, 2016.

[2] G. Fonarow, K. Adams Jr, W. Abraham, C. Yancy, and W. Boscardin, "ADHERE Scientific Advisory Committee Study Group and Investigators. Risk stratification for in-hospital mortality in acutely decompensated heart failure: classification and regression tree analysis," Jama, vol. 293, pp. 572-80, 2005.

[3] M. B. Yılmaz, A. Çelik, Y. Çavuşoğlu, L. Bekar, E. Onrat, M. Eren, et al., "Snapshot evaluation of heart failure in Turkey: Baseline characteristics of SELFIE-TR," Turk Kardiyoloji Dernegi arsivi: Turk Kardiyoloji Derneginin yayin organidir, vol. 47, pp. 198-206, 2019.

[4] H. Akpınar, "Veri tabanlarında bilgi keşfi ve veri madenciliği," İstanbul Üniversitesi İşletme Fakültesi Dergisi, vol. 29, pp. 1-22, 2000.

[5] F. A. Thabtah, "A review of associative classification mining," Knowledge Engineering Review, vol. 22, pp. 37-65, 2007.

[6] D. Chicco and G. Jurman, "Machine learning can predict survival of patients with heart failure from serum creatinine and ejection fraction alone," BMC Medical Informatics and Decision Making, vol. 20, p. 16, 2020/02/03 2020.

[7] U. M. Fayyad, G. Piatetsky-Shapiro, P. Smyth, and R. Uthurusamy, "Advances in knowledge discovery and data mining," 1996.

[8] D. T. Larose and C. D. Larose, Discovering knowledge in data: an introduction to data mining vol. 4: John Wiley & Sons, 2014.

[9] N. Ye, The handbook of data mining: CRC Press, 2003.

[10] F. Thabtah, "A review of associative classification mining," The Knowledge Engineering Review, vol. 22, pp. 37-65, 2007.

[11] G. S. Bleumink, A. M. Knetsch, M. C. Sturkenboom, S. M. Straus, A. Hofman, J. W. Deckers, et al., "Quantifying the heart failure epidemic: prevalence, incidence rate, lifetime risk and prognosis of heart failure: The Rotterdam Study," European heart journal, vol. 25, pp. 1614-1619, 2004.

[12] A. Mosterd and A. W. Hoes, "Clinical epidemiology of heart failure," Heart, vol. 93, pp. 1137-1146, 2007.

[13] İ. Perçın, F. H. Yağin, E. Güldoğan, and S. Yoloğlu, "ARM: An Interactive Web Software for Association Rules Mining and an Application in Medicine," in 2019 International Artificial Intelligence and Data Processing Symposium (IDAP), 2019, pp. 1-5.

BIOGRAPHIES

Zeynep KÜÇÜKAKÇALI obtained her BSc. degree in mathematics from Çukurova University in 2010. She received MSc. degree in biostatistics and medical informatics from the Inonu University in 2018. She currently continues Ph.D. degrees in biostatistics and medical informatics from the Inonu University. In 2014, she joined the Department of Biostatistics and Medical Informatics at Inonu University as a researcher assistant. Her research interests are cognitive systems, data mining, machine learning, deep learning.

İpek BALIKÇI ÇİÇEK obtained her BSc. degree in mathematics from Çukurova University in 2010. She received MSc. degree in biostatistics and medical informatics from the Inonu University in 2018. She currently continues Ph.D. degrees in biostatistics and medical informatics from the Inonu University. In 2014, she joined the Department of Biostatistics and Medical Informatics at Inonu University as a researcher assistant. Her research interests are cognitive systems, data mining, machine learning, deep learning.

Emek GÜLDOĞAN obtained his BSc. degree in Computer Engineering from Middle East Technical University in 2001. He received MSc. degree in biostatistics and medical informatics from the Inonu University in 2005, and Ph.D. degrees in biostatistics and medical informatics from the Inonu University in 2017. He is currently working as an assistant professor of the Department of Biostatistics and Medical Informatics at Inonu University and as the information processing manager at Turgut Özal Medical Center. His research interests are cognitive systems, data mining, machine learning, deep learning.

Cemil ÇOLAK obtained his BSc. degree in statisticsfrom Ondokuz Mayıs University in 1999. He received MSc. degree in Biostatistics from the Inonu University in 2001, and Ph.D. degree in the Graduate Department of Biostatistics and Medical Informatics of Ankara University in 2005. His research interests are cognitive systems, data mining, reliability, and biomedical system, genetics, and bioengineering. In 2016, he joined the Department of Biostatistics and Medical Informatics at Inonu University as a Professor, where he is presently a professor. He is active in teaching and research in the general image processing, artificial intelligence, data mining, analysis.

A COGNITIVE INTEGRATED MULTI-MODAL PERCEPTION MECHANISM AND DYNAMIC WORLD MODELING FOR SOCIAL ROBOT ASSISTANTS

E. Daglarli

Abstract**—** In this study, it is investigated that robots may acquire environmental awareness based on perceptual attention. The purpose of this study is to model a human-like perception system architecture for a humanoid robot to establish robust and efficient communication with humans and its environment. The modeling of the robot's environment and deficiencies in the coordination of multi-modal perceptual stimuli are the main challenges for achieving this purpose. Previous works do not fulfill some of these requirements. We present a novel solution which covers a cognitive multi-modal integrated perception system. The computational framework contains features of basic feature extraction, recognition tasks, and spatial-temporal inference. In addition, it provides to help modeling perceptual attention and awareness. It is convenient that implementation works are carried out on the developed open-source software involving this architecture for social robots. The model's performance can be evaluated by various interaction scenarios. In the future, it is considered that the framework presented in this study will guide to develop next-generation cognitive models for social robots.

Keywords**—** *Cognitive perception, attention modeling, perceptual awareness, human-robot interaction.*

1. INTRODUCTION

THE perceptual ability to interact with their social environment is very important not only for humans but also for social robots. These skills depend on their spatial-temporal world model representation, perceptual awareness, and attention (focus) abilities [1]. In nature, these functions and features are biologically achieved by sensory and perceptual regions of the neocortex in the human brain. The anatomical construct of the neocortex contains two major cerebral structures such as frontal and posterior parts [2]. Cognitive skills related to perception functions are involved in the posterior part of the cerebral cortex [2, 3]. This part of the cortex is divided into three sub-regions such as occipital, parietal, and temporal lobes. The occipital lobe which hosts regions of the primary visual cortex realizes post feature extraction on visual stimuli. The temporal lobe involves pattern recognition on visual and auditory stimuli. The parietal lobe which accepts visual and somatosensory stimuli is responsible for spatial perception [3]. However, smart digital assistants or social robots have confronted severe difficulties in accomplishing these abilities during human-machine interaction experiments [4]. To establish better interaction between the social robot and the human, cognitive perception systems are currently very critical issues for human-robot interaction (HRI) studies and social robotics [5].

Evren DAGLARLI. is with Computer and Informatics Engineering Faculty Istanbul Technical University, Istanbul, Turkey,
(e-mail: evren.daglarli@itu.edu.tr, evrendaglarli@ieee.org)

Manuscript received Sep 15, 2020; accepted Oct 18, 2020.

In daily life, as personal assistants, social robots which contain cognitive-perceptual functions can be utilized to support individuals attempting to interact with their social environment [6, 7]. Therefore, social robots should be equipped with a human-like perception system that incorporates spatial-temporal cognitive perception skills to interpret world model representation and evaluates the human-machine interaction via joint attention in a shared workspace [1].

The spatial cognition deals with environmental (situation) awareness which involves spatial perception (e.g. locations, orientations, distances, and movements) of the objects [1, 8]. The temporal cognition deals with mid-level abstraction processes that involve temporal or non-spatial encoding (e.g. color, shape) of the objects, recognition of the patterns (e.g. objects, faces, spoken words). It is a very hard problem that high-level cognitive skills that make it possible to respond to multi-modal perceptual stimuli can provide some properties having environmental awareness as well as capabilities of pattern recognition and modeling of attention [1, 9]. The modeling of the robot's social environment (e.g. the spatial world model representation and the interaction of physical behavior models) is one of the biggest problems [10]. Another major issue is the temporal perception that involves event (or situation) based world model representation [11-13]. Deficiencies in the representation of the world modeling or the coordination of multi-modal perceptual stimuli may cause interaction failures.

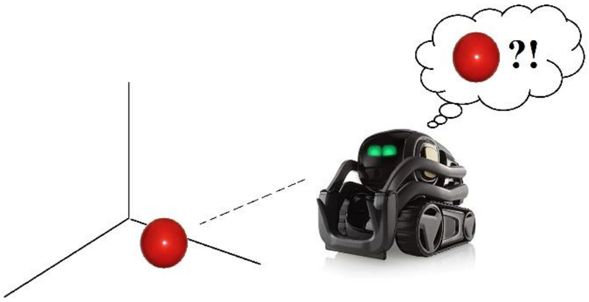


Fig.1. Cognitive perception in social robots.

The purpose of this study is to sketch a cognitive integrated multi-modal perception system for mobile robots that can be used as a social assistant. The computational approximation of the posterior neocortex may guide in which the cognitive perception system is developed in a software framework. It is expected that this solution may succeed in the case of world model representation having dynamic environments with uncertainties. This approach provides several contributions. For example, cognitive perception functions accepting multi-modal stimuli (e.g. visual, auditory, and somatosensory) perform several tasks including feature extraction, pattern recognition, and spatial perception. Perceptual coordination property includes several skills such as perceptual association (or sensory fusion) and competition between stimuli. This property plays an important role in the modeling of perceptual attention. To achieve them, supervised and unsupervised learning progress are performed on different modules of the cognitive perception system.

The paper follows with section 2 that express related works. In section 3, design principles of the computational framework of cognitive integrated multi-modal perception system realizing the world model representation model and spatial-temporal situation awareness under dynamic environments with uncertainties are presented. After all, discussion, concluding remarks, and future works are expressed in section 4.

2. RELATED WORKS

Computational cognitive architectures have been developed to solve perceptual and environmental modeling problems for social robots. The number of projects are rapidly increasing, and they promise to increase as a developing subject in the future. Certain quite interesting instances of computational architectures based on cognitive perception ensuring world model representation and attention model have been presented in recent years.

Inceoglu et al. present a visual scene representation framework for service robots to produce and keep up exact models of their workplaces for object manipulation [14]. Their framework is intended to ensure a conventional system to both humanoid and manipulator utilizing various sorts of algorithms and sources of vision data streams. Different perception algorithms processing visual data are implemented to develop and persistently refreshing a world model representation.

Kim et al. [15] investigated a curiosity-driven Dynamic World Model Learning (AWML) framework. To realize that, as visually exploring a 3D physical workplace plenty with the refinement of representative real-world agents, a curious agent building world models was built [16]. They aimed an AWML framework guided by -Progress: efficient and adaptive learning progress-based curiosity indicator and introduce that -Progress inherently provides ascend to an exploration policy. Thus, their - Progress-driven controller accomplished altogether higher AWML performance than controllers embodied with cutting edge exploration methodologies like random network distillation and model disagreement.

Riedelbauch and Henrich introduced a highly adaptable method to the human-robot collaboration framework in which a robot dynamically chooses actions contributing to a common objective from a given behavior model [17]. According to this, a world model built from eye-in-hand camera images derived knowledge on the task progress. They utilized a human-aware world model sustaining an observation for trust in stored items concerning the ongoing human presence and past assignment progress, since data created by fractional workspace perceptions is not suitable over time, as humans may interact with resources. Their contribution was a decision-making mechanism utilizing this confidence indication score to interleave task operations with an active vision to replenish the world model. Comprehensive system tests spread different kinds of human interest in various benchmark assignments through the recreation of streamlined, fractionally randomized human models. The results of their system displayed scores related to functions for various parametrizations of their human-robot joining structure.

Rosinol et al. offered an integrated model for dynamic spatial perception: 3D Dynamic environment networks. The term environment networks expressed in their framework composed of a directed network including nodes as entities in the scene (e.g., objects, walls, rooms), and edges as relations (e.g., inclusion, adjacency) between nodes [18]. This notion is extended by Dynamic environment networks (DENs) to show dynamic scenes with moving agents (e.g., humans, robots), and to incorporate dynamic data that assists planning and decision-making (e.g., spatiotemporal relations, topology at different levels of abstraction). Another novelty is to realize an automatic Spatial Perception eNgine (SPIN) to develop a DEN from visual-inertial data. They studied cutting edge strategies for human and object recognition and posture computation, and depicted how to detect objects, robot, and human nodes in crowded places. The visual-inertial SLAM and the dense human mesh tracking are incorporated by their work. Besides, they also offered algorithms to get hierarchical models of indoor areas (e.g., corridors, warehouses, hall, lobby, rooms) and their relations. Their last innovation is to show a spatial perception engine in a photo-realistic Unity-based simulator. 3D Dynamic environment networks technique seems a deep effect on action selection, task planning, human-robot interaction, long-term autonomy, and environment modeling.

Venkataraman et al. dealt with the issue of creating overall 3D models for genuine items utilizing a robot, that can expel items from the mess for better order [19]. They realized models of grasped objects using simultaneous manipulation and monitoring. Then their model processed visual data utilizing a kinematic representation of the robot to incorporate observations from various scenes and cancel background noise. For evaluation of their model, they employed a robot composed of a mobile platform with a manipulator and mounted with an RGBD camera to assemble voxelized representations of undefined items and then classify them into new categories.

Persson et al. dealt with the issue of semantic world representation by merging statistical training and item linking. Their paradigm employs a top-down item binding approach based on full permanent property scores observed from perceptual sensor data [20]. According to their study, a binding pairing model trains to sustain item entities and is verified utilizing a big set of trained manually labeled ground truth data of real-world items. To add more complicated instances, a high-level probabilistic item tracker was organized with the binding architecture and handled the tracking of occluded items with reasoning about the state of unobserved items. They displayed the performance of their system with scenarios including the shell game scenario. In this scenario, it is explained how binder items are stored by maintaining relations through probabilistic reasoning.

Martires et al. aimed a semantic scene representation paradigm based on top-down item linking utilizing an item-originated model of the world [21]. Perceptual linking processes continuous perceptual sensor data and sustains a correlation to a symbolic model. They continued the descriptions of linking to conduct multi-modal probability distributions and symbol linking model to a probabilistic logic reasoning for performing inference. In addition, they utilized stochastic correlation learning to allow the linking system to train symbolic information in the form of a set of probabilistic relationships of the world model from noisy and sub-symbolic sensor input. As exploiting the significance of relationships to reason about the state of items which are not directly detected by sensory input data, their system that incorporates perceptual linking and statistical relational learning, may sustain a semantic world model of all the items that have been perceived over time. They indicated the performance of the framework to verify their system, to execute probabilistic reasoning over multi-modal likelihood, the learning of probabilistic logical rules from linked items generated by perceptual observations.

3. COGNITIVE PERCEPTION

The main responsibilities of the cognitive integrated multi-modal perception system are to build world model representation for dynamic and uncertain environments and to support to constitute joint attention with robot's partners. In this study, the proposed framework providing initialize and establish social interaction evaluates its operations according to three main grounds such as human, robot's entity, and environment. The proposed novel structure will represent spatio-temporal relations or features of the dynamic interaction between the robot and the world model. To improve the world model of a robot, the attention model is extracted from high-level perceptual processing. This is a key element for measuring or detecting the spatio-temporal situation awareness level of the robot during interaction with its environment.

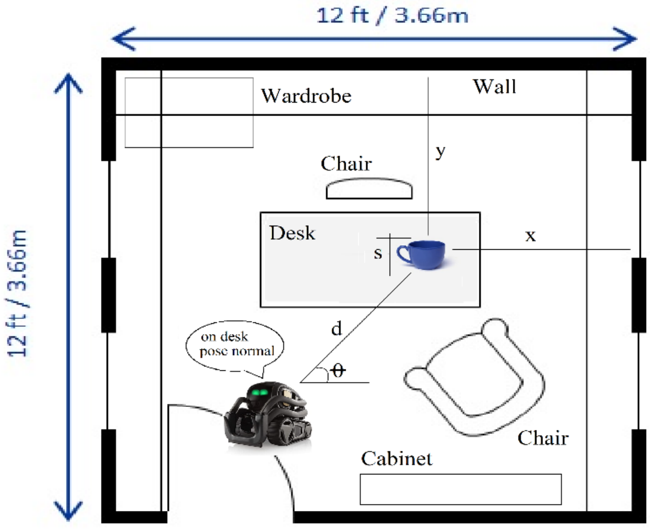


Fig.2. Spatial perception.

In order to constitute the robot's world model representation, the perceptual cognition need to deal with multi-modal perceptual data fusion and anchoring of perceptual concepts [22]. In addition, perceptual cognition of the robot is organized by its world model representation and joint attention ability involving to establish a communicational link with the other (human or robot) during the interaction. The realizing situation awareness that is the ability to resolve and discover all perceptual relations built by the robot's attention and world model representation model may be a big problem for the efficient social interaction between the human and the social robot. On the other hand, problems in manipulation between multi-modal perceptual stimuli make it difficult for the building of the robot's attention model and world model representation during interaction with the social robot. Another issue in human-robot interaction might exhibit difficulties in modeling dynamic environments and associative learning of perceptual response with multi-modal stimuli. Perceptual corruption in processes of recognition and joint attention drastically restrain human-robot interaction (HRI) in socially interactive and shared workspaces with uncertainties.

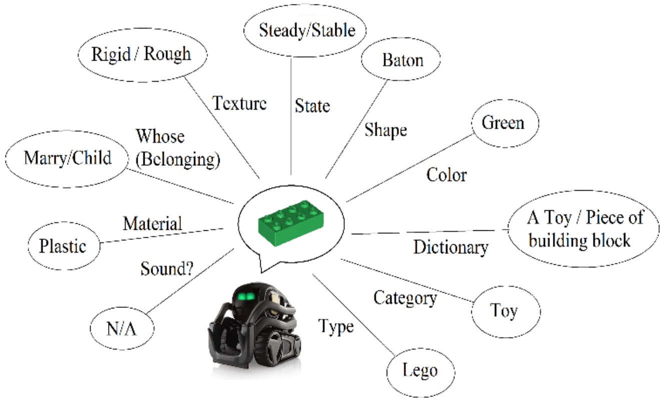
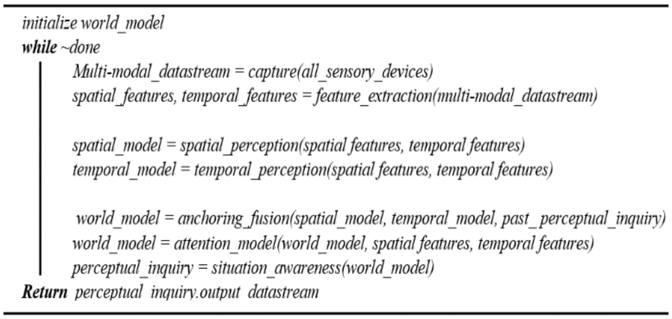


Fig.3. Temporal or non-spatial perception.

The proposed framework has multiple modalities for visual perception, auditory perception, and somatosensory (physical) perception. According to the order of data flow, the pipeline of the generic algorithm is proceeded by several main tasks such as pre-processing, feature extraction, basic perceptual operations, and processing (perceptual cognition) respectively. In our cognitive framework, the most basic functions are feature extraction and grabbers capturing multi-modal input data streams (e.g. visual, auditory, and somatosensory). It is expected that the integration of visual data sources beside of the other data streams coming from non-visual sensor modalities (e.g., microphone array, tactile sensor, laser range finder, etc.) are considered to help to achieve human-like perceptual cognition [23]. In this section, the computational framework of the cognitive perception system is described so that it constitutes the perceptual model of the robot interacting with the environment and humans. Before cognitive perception processes involving world model representation and situation awareness, some feature extraction tasks realizing segmentation, edge detection, and filtering need to be performed as data pre-processing activities. After these preliminary processes are completed, higher-level cognitive modeling is realized by developing situation awareness and attention models for a social robot so that social interaction skills such as human-like communication between robots and humans are established. Our proposed cognitive perception system introduces a generic template for world modeling which is based on anchoring knowledge between semantic concepts and perceptual relationships for social robots. It has to be analyzed about how knowledge representation can be used together with an anchoring mechanism. As a computational framework, the proposed cognitive integrated multi-modal perception system for social robot assistants annotates descriptions of recognized objects and makes anchoring these annotated representations to build a semantic network of perceptual relationships. It is considered this integrated solution in the context of social robots having perceptual cognition. In particular, we study how to constitute semantic relationships between percepts and concepts referring to features of items for social robots.



```
initialize world_model
while ~done
    Multi-modal_datastream = capture(all_sensory_devices)
    spatial_features, temporal_features = feature_extraction(multi-modal_datastream)

    spatial_model = spatial_perception(spatial features, temporal features)
    temporal_model = temporal_perception(spatial features, temporal features)

    world_model = anchoring_fusion(spatial_model, temporal_model, past_perceptual_inquiry)
    world_model = attention_model(world_model, spatial features, temporal features)
    perceptual_inquiry = situation_awareness(world_model)
Return perceptual  inquiry.output  datastream
```

Fig.4. Pseudo algorithm related to the framework.

In this system architecture, for multi-modal perceptual data fusion and knowledge anchoring in world model representation, various machine learning methods may be utilized such as Bayesian networks, deep belief networks, support vector machine (SVM), Helmholtz machines and convolutional neural network (CNN or ConvNet). Recognition tasks include supervised learning methodologies. Generalization, classification, and clustering tasks are usually driven by unsupervised learning procedures. The hybrid formulation of methods depicts the effectiveness of the computational cognitive perception system, which realizes related perceptual-cognitive skills for a social robot.

This framework which utilizes the knowledge anchoring model and determines semantic relations via machine learning tools offers human-like perception abilities enabling social robots to continuously learn object categories and affordances for achieving augmented world model representation and situation awareness with attention model.

4. CONCLUSIONS

Social robots need to interact with users to establish joint attention enabling behavioral, emotional, and intentional synchronization. To achieve these robots to do that, they should be equipped with a human-like cognitive perception system. In the present study, it is investigated design principles of a novel computational framework of cognitive perception system for social robots as the digital smart assistant. The cognitive perception system is a framework incorporating multi-modal sensor fusion to build a more realistic world model [24].

3D physical environment and its dynamics of the world model representation are evaluated by the spatial perception model. The temporal perception model of this framework is responsible for the recognition of non-spatial and event-based features. Also, this framework has an attention mechanism assisting to handle situation awareness under the supervision of semantic memory.

It is expected that this sketched cognitive framework ensuring a human-like multi-modal perception mechanism will have robust performance against dynamic environments having uncertainties. In addition, a social robot with this proposed cognitive framework that allows constituting joint attention will be successful in social environments such as interactive areas existing humans.

The proposed architecture represents a prospective model for the perceptual cognition system of social robots. Hence, the framework can be employed by digital assistants, smart devices, or social robots. The presented structure can be further developed in the future, by incorporating approximate models of the other cortical regions related to cognitive perception.

REFERENCES

[1] Yan, Z., Schreiberhuber, S., Halmetschlager, G., Duckett, T., Vincze, M., & Bellotto, N. (2020). Robot Perception of Static and Dynamic Objects with an Autonomous Floor Scrubber. arXiv preprint arXiv:2002.10158.

[2] Freud, E., Behrmann, M., & Snow, J. C. (2020). What Does Dorsal Cortex Contribute to Perception?. Open Mind, 1-18.

[3] Bear, M., Connors, B., & Paradiso, M. A. (2020). Neuroscience: Exploring the brain. Jones & Bartlett Learning, LLC.

[4] Taylor, A., Chan, D. M., & Riek, L. D. (2020). Robot-centric perception of human groups. ACM Transactions on Human-Robot Interaction (THRI), 9(3), 1-21.

[5] Ronchi, M. R. (2020). Vision for Social Robots: Human Perception and Pose Estimation (Doctoral dissertation, California Institute of Technology).

[6] Müller, S., Wengefeld, T., Trinh, T. Q., Aganian, D., Eisenbach, M., & Gross, H. M. (2020). A Multi-Modal Person Perception Framework for Socially Interactive Mobile Service Robots. Sensors, 20(3), 722.

[7] Russo, C., Madani, K., & Rinaldi, A. M. (2020). Knowledge Acquisition and Design Using Semantics and Perception: A Case Study for Autonomous Robots. Neural Processing Letters, 1-16.

[8] Lee, C. Y., Lee, H., Hwang, I., & Zhang, B. T. (2020, June). Visual Perception Framework for an Intelligent Mobile Robot. In 2020 17th International Conference on Ubiquitous Robots (UR) (pp. 612-616). IEEE.

[9] Mazzola, C., Aroyo, A. M., Rea, F., & Sciutti, A. (2020, March). Interacting with a Social Robot Affects Visual Perception of Space. In Proceedings of the 2020 ACM/IEEE International Conference on Human-Robot Interaction (pp. 549-557).

[10] Mariacarla, B. Special Issue on Behavior Adaptation, Interaction, and Artificial Perception for Assistive Robotics.

[11] Sanneman, L., & Shah, J. A. (2020, May). A Situation Awareness-Based Framework for Design and Evaluation of Explainable AI. In International Workshop on Explainable, Transparent Autonomous Agents and Multi-Agent Systems (pp. 94-110). Springer, Cham.

[12] Kridalukmana, R., Lu, H. Y., & Naderpour, M. (2020). A supportive situation awareness model for human-autonomy teaming in collaborative driving. Theoretical Issues in Ergonomics Science, 1-26.

[13] Tropmann-Frick, M., & Clemen, T. (2020). Towards Enhancing of Situational Awareness for Cognitive Software Agents. In Modellierung (Companion) (pp. 178-184).

[14] Inceoglu, A., Koc, C., Kanat, B. O., Ersen, M., & Sariel, S. (2018). Continuous visual world modeling for autonomous robot manipulation. IEEE Transactions on Systems, Man, and Cybernetics: Systems, 49(1), 192-205.

[15] Kim, K., Sano, M., De Freitas, J., Haber, N., & Yamins, D. (2020). Active World Model Learning in Agent-rich Environments with Progress Curiosity. In Proceedings of the International Conference on Machine Learning (Vol. 8).

[16] Kim, K., Sano, M., De Freitas, J., Haber, N., & Yamins, D. (2020). Active World Model Learning with Progress Curiosity. arXiv preprint arXiv:2007.07853.

[17] Riedelbauch, D., & Henrich, D. (2019, May). Exploiting a Human-Aware World Model for Dynamic Task Allocation in Flexible Human-Robot Teams. In 2019 International Conference on Robotics and Automation (ICRA) (pp. 6511-6517). IEEE.

[18] Rosinol, A., Gupta, A., Abate, M., Shi, J., & Carlone, L. (2020). 3D Dynamic Scene Graphs: Actionable Spatial Perception with Places, Objects, and Humans. arXiv preprint arXiv:2002.06289.

[19] Venkataraman, A., Griffin, B., & Corso, J. J. (2019). Kinematically-Informed Interactive Perception: Robot-Generated 3D Models for Classification. arXiv preprint arXiv:1901.05580.

[20] Persson, A., Dos Martires, P. Z., De Raedt, L., & Loutfi, A. (2019). Semantic relational object tracking. IEEE Transactions on Cognitive and Developmental Systems, 12(1), 84-97.

[21] Zuidberg Dos Martires, P., Kumar, N., Persson, A., Loutfi, A., & De Raedt, L. (2020). Symbolic Learning and Reasoning with Noisy Data for Probabilistic Anchoring. arXiv, arXiv-2002.

[22] Chiu, H. P., Samarasekera, S., Kumar, R., Matei, B. C., & Ramamurthy, B. (2020). U.S. Patent Application No. 16/523,313.

[23] Wang, S., Wu, T., & Vorobeychik, Y. (2020). Towards Robust Sensor Fusion in Visual Perception. arXiv preprint arXiv:2006.13192.

[24] Xue, T., Wang, W., Ma, J., Liu, W., Pan, Z., & Han, M. (2020). Progress and prospects of multi-modal fusion methods in physical human-robot interaction: A Review. IEEE Sensors Journal.

BIOGRAPHIES

Evren Daglarli received the B.Sc. from Marmara University in 2004, the M.Sc. degree in mechatronics engineering with a focus on intelligent systems and robotics from Istanbul Technical University (ITU) in 2007, and the Ph.D. degree in control and automation engineering with a focus on computational cognitive-neuroscience, human-robot interaction from ITU in 2019. He has worked as a Research Assistant with the Department of Electrical and Electronics Engineering, Atilim University. Dr. Daglarli has many publications that are published in international journals, conferences/symposiums related to mechatronics, intelligent control systems, and robotic areas. As a Researcher, he took some duties and responsibilities in several national/international projects. Also, he has worked as a Project Engineer and the Department Manager at a private sector technology and engineering company. He is currently working as a Faculty Member and a Lecturer/Instructor with the Computer Engineering Department, Faculty of Computer and Informatics Engineering, Istanbul Technical University. Dr. Daglarli is a member of IEEE.

ASSOCIATIVE CLASSIFICATION APPROACH CAN PREDICT PROSTATE CANCER BASED ON THE EXTRACTED ASSOCIATION RULES

I. Balikci Cicek, Z. Kücükakcali and C. Colak

Abstract*— *Aim: This study aims to classify the diagnosis status of prostate cancer and determine the related factors by applying the associative classification method, one of the data mining methods, to the open-access prostate cancer data set.

Materials and Methods: In the current study, an open-access data set named "Prostate Cancer" is used for classification. The performance of the associative classification model is evaluated using the classification performance metrics such as sensitivity, selectivity, accuracy, balanced accuracy, negative predictive value, positive predictive value, and F1-score.

Results: According to the prostate cancer classification results obtained from the associative classification model, the accuracy, balanced accuracy, sensitivity, specificity, positive predictive value, negative predictive value, and F1 score values were obtained as 0.968, 0.789, 0.9, 0.879, 0.938, 0.882 and 0.923, respectively.

Conclusion: In the analysis of the open-access data set, the proposed associative classification model has distinctively successful results in classifying prostate cancer on the performance metrics.

Keywords— Prostate cancer, data mining, association rules, classification, associative classification.

1. INTRODUCTION

PROSTATE cancer is a type of cancer characterized by the abnormal division of cells in the prostate gland. Prostate cancer is the second most common cancer in men and the fifth leading cause of death worldwide [1]. However, it is the most common cancer diagnosed in men over middle age, both in developed and developing countries. While the probability of developing prostate cancer in a male between the ages of 0-39 is 0.01%, it is 2.58% for a man between the ages of 40-59, and 14.7% between the ages of 60-79. Besides, the probability of developing prostate cancer in a man during his lifetime is around 17.8% [2]. The risk factors of prostate cancer include familial predisposition, advanced age, race, genetics, diet, environmental factors, and hormonal factors [3]. Because it is seen too much, the correct management of the treatment, diagnosis, and follow-up process of prostate cancer is important for the patient and the doctor, as well as for national health policies [4].

Data mining is an interdisciplinary field that acts as a bridge in many areas such as statistics, data visualization, and pattern recognition, database technology, machine learning, and artificial intelligence [5]. Data mining can be defined as discovering hidden relationships and patterns in data. It focuses on the methods required to define a valid, useful, and new model [6]. In general, data mining models are divided into two as descriptive and predictive. One of its descriptive models is association rules [7].

Data mining methods that analyze the co-occurrence of events are called association rules. Association rules express the co-occurrence of events together with certain possibilities. Association rules are methods that help reveal a seemingly unrelated relationship between data [8]. Association rules analyzes are aimed at finding which variables are “together”. Association rules analysis uses data mining methods to reveal hidden relationships between data [9].

Associative classification is a data mining model and is based on the logic of integrating the association rule model and classification. Associative classification is a type of classification approach that is created with a set of rules obtained by association rule mining to create classification models. The presence of the target/response variable on the right side of the rule obtained in the associative classification algorithm makes it easier for the user to understand and interpret it [10].

In the present study, it is aimed to classify prostate cancer as benign malignant by applying the associative classification model using association rules on open access prostate cancer data set, and to determine the rules related to the diagnosis of prostate cancer.

2. MATERIAL AND METHODS

2.1. Dataset

In the study, in order to examine how the associative classification method works and to evaluate the model, the open-access data set named "Prostate Cancer Data Set" was obtained from the address https://www.kaggle.com/sajidsaifi/prostate-cancer [11]. There are 100 patients diagnosed with prostate cancer in this open access data set. Of the patients diagnosed with cancer, 38 (38%) were diagnosed as benign, and 62 (62%) were diagnosed as malignant. The variables and the descriptive properties of the variables in the relevant data set are given in Table I.

İpek BALIKCI CICEK, Inonu University Department of Biostatistics and Medical Informatics, Faculty of Medicine, Malatya, Turkey, (ipek.balikci@inonu.edu.tr)

Zeynep KUCUKAKCALI, Inonu University Department of Biostatistics and Medical Informatics, Faculty of Medicine, Malatya, Turkey, (zeynep.tunc@inonu.edu.tr)

✉ **Cemil COLAK**, Inonu University Department of Biostatistics and Medical Informatics, Faculty of Medicine, Malatya, Turkey, (cemil.colak@inonu.edu.tr)

Manuscript received Sep 19, 2020; accepted Nov. 5, 2020.
Digital Object Identifier:

TABLE I
VARIABLES IN THE DATA SET AND THEIR DESCRIPTIVE PROPERTIES

Variable	Variable Description	Variable Type	Variable Role
Diagnosis	The diagnosis of breast tissues (M = malignant, B = benign)	Qualitative	Dependent/ Target
Radius	Mean distances from the center to perimeter points	Quantitative	Independent / Predictor
Texture	The standard deviation of gray-scale values	Quantitative	Independent / Predictor
Perimeter	Mean size of the core tumor	Quantitative	Independent / Predictor
Area	-	Quantitative	Independent / Predictor
Smoothness	Mean of local variation in radius lengths	Quantitative	Independent / Predictor
Compactness	(mean of perimeter)2 / (area - 1)	Quantitative	Independent / Predictor
Symmetry	-	Quantitative	Independent / Predictor
Fractal dimension	mean for "coastline approximation" - 1	Quantitative	Independent / Predictor

3. ASSOCIATIVE CLASSIFICATION MODEL

Association rules are an unsupervised data mining method that searches for relationships between data records. Association rules are used to detect useful, consistent, and interesting relationships from large data sets that cannot be deduced at first glance. Association rules, in other words, are useful for detecting frequently seen situations in the data set. The main purpose of association rule algorithms is to determine the probability of occurrence of two or more events together [12]. Relationships between data in association rules analysis are shown with IF-THEN expressions. "IF <if certain conditions are met>" is in the form "THEN <estimate the values of some attributes>" and are rules that are measures of support and trust. Association rules are an approach that supports future studies by analyzing past data and determining association behaviors in these data. One of the data mining models, association rules, is frequently preferred in terms of applicability in almost every field [13].

Associative classification is a supervised learning model that classifies based on the use of association rules. Association rules are derived using "if-then" clauses called precursor-successor. In associative classification models, the right side (successor) of association rules consists only of the categories of the class/response / dependent variable. Thus, the relevant associative classification model classifies the unlabeled test data set by using these association rules [14].

3.1. Performance evaluation metrics

While comparing the classification performances, performance metrics such as sensitivity, specificity, accuracy, balanced accuracy, negative predictive value, positive predictive value, and F1-score were used. The classification matrix of performance metrics is given in Table II.

TABLE II
CLASSIFICATION MATRIX OF PERFORMANCE METRICS

		Real		
		Positive	**Negative**	**Total**
Predicted	Positive	True positive (TP)	False negative (FN)	TP+FN
	Negative	False positive (FP)	True negative (TN)	FP+TN
	Total	TP+FP	FN+TN	TP+TN+FP+FN

Sensitivity = TP/(TP+FP)

Specificity = TN/(TN+FN)

Accuracy = (TP+TN)/(TP+TN+FP+FN)

Balanced accuracy = [[TP/(TP+FP))]+[TN/(TN+FN)]]/2

Negative predictive value =TN/(TN+FP)

Positive predictive value = TP/(TP+FN)

F1-score = (2*TP)/(2*TP+FP+FN)

4. DATA ANALYSIS

Quantitative data were expressed as mean ± standard deviation, median (minimum-maximum), and qualitative data as number (percentage). Conformity to normal distribution was assessed using the Shapiro-Wilk test. Whether there is a statistically significant difference between the "Bening" and "Malignant" groups, which are the categories of dependent/target variable (prostate cancer) in terms of independent variables, was examined using the Mann-Whitney U test and the independent-samples t-test. Values of $p < 0.05$ were considered statistically significant. IBM SPSS Statistics 26.0 package program was used for all analyzes.

5. RESULTS

Descriptive statistics for the independent variables examined in this study are given in Table III. There is a statistically significant difference between the dependent/target variable groups in terms of the perimeter, area, compactness, symmetry, smoothness variables ($p<0.05$).

TABLE III
DISTRIBUTION TABLE OF THE DEPENDENT/TARGET VARIABLE

Benign		Malignant	
Count	Percentage	Count	Percentage
38	38	62	62

The distribution table for the dependent/target variable in the data set is given in Table III.

TABLE IV
DESCRIPTIVE STATISTICS FOR QUANTITATIVE INDEPENDENT VARIABLES

Variables	Diagnosis				p-value
	Benign		Malignant		
	Median (min-max)	**Mean± Standard deviation**	**Median (min-max)**	**Mean± Standard dev.**	
radius	18 (9-25)	-	16 (9-25)	-	0.09*
texture	17 (11-27)	-	18 (11-27)	-	0.450*
perimeter	78.5 (52-133)	-	104 (72-172)	-	<0.001*
area	458.5 (202-1326)	-	790.5 (371-1878)	-	<0.001*
compactnes	0.0785 (0.038-0.246)	-	0.1405 (0.051-0.345)	-	<0.001*
symmetry	0.182 (0.135-0.274)	-	0.193 (0.153-0.304)	-	0.013*
fractal_ dimension	0.0635 (0.053-0.09)	-	0.063 (0.053-0.097)	-	0.963*
smoothness	-	0.099± 0.015	-	0.105± 0.014	0.049**

***: Mann Whitney U test, **: Independent samples t-test**

The distribution table for the dependent/target variable in the data set is given in Table IV.

In this study, the classification matrix of the associative classification model used in classifying the prostate cancer dataset is given in Table V.

TABLE V
CLASSIFICATION MATRIX FOR THE ASSOCIATIVE CLASSIFICATION MODEL

Prediction	**Reference**		
	Bening	**Malingnant**	**Total**
Bening	30	2	32
Malingnant	8	60	68
Total	38	62	100

The values of the classification performance metrics for the associative classification model are shown in Table 6. The sensitivity obtained from the model was calculated as 0.968, selectivity 0.789, accuracy 0.9, balanced accuracy 0.879, negative predictive value 0.938, positive predictive value 0.882, and F1-score 0.923.

Table 7 shows the association rules used by the classification algorithm. As expressed in Table 7, when perimeter=[52, 87.5) and compactness=[0.038, 0.097) are considered, the probability of a male not having prostate cancer is 96.6%. Similarly, as texture=[17.5, 27) and perimeter=[87.5, 172) are taken into account, the probability of a man getting prostate cancer is 96.3%, and when area=[575, 1.88e+03) and smoothness=[0.0905, 0.143) are regarded, the probability of a man with prostate cancer is 95.7%. If area=[575, 1.88e+03) and compactness=[0.097,0.345) are considered, the probability of a man with prostate cancer is 95.7 %.

TABLE VI
VALUES OF THE CLASSIFICATION PERFORMANCE METRICS OF THE MODEL

Metric	**Value**
Sensitivity	0.968
Specificity	0.789
Accuracy	0.9
Balanced accuracy	0.879
Negative predictive value	0.938
Positive predictive value	0.882
F1-score	0.923

The other rules generated from the classification based on the association rules model can be interpreted as the rules described earlier (Table VII).

TABLE VII
ASSOCIATION RULES USED BY THE CLASSIFICATION ALGORITHM

Left-hand side rules	**Right-hand side rules**	**Sup.**	**Conf.**	**Freq.**
{perimeter=[52, 87.5), compactness=[0.038,0.097)}	{diagnosis=Benign}	0.28	0.966	28
{texture=[17.5,27), perimeter=[87.5,172)}	{diagnosis=Malignant}	0.26	0.963	26
{area=[575,1.88e+03), smoothness=[0.0905,0.143)}	{diagnosis=Malignant}	0.45	0.957	45
{area=[575,1.88e+03), compactness=[0.097,0.345)}	{diagnosis=Malignant}	0.44	0.957	44
{texture=[11,16.5), perimeter=[87.5,172), compactness=[0.097,0.345)}	{diagnosis=Malignant}	0.2	0.952	20
{perimeter=[87.5,172), smoothness=[0.0905,0.143)}	{diagnosis=Malignant}	0.47	0.94	47
{area=[575,1.88e+03)}	{diagnosis=Malignant}	0.5	0.926	50
{texture=[11,16.5), compactness=[0.097,0.345)}	{diagnosis=Malignant}	0.25	0.926	25
{radius=[9,24.5), texture=[17.5,27), compactness=[0.097,0.345)}	{diagnosis=Malignant}	0.25	0.893	25

6. DISCUSSION

Prostate cancer is a disease that can start anywhere in the prostate gland, slow in the first 5-10 years, then grow rapidly and can spread to other organs [15]. Prostate cancer is an important cause of illness and death among men. Prostate cancer-initiating causes of prostate cancer are not fully known yet, factors such as genetic factors, chronic inflammation, and infection, high-fat diet, smoking, alcohol use, obesity [16]. The global burden of prostate cancer is among the top five cancers in terms of both incidence and mortality. Therefore, it is possible to prevent the progression of the disease and to apply alternative treatment protocols by diagnosing prostate cancer in the early stages [1].

Data Mining is the analysis and summarization of large amounts of data by various methods in order to detect unexpected or unpredictable relationships and obtain understandable and useful information [17]. One of the most important areas of achieving attractive results in data mining studies is association rules mining. Association rules are mostly used in explanatory data analysis, data preprocessing, determining discrete values, finding trends, and relationships. The relationships between the obtained rules and object or object groups can be determined, and they can be used as a guide in decisions to be made and definitions to be made [18]. Associative classification is a data mining model that classifies by combining association rules and classification methods. The associative classification model gives high accuracy results compared to traditional methods. The main advantage of associative classifiers is that they are easy to interpret. Other models found in data mining complicate the interpretability of results due to a large number of unnecessary and contradictory rules. Associative classification, on the other hand, provides great success in interpreting and classifying results by trimming unnecessary rules [10].

In this study, an associative classification model, which combines the association rule and classification, which is one of the data mining methods, was applied to the data set named "Prostate Cancer", which is an open-access data set [11]. In this context, different factors (independent variables) that may be associated with a prostate cancer diagnosis (the dependent variable) are estimated with the associative classification model, and association rules have been found. According to the results of the associative classification analysis, the performance metrics obtained from the model were obtained as sensitivity 0.968, selectivity 0.789, accuracy 0.9, balanced accuracy 0.879, negative predictive value 0.938, positive predictive value 0.882, and F1-score 0.923.

In a study using the same data set, the accuracy results obtained with different machine learning methods were compared. According to the results of the current study, the highest accuracy was obtained as 0.80 with the k-Nearest Neighbor and Naive Bayes Classification models. In this study, an accuracy of 0.9 was obtained, and the rules related to the disease were also obtained [19].

As a result, the associative classification model proposed as a result of the analysis of the open-access data set produces distinctively successful predictions in classifying prostate cancer according to performance metrics.

REFERENCES

[1] F. Bray, J. Ferlay, I. Soerjomataram, R. L. Siegel, L. A. Torre, and A. Jemal, "Global cancer statistics 2018: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries," CA: a cancer journal for clinicians, vol. 68, pp. 394-424, 2018.

[2] A. Jemal, R. C. Tiwari, T. Murray, A. Ghafoor, A. Samuels, E. Ward, et al., "Cancer statistics, 2004," CA: a cancer journal for clinicians, vol. 54, pp. 8-29, 2004.

[3] A. Jemal, A. Thomas, T. Murray, and M. Thun, "Cancer statistics, 2002," Ca-A Cancer Journal for Clinicians, vol. 52, pp. 23-47, 2002.

[4] R. L. Siegel, K. D. Miller, and A. Jemal, "Cancer statistics, 2016," CA: a cancer journal for clinicians, vol. 66, pp. 7-30, 2016.

[5] S. Özekes, "Veri madenciliği modelleri ve uygulama alanları," 2003.

[6] A. K. Pujari, Data mining techniques: Universities press, 2001.

[7] H. Akpınar, "Veri tabanlarında bilgi keşfi ve veri madenciliği," İstanbul Üniversitesi İşletme Fakültesi Dergisi, vol. 29, pp. 1-22, 2000.

[8] L. Boney, A. Tewfik, and K. Hamdy, "Minimum association rule in large database," in Proceedings of Third IEEE International Conference on Computing, 2006, pp. 12-16.

[9] S. Vinodh, N. H. Prakash, and K. E. Selvan, "Evaluation of leanness using fuzzy association rules mining," The International Journal of Advanced Manufacturing Technology, vol. 57, pp. 343-352, 2011.

[10] M. Azmi, G. C. Runger, and A. Berrado, "Interpretable regularized class association rules algorithm for classification in a categorical data space," Information Sciences, vol. 483, pp. 313-331, 2019.

[11] Available: https://www.kaggle.com/sajidsaifi/prostate-cancer

[12] Y. Köse, "Değerli müşterilerde ürün kategorileri arasındaki satış ilişkilerinin veri madenciliği yöntemlerinden birliktelik kuralları ve kümeleme analizi ile belirlenmesi ve ulusal bir perakendecide örnek uygulama," Selçuk Üniversitesi Sosyal Bilimler Enstitüsü, 2015.

[13] A. S. Albayrak and S. K. Yilmaz, "Veri Madenciliği: Karar Ağacı Algoritmaları ve İMKB Verileri Üzerine Bir Uygulama," Suleyman Demirel University Journal of Faculty of Economics & Administrative Sciences, vol. 14, 2009.

[14] F. A. Thabtah, "A review of associative classification mining," Knowledge Engineering Review, vol. 22, pp. 37-65, 2007.

[15] T. Kahvecioğlu and F. E. Güneş, "İnek sütü ve prostat kanseri ilişkisi," Sağlık ve Yaşam Bilimleri Dergisi, vol. 1, pp. 44-49, 2019.

[16] A. Ahlbom, P. Lichtenstein, H. Malmström, M. Feychting, N. L. Pedersen, and K. Hemminki, "Cancer in twins: genetic and nongenetic familial risk factors," Journal of the National Cancer Institute, vol. 89, pp. 287-293, 1997.

[17] U. M. Fayyad, G. Piatetsky-Shapiro, P. Smyth, and R. Uthurusamy, "Advances in knowledge discovery and data mining," 1996.

[18] A. S. Kumar and R. Wahidabanu, "Data Mining Association Rules for Making Knowledgeable Decisions," in Data Mining Applications for Empowering Knowledge Societies, ed: IGI Global, 2009, pp. 43-53.

[19] Available: https://www.kaggle.com/alihantabak/prostate-cancer-predictions-with-ml-and-dl-methods.

.

BIOGRAPHIES

İpek BALIKÇI ÇİÇEK obtained her BSc. degree in mathematics from Çukurova University in 2010. She received MSc. degree in biostatistics and medical informatics from the Inonu University in 2018. She currently continues Ph.D. degrees in biostatistics and medical informatics from the Inonu University. In 2014, she joined the Department of Biostatistics and Medical Informatics at Inonu University as a researcher assistant. Her research interests are cognitive systems, data mining, machine learning, deep learning.

Zeynep KÜÇÜKAKÇALI obtained her BSc. degree in mathematics from Çukurova University in 2010. She received MSc. degree in biostatistics and medical informatics from the Inonu University in 2018. She currently continues Ph.D. degrees in biostatistics and medical informatics from the Inonu University. In 2014, she joined the Department of Biostatistics and Medical Informatics at Inonu University as a researcher assistant. Her research interests are cognitive systems, data mining, machine learning, deep learning.

Cemil ÇOLAK obtained his BSc. degree in statisticsfrom Ondokuz Mayıs University in 1999. He received MSc. degree in Biostatistics from the Inonu University in 2001, and Ph.D. degree in the Graduate Department of Biostatistics and Medical Informatics of Ankara University in 2005. His research interests are cognitive systems, data mining, reliability, and biomedical system, genetics, and bioengineering. In 2016, he joined the Department of Biostatistics and Medical Informatics at Inonu University as a Professor, where he is presently a professor. He is active in teaching and research in the general image processing, artificial intelligence, data mining, analysis.

PERFORMANCE EVALUATION OF THE ENSEMBLE LEARNING MODELS IN THE CLASSIFICATION OF CHRONIC KIDNEY FAILURE

Z. Kucukakcali, and I. Balikci Cicek

Abstract— *Aim:* This study aims to classify the CKF by applying the ensemble learning method, which is an important sub-field of machine learning, on the open access CKF data set.

Materials and Methods: In this study, the ensemble learning methods Bagging, Boosting and Stacking methods were applied to the open access data set named “Chronic Kidney Disease”. The performance of the models used was evaluated with accuracy, sensitivity, specificity, positive predictive value, and negative predictive value.

Results: Accuracy, , sensitivity, specificity, positive predictive value and negative predictive value obtained from the Bagging model were 96.5, 96.8, 96, 97.5 and 94.7 respectively. Accuracy, , sensitivity, specificity, positive predictive value and negative predictive value obtained from the Boosting model were 98.75, 98, 1, 1 and 96.7 respectively. Accuracy, , sensitivity, specificity, positive predictive value and negative predictive value obtained from the Stacking model were 99.25, 99.6, 98.9, 99.2 and 99.3 respectively.

Conclusion: The findings obtained from this study showed that successful results were obtained with the ensemble learning model for the kidney failure data set.

Keywords— Chronic kidney failure, classification, machine learning, ensemble learning.

1. INTRODUCTION

CHRONIC kidney failure (CKF), which has become an important public health problem in the world and in our country today, is a disease that can occur due to many reasons, results in irreversible loss of kidney functions, negatively affects the quality of life of individuals and requires lifelong treatment and follow-up [1]. Clinically, CKF is defined as a structural and functional disorder that can be demonstrated by blood, urine, and imaging methods resulting in a decrease in nephron count and nephron functions as a result of a decrease in glomerular filtration rate (GFR) that lasts for more than three months. [2]. Nowadays, it is reported that the incidence of CKF is increasing rapidly. Considering the results of the studies, it is seen that the rate of chronic kidney failure varies between 10-16% in the world.

It is estimated that more than 500 million people worldwide have CKF. [3]. Chronic kidney failure (CKF) is an increasingly common health problem worldwide. This disease, which is very expensive to treat, can lead to negative consequences when evaluated from a prognostic point of view. The most important consequences are the progression of kidney disease, acute and chronic complications due to renal dysfunction, cardiovascular mortality, and morbidity. Therefore, CKF, which can occur due to many reasons and results in irreversible loss of kidney functions, is a disease that negatively affects the quality of life of individuals and requires lifelong treatment and follow-up [4].

Data mining is the finding of relationships and rules that will enable us to make predictions about the future from a large amount of data with the help of computer programs. [5]. Data mining; includes a combination of techniques from different disciplines such as database technology, statistics, machine learning, pattern recognition, neural networks, data visualization, and spatial data analysis [6]. Machine learning, one of these techniques, is a subfield of artificial intelligence that aims to make predictions about new data when they are exposed to new data by performing data-based learning. Machine learning systems aim to completely eliminate the need for human intuition or to gain the ability to make decisions through human-machine cooperation [7]. The logic of ensemble learning, which is an important subfield of machine learning, is based on the idea that many classifiers can be combined to increase the rate of correct prediction using a single basic classifier. In other words, the method of ensemble learning is based on the idea of combining many basic classifiers to obtain a more accurate and reliable model (meta classifier) compared to the classification success that a basic classifier (model) can achieve [8]. In this way, ensemble learning methods increase the predictive power of weak classifiers [9]. For these reasons, ensemble learning methods are highly preferred recently.

The purpose of this study is to classify the CKF by applying the ensemble learning method, which is an important sub-field of machine learning, on the open-access CKF data set.

2. MATERIAL AND METHODS

2.1. Dataset

In the study, the ensemble learning method, which is an important sub-field of machine learning, was applied to an open-access data set called “Chronic Kidney Disease”. Open access data set named “Chronic Kidney Disease” was obtained from https://www.kaggle.com/abhia1999/chronic-kidney-disease. There are 400 patients in the data set used. 250 (62.5%) of these patients have chronic kidney failure. Explanations

✉ **Zeynep KUCUKAKCALI,** Inonu University Department of Biostatistics and Medical Informatics, Faculty of Medicine, Malatya, Turkey, (zeynep.tunc@inonu.edu.tr)

İpek BALIKCI CICEK, Inonu University Department of Biostatistics and Medical Informatics, Faculty of Medicine, Malatya, Turkey, (ipek.balikci@inonu.edu.tr)

Manuscript received Sep 19, 2020; accepted Oct 21, 2020.
Digital Object Identifier:

about the variables and their properties in the data set are given in Table 1.

TABLE I
EXPLANATIONS ABOUT THE VARIABLES IN THE DATASET AND THEIR PROPERTIES

Variable	**Variable Description**	**Variable Type**	**Variable Role**
Bp	Blood Pressure	Quantitative	Predictor
Sg	Specific Gravity	Quantitative	Predictor
Al	Albumin	Qualitative	Predictor
Su	Sugar	Qualitative	Predictor
Rbc	Red Blood Cell	Qualitative	Predictor
Bu	Blood Urea	Quantitative	Predictor
Sc	Serum Creatinine	Quantitative	Predictor
Sod	Sodium	Quantitative	Predictor
Pot	Pottasium	Quantitative	Predictor
Hemo	Hemoglobin	Quantitative	Predictor
Wbcc	White Blood Cell Count	Quantitative	Predictor
Rbcc	Red Blood Cell Count	Quantitative	Predictor
Htn	Hypertension	Qualitative	Predictor
Class	Predicted Class	Qualitative	Output

3. ENSEMBLE LEARNING METHOD

Machine learning methods classify and infer by learning the pattern in the data stack. Machine learning has developed rapidly in recent years. The rapid development of machine learning has been dependent on the development of statistical algorithms that can extract information from these data with the rapid increase in data stacks in the computer and internet environment. For this reason, many machine learning methods have been developed, some of which are k-nearest neighbor algorithm, simple (naive) Bayes classifier, decision trees, logistic regression analysis, k-means algorithm, support vector machines, and artificial neural networks. Some of these approaches have the ability to predict, some to cluster and some to classify [10]. An important subfield of machine learning is ensemble learning methods. Ensemble learning methods provide a common classification result from the estimates of each classifier by classifying the data of more than one machine learning algorithm separately, rather than classifying the data set of a machine learning algorithm. Thus, according to the prediction results of a machine learning method, the common prediction results obtained from more than one machine learning method provide more accurate, more reliable, and higher performance [7]. Ensemble learning methods are based on the principle that more than one classifier can perform classification with higher accuracy than a single classifier predicts. Ensemble learning methods have found a wide range of applications in recent years with their successful results. Commonly used ensemble learning methods have been successfully applied in many diagnostic and diagnostic studies [11, 12].

An important issue affecting the classification performance in the ensemble learning method is the selection of the appropriate joining method. In determining the joining technique, attention should be paid to the selection of the appropriate joining technique for classifiers. There are different ensemble learning methods according to the joining techniques, the sample selection for the training data set, and the process steps. These methods are the bagging ensemble learning method, the boosting ensemble learning method, and the stacking ensemble learning method [13, 14].

3.1. Bagging

Bagging method is a method that aims to retrain the basic learner by creating new training sets by random selection by substituting from a known training set [15]. In summary, the main purpose of the Bagging method is to obtain new data sets randomly using training data and to increase the success of classification by creating differences. In the bagging method, first, the data set is divided into training and test data. One or more new training sets consisting of n samples are obtained by random selection method by replacing the training set containing N samples. Each basic classifier in the community obtained by the bagging method is trained with training sets containing different examples obtained in this way. Finally, the result of each major classifier is combined with the majority vote [16].

3.2.Boosting

Boosting yöntemi, Schapire tarafından 1990 yılında tanıtılan ve 2000'li yıllara kadar geliştirilen bir toplu öğrenme yöntemi [17-19]. The term "boost" refers to a family of algorithms that transform poor learning methods into powerful learning techniques. Boosting is an ensemble method to improve the model predictions of any learning algorithm, and unlike the Bagging method, the predictors in the Boosting method are created sequentially, although they are not independent of each other. The aim of this method is to combine weak estimators to obtain strong estimator (s). Models are created by assigning weight to observations. In the Boosting method, as in the bagging method, N training sets are created. In this method, models with low variance and bias are obtained by both the presence of the bagging method and the assignment of weight to the observations [19].

3.3.Stacking

The stacking method is a simple ensemble learning technique that creates a meta classifier by combining two or more basic multiple classification models. It is an ensemble model that is trained by combining the estimates of the classification models used. Predictions made from models created by the basic classifier are used as input for each ordered layer and are combined to create a new set of predictions. In the stacking method, basic classification models are trained on the original training data set and then created based on the outputs (estimates) of the basic classification models in the meta-classifier community. The meta-classifier performs the classification process by training on the predicted class labels [20].

3.4. Performance evaluation criteria

The classification matrix for the calculation of performance metrics is given in Table II.

TABLE II

THE METRICS OF THE MODEL'S CLASSIFICATION PERFORMANCE

		Real		
		Positive	**Negative**	**Total**
Predicted	Positive	True positive (TP)	False negative (FN)	TP+FN
	Negative	False positive (FP)	True negative (TN)	FP+TN
	Total	TP+FP	FN+TN	TP+TN+FP+FN

Accuracy = (TP+TN)/(TP+TN+FP+FN)

Sensitivity = TP/(TP+FP)

Specificity = TN/(TN+FN)

Positive predictive value = TP/(TP+FN)

Negative predictive value =TN/(TN+FP)

4. Data Analysis

Quantitative data are summarized by median (minimum-maximum) and qualitative variables are given by number and percentage. Normal distribution was evaluated with the Kolmogorov-Smirnov test. In terms of input variables, the existence of a statistically significant difference and relationship between the categories of output variable, "ckd" and "notckd" groups, was examined using Mann-Whitney U, Pearson Chi-square test and Yates's correction chi-square test. p<0.05 values were considered statistically significant. In all analyzes, IBM SPSS Statistics 26.0 for the Windows package program was used. Random Forest, Neural Network, Support Vector Machine were used as classifiers in the stacking method of the ensemble learning models. Rapidminer Studio Free version 9.4 is used for ensemble learning methods analysis.

5. Results

Descriptive statistics related to the target variable examined in this study are presented in Table 3 and Table 4. There is a statistically significant difference between the dependent variable classes in terms of other variables other than the “Pot” variable.

TABLE III

DESCRIPTIVE STATISTICS FOR QUANTITATIVE INPUT VARIABLES

Variables	Predicted Class Not-ckd **Median (min-max)**	 ckd **Median (min-max)**	**P* value**
Bp	70(60-80)	80(50-180)	<0,001*****
Sg	1,02(1,02-1,03)	1,02(1,01-1,03)	<0,001*****
Bu	33,5(10-57)	55(1,5-391)	<0,001*****
Sc	0,9(0,4-3,07)	2,45(0,5-76)	<0,001*****
Sod	141(135-150)	137,53(4,5-163)	<0,001*****
Pot	4,5(3,3-5)	4,63(2,5-47)	0,515
Hemo	15(12,53-17,8)	11,3(3,1-16,1)	<0,001*****
Wbcc	7750(4300-11000)	8406(2200-26400)	<0,001*****
Rbcc	5,25(4,4-6,5)	4,71(2,1-8)	<0,001*****

***: Mann Whitney U test**

TABLE IV

DESCRIPTIVE STATISTICS FOR QUALITATIVE INPUT VARIABLES

Variables		Predicted Class Not-ckd	 ckd	**P** value**
Al	0	145(96,7%)	54(21,6%)	<0,001*****
	1	5(3,3%)	85(34,0%)	
	2	0(0%)	43(17,2%)	
	3	0(0%)	43(17,2%)	
	4	0(0%)	24(9,6%)	
	5	0(0%)	1(0,4%)	
Su	0	150(100%)	189(75,6%)	<0,001*****
	1	0(0%)	13(5,2%)	
	2	0(0%)	18(7,2%)	
	3	0(0%)	14(5,6%)	
	4	0(0%)	13(5,2%)	
	5	0(0%)	3(1,2%)	
rbc	0	0(0%)	47(18,8%)	<0,001******
	1	150(100%)	203(81,2%)	
htn	0	150(100%)	103(41,2%)	<0,001******
	1	0(0%)	147(58,8%)	

***: Pearson's chi-square test;** Yates's correction chi-square test**

In this study, the classification matrices of Bagging, Boosting, and Stacking models, which are among the ensemble learning methods used to classify the CKF dataset, are given in Table V.

TABLE V
CLASSIFICATION MATRICES OF BAGGING, BOOSTING AND STACKING MODELS

Classification Matrix of the Bagging Model			
Prediction	**Reference**		
	CKD	not CKD	**Total**
CKD	242	6	248
not CKD	8	144	152
Total	250	150	400

Classification Matrix of the Boosting Model			
Prediction	**Reference**		
	CKD	not CKD	**Total**
CKD	245	0	245
not CKD	5	150	155
Total	250	150	400

Classification Matrix of the Stacking Model			
Prediction	**Reference**		
	CKD	not CKD	**Total**
CKD	249	2	251
not CKD	1	148	149
Total	250	150	400

The values for the metrics of the classification performance of Bagging, Boosting and Stacking models are given in Table VI.

TABLE VI
VALUES FOR THE METRICS OF THE CLASSIFICATION PERFORMANCE OF BAGGING, BOOSTING AND STACKING MODELS

Models	**Metric**	**Value**
Bagging	Accuracy	96.5
	Sensitivity	96.8
	Specitivity	96
	Positive predictive value	97.5
	Negative predictive value	94.7
Boosting	Accuracy	98.75
	Sensitivity	98
	Specitivity	1
	Positive predictive value	1
	Negative predictive value	96.7
Stacking	Accuracy	99.25
	Sensitivity	99.6
	Specitivity	98.9
	Positive predictive value	99.2
	Negative predictive value	99.3

Accuracy, sensitivity, specificity, positive predictive value and negative predictive value obtained from the Bagging model were 96.5, 96.8, 96, 97.5 and 94.7 respectively. Accuracy, sensitivity, specificity, positive predictive value and negative predictive value obtained from the Boosting model were 98.75, 98, 1, 1 and 96.7 respectively. Accuracy, sensitivity, specificity, positive predictive value and negative predictive value obtained from the Stacking model were 99.25, 99.6, 98.9, 99.2 and 99.3 respectively.

6. DISCUSSION

Chronic kidney failure (CKF) is an important public health problem with increasing frequency in the world and in our country. KRG; it is an important health problem that is commonly seen with objective kidney damage lasting at least three months and/or GFR below 60ml / min, chronic and progressive impairment in the fluid-electrolyte balance, endocrine and metabolic functions of the kidney, increased mortality and decreased quality of life. In population-based studies investigating the prevalence of CKF, similar results have been obtained in the world and in our country. The prevalence of CKF in the world was found to be 11.7-15.1% (average 13.4%) according to the results of the meta-analysis study conducted in 2016. CKF is considered to be a serious public health problem in the world due to its high morbidity rate and increased health expenditures. Therefore, it is an open area for research and new developments. [21-23].

Machine learning methods perform classification and estimation by learning the pattern in the data stack. Machine learning has developed rapidly in recent years. Machine learning methods are one of the technologies used in the diagnosis of diseases and clinical decision support systems in recent years. [24]. Ensemble learning methods, one of the machine learning methods, provide a common classification result from the predictions of each classifier by classifying the data of more than one machine learning algorithm rather than classifying the data set of a machine learning algorithm. Thus, according to the prediction results of a machine learning method, common prediction results obtained from more than one machine learning method provide more accurate, more reliable, and higher performance. Ensemble learning methods have found a wide range of applications in recent years with their successful results. [11, 25].

In this study, ensemble learning models, one of the machine learning methods, were applied to the data set named “Chronic Kidney Disease” which is an open-source data set. According to the results of 3 different models used, the method for the best classification performance is the Stacking method. Accuracy, sensitivity, specificity, positive predictive value and negative predictive value obtained from this model were 99.25, 99.6, 98.9, 99.2 and 99.3 respectively.

In a study conducted with the same data set, the effects of variable selection on the support vector machine method were investigated. According to the results of this study, the highest accuracy was obtained as 98.5 from the support vector machine results created with different variable selection methods [26]. In this study, the disease was classified by obtaining an accuracy of 99.25 with the Stacking method.

As a result, the ensemble learning methods used have produced very successful results in the study conducted with the chronic kidney failure data set. Ensemble learning methods offer very high classification performance when correct classifiers and joining techniques are used.

REFERENCES

[1] N. Erol, "Diyaliz tedavisine başlanmayan kronik böbrek yetmezliği hastaları ile hemodiyaliz tedavisi olan hastaların yaşam kalitelerinin karşılaştırılması," Sağlık Bilimleri Enstitüsü, 2010.

[2] E. Crowe, D. Halpin, and P. Stevens, "Early identification and management of chronic kidney disease: summary of NICE guidance," Bmj, vol. 337, pp. a1530, 2008.

[3] M. Eskicioglu, Ü. Eda, and A. Özdemir, "Böbrek hastalarının klinikte yattığı sürede öğrenim gereksinimlerinin tespiti," Uludağ Üniversitesi Tıp Fakültesi Dergisi, vol. 45, pp. 205-210.

[4] Ü. Yıldırım, "HİpervolemİK kronİk böbrek yetmezlİğİ hastalarinda troponin düzeyleri ve medİkal dİürez tedavisinin troponin düzeyi üzerine etkileri," Uzmanlık, Van yyü üniversitesi, 2020.

[5] H. Akpınar, "Veri tabanlarında bilgi keşfi ve veri madenciliği," İÜ İşletme Fakültesi Dergisi, vol. 29, pp. 1-22, 2000.

[6] D. J. Hand, "Principles of data mining," Drug safety, vol. 30, pp. 621-622, 2007.

[7] R. Polikar, "Ensemble learning," in Ensemble machine learning, ed: Springer, 2012, pp. 1-34.

[8] L. Rokach and O. Maimon, "Clustering methods," in Data mining and knowledge discovery handbook, ed: Springer, 2005, pp. 321-352.

[9] T. G. Dietterich, "Ensemble methods in machine learning," in International workshop on multiple classifier systems, 2000, pp. 1-15.

[10] M. Atalay and E. Çelik, "Büyük Veri Analizinde Yapay Zekâ Ve Makine Öğrenmesi Uygulamalari-Artificial Intelligence and Machine Learning Applications in Big Data Analysis," Mehmet Akif Ersoy Üniversitesi Sosyal Bilimler Enstitüsü Dergisi, vol. 9, pp. 155-172, 2017.

[11] S. Sun, "Ensemble learning methods for classifying EEG signals," in International Workshop on Multiple Classifier Systems, 2007, pp. 113-120.

[12] S.-L. Hsieh, S.-H. Hsieh, P.-H. Cheng, C.-H. Chen, K.-P. Hsu, I.-S. Lee, et al., "Design ensemble machine learning model for breast cancer diagnosis," Journal of medical systems, vol. 36, pp. 2841-2847, 2012.

[13] C. Zhang and Y. Ma, Ensemble machine learning: methods and applications: Springer, 2012.

[14] L. Rokach, Pattern classification using ensemble methods vol. 75: World Scientific, 2010.

[15] L. Breiman, "Bagging predictors," Machine learning, vol. 24, pp. 123-140, 1996.

[16] A. J. Ferreira and M. A. Figueiredo, "Boosting algorithms: A review of methods, theory, and applications," in Ensemble machine learning, ed: Springer, 2012, pp. 35-85.

[17] J. H. Friedman, "Greedy function approximation: a gradient boosting machine," Annals of statistics, pp. 1189-1232, 2001.

[18] J. Friedman, T. Hastie, and R. Tibshirani, "Additive logistic regression: a statistical view of boosting (with discussion and a rejoinder by the authors)," The annals of statistics, vol. 28, pp. 337-407, 2000.

[19] Y. Freund and R. E. Schapire, "Experiments with a new boosting algorithm," in icml, 1996, pp. 148-156.

[20] F. Divina, A. Gilson, F. Goméz-Vela, M. García Torres, and J. F. Torres, "Stacking ensemble learning for short-term electricity consumption forecasting," Energies, vol. 11, pp. 949, 2018.

[21] N. R. Hill, S. T. Fatoba, J. L. Oke, J. A. Hirst, C. A. O'Callaghan, D. S. Lasserson, et al., "Global prevalence of chronic kidney disease–a systematic review and meta-analysis," PloS one, vol. 11, pp. e0158765, 2016.

[22] S. F. Yalin, E. Parmaksiz, M. Mese, Z. Dogu, N. D. Çeçen, and Z. B. Bahçebasi, "Evaluating the causes for rejection of potential live-renal donors: Single center experience/potansiyel canli bobrek vericilerinin reddedilme nedenlerinin degerlendirilmesi: Tek merkez deneyimi," Journal of Istanbul Faculty of Medicine, vol. 82, pp. 127-131, 2019.

[23] A. C. Webster, E. V. Nagler, R. L. Morton, and P. Masson, "Chronic kidney disease," The lancet, vol. 389, pp. 1238-1252, 2017.

[24] A. Doganer, "Topluluk öğrenme yöntemleri ile renal hücreli karsnom'un tahmin edilmesi," Doktora, İnönü Üniversitesi, 2020.

[25] Y. Qi, "Random forest for bioinformatics," in Ensemble machine learning, ed: Springer, 2012, pp. 307-323.

[26] H. Polat, H. D. Mehr, and A. Cetin, "Diagnosis of chronic kidney disease based on support vector machine by feature selection methods," Journal of medical systems, vol. 41, pp. 55, 2017.

BIOGRAPHIES

Zeynep KÜÇÜKAKÇALI obtained her BSc. degree in mathematics from Çukurova University in 2010. She received MSc. degree in biostatistics and medical informatics from the Inonu University in 2018. She currently continues Ph.D. degrees in biostatistics and medical informatics from the Inonu University. In 2014, she joined the Department of Biostatistics and Medical Informatics at Inonu University as a researcher assistant. Her research interests are cognitive systems, data mining, machine learning, deep learning.

İpek BALIKÇI ÇİÇEK obtained her BSc. degree in mathematics from Çukurova University in 2010. She received MSc. degree in biostatistics and medical informatics from the Inonu University in 2018. She currently continues Ph.D. degrees in biostatistics and medical informatics from the Inonu University. In 2014, she joined the Department of Biostatistics and Medical Informatics at Inonu University as a researcher assistant. Her research interests are cognitive systems, data mining, machine learning, deep learning.

A PROPOSED MODEL CAN CLASSIFY THE COVID-19 PANDEMIC BASED ON THE LABORATORY TEST RESULTS

S. Yasar and C. Colak

***Abstract*—** As reported by the World Health Organization (WHO) in March 2020, COVID-19 is a worldwide epidemic. Since the rapid spread of the epidemic harms humans, the need for methods that enable early diagnosis and treatment has increased. Machine learning (ML) methods can play a vital role in identifying COVID-19 patients. In this study, the classification algorithms of ML methods (CART), Support Vector Machine (SVM-Radial), K Nearest Neighbors (K-NN) and Random Forest are used to determine the best model that diagnoses COVID-19 from the person's complete blood counts (positive/negative). According to the experimental results, the Random Forest algorithm gives the best predictions in the classification of COVID-19 (99.76% of accuracy). Besides, in the classification of Covid-19, it can be recommended to apply meta-learning algorithms as they can give high predictive results.

***Keywords*—** *COVID-19, Machine Learning, CART, SVM, K-NN, Random Forest*

1. INTRODUCTION

THE COVID-19 outbreak, caused by the Severe Acute Respiratory Syndrome (SARS-CoV-2) virus, which first appeared in Wuhan, China's Hubei province on December 31, 2019, has rapidly spread to hundreds of countries. The epidemic process, which started in our country with the identification of the first positive case on March 11, 2020, continues increasingly. This disease agent has become the most critical health problem of the 21st century due to its high contagious feature, its unfavorable clinical prognosis, and its lethal effect in almost every age group, especially those aged 65 and above. Therefore, since the isolation of the new type of coronavirus, researches on COVID-19 disease and SARS-CoV-2 virus have been started in many countries, and it is possible to provide accurate information about the disease with these studies. However, as of the current situation, there are many issues related to the COVID-19 disease that have not yet been clarified [1].

Although developments in the field of health and technology are pleasing, some uncertainties encountered in medicine make it difficult for clinicians to make decisions. In recent years, artificial intelligence (AI) and machine learning (ML)-based studies have gained importance in terms of supporting the decision of doctors to solve problems such as the diagnosis of diseases in the field of health.

Artificial intelligence is the general name of the technology to create and develop machines that can exhibit human mental activities and behaviors without the support of a living being [2]. Machine learning (ML), a sub-branch of artificial intelligence (AI), is a branch of computer science that makes decisions using past experiences when it is necessary to make decisions for the future. It allows a model to learn automatically from experience based on data, without having to model it just like statistical models. ML creates a new rule from the given samples [3]. It is also relevant to other disciplines such as machine learning, artificial neural networks, pattern recognition, information retrieval, artificial intelligence, data mining, and function approach. Compared to these areas, machine learning focuses more on how to model, optimize, and get organized for the best use of accessible training data. Classification algorithms used when data is tagged in machine learning are also frequently used in this field [4]. The purpose of classification algorithms is to predict which of the pre-labelled data groups similar data belongs to.

Many different algorithms have been developed for the classification process in the literature. Some of these are Decision Trees (CART algorithm), Support Vector Machines (SVM-Radial), K- nearest neighborhood (K-NN) and Random Forest (RF) algorithms. In the decision trees algorithm, the prediction can be obtained by building a decision tree with test points and branches. Each test point of a decision tree involves testing a specific input variable, and each branch represents the decision being made. A node that does not contain more branches is called a leaf node. The depth of the node is the minimum number of steps required to get from the node to the root [5]. In support vector machines algorithm, a technique called kernel trick is used to transform data. Kernel trick methods determine the optimal boundary among possible outcomes based on data transformation models. That is, kernel solution methods first perform complex data transformations and then determine how this data is to be separated based on the defined classes or results [6]. K-Nearest neighbor algorithm is one of the most known and used algorithms in machine learning algorithms. In this algorithm, a selected feature is classified using the distance between its closest feature. The K value found here is expressed as a number such as 3 or 5, for example [7]. The working principle of the Random Forest algorithm creates multiple decision trees and combines them to obtain a more accurate and stable forecast. Each tree is trained with different training sets and uses the results obtained. Each tree produces a classifier. These produced classifiers vote among

Seyma YASAR, Inonu University Department of Biostatistics and Medical Informatics, Faculty of Medicine, Malatya, Turkey, (seyma.yasar@inonu.edu.tr)

✉ **Cemil COLAK,** Inonu University Department of Biostatistics and Medical Informatics, Faculty of Medicine, Malatya, Turkey, (cemil.colak@inonu.edu.tr)

Manuscript received Sep 21, 2020; accepted Nov 16, 2020.
Digital Object Identifier:

themselves, and the algorithm determines the classifier with the most votes. This selected classifier is used to classify the data when new data is given [8].

This study aims to examine the data set consisting of laboratory test results obtained from patients with negative or positive SARS-Cov-2 test results with learning methods [such as decision trees (CART), support vector machines (SVM-Radial), K nearest neighborhood (K-NN) and random forest] to predict the COVID-9. It is foreseen that the best classification method can be used as an auxiliary system to diagnose the COVID-19 disease. Besides, a prediction model that can provide diagnostic support to physicians and other healthcare professionals is proposed to classify the COVID-19 pandemic.

2. MATERIAL AND METHODS

2.1. Dataset

The data set used in this study is the anonymized version of the SARS-CoV-2 RT-PCR and complete blood count (14 different) results of 5644 patients who applied to Israelita Albert Einstein Hospital in São Paulo, Brazil as suspicious. Complete blood count data are standardized to have mean zero and unit standard deviation before sharing. Some variables in the data set were excluded from the analysis because the missing data rate was over 90%. Explanations regarding the variables and their properties used in the creation of algorithms are given in Table 1.

TABLE I
EXPLANATIONS ABOUT THE VARIABLES IN THE DATASET AND THEIR PROPERTIES

Variable	Variable Description	Variable Type	Variable Role
Class	SARS-Cov-2 exam result (negative/positive)	Qualitative	Output
HCT	Hematocrit	Quantitative	Predictor
HGB	Hemoglobin	Quantitative	Predictor
PLT	Platelets	Quantitative	Predictor
MPV	Mean platelet volume	Quantitative	Predictor
RBC	Red blood Cells	Quantitative	Predictor
LYM	Lymphocytes	Quantitative	Predictor
MCHC	Mean corpuscular hemoglobin concentration	Quantitative	Predictor
LEU	Leukocytes	Quantitative	Predictor
BASO	Basophils	Quantitative	Predictor
MCH	Mean corpuscular hemoglobin	Quantitative	Predictor

The class variable included in the study consists of a total of 601 samples, 83 positive samples and 518 negative samples. Since it is thought that this class imbalance problem will negatively affect the training and test performances of the algorithms, the R programming language has been balanced using the "unbalanced" package.

2.2. Decision Trees

Decision trees technique are one of the simplest forms of the decision model, and they use sample data properties to create their rules, which are in the form of a tree structure. Classification of data using the decision tree technique is a two-step process, which is learning and classification. In the learning stage, previously known training data is analyzed by the classification algorithm in order to create a model. The learned model is shown as classification rules or decision trees. In the classification step, test data are used to determine the accuracy of the classification rules or decision tree. If the accuracy is acceptable, the rules are used to classify new data. It should be determined in which order which fields in the training data will be used to create the tree. The most widely used measurement for this purpose is the Entropy measurement. The greater the measure of entropy, the more uncertain and unstable the results obtained using that field. Therefore, fields with the least entropy measure are used in the root of the decision tree [9]. There are many decision tree algorithms. The Classification and Regression Tree (CART) used in this study is a binary decision tree created by dividing a variable into two consecutive nodes (by repeating the process until the homogeneity criterion is reached), starting from the root node containing two learning steps [10]. The essential step in the creation of the CART decision tree is that branching to classify variables in the data set to determining according to which the criteria or which variable to do. At this stage, the variable with the lowest uncertainty is processed and used for testing at the root node [11].

2.3. Support Vector Machines (SVM)

A support vector machine creates an n-dimensional hyperplane that optimally divides the data into two categories. SVM models are closely related to artificial neural networks, SVM using a sigmoid kernel function; It has a two-layer feed-forward neural network. SVM uses kernel functions with high dimensional properties to ensure high performance when performing classification. In this study, a radial based kernel function is used as kernel function [12].

2.4. K Nearest Neighbors Algorithm (K-NN)

The nearest neighbors algorithm is a robust and versatile classifier often used as a reference point for more complex classifiers such as SVMs. In the K-NN algorithm, the samples in the training set are specified with n-dimensional numerical attributes. All training samples are kept in an n-dimensional sample space, with each sample representing a point in n-dimensional space. When an unknown sample is encountered, k nearest samples are determined from the training set, and the class label of the new sample is assigned according to the majority vote of the class labels of the k nearest neighbors [13].

2.5. Random Forest

Developed by Breiman, the Random Forest classifier creates a community with many decision trees it creates [14]. With the Bootstrap sampling method, different subsets are created from the data set and each decision tree trains the decision tree in the community. The best of the randomly determined variables for each node is considered, and the nodes are branched [15]. In the Random Forest algorithm, classification is started by determining two main parameters by the user. One of these parameters is the number of decision trees planned to be created, and the other parameter is the number of variables that perform the split. The Random Forest classifier uses classification and regression trees (CART) method to generate trees. In the classification process, the result of the estimation of each tree is taken, and the majority decides what the classifier will be of votes method [14]. Random Forest classifiers are fast working classifiers as well as being resistant to overfitting problem. Two-thirds of the data set is divided into training data set and the remaining part as test data. Using a large number of decision trees as classifiers and generating a general estimate with the majority vote from the estimates obtained from each classifier decreases the bias while reducing the error rate in the estimates [15].

3. RESULTS

Performance criteria for making a COVID-19 diagnosis (negative/positive) using decision trees (CART), support vector machines (SVM), K- nearest neighborhood (K-NN) and random forest algorithms used in the study and 95% confidence intervals for these criteria are given in Table II.

TABLE II

PERFORMANCE METRICS AND A 95% CONFIDENCE INTERVAL FOR THE DATASET

Classification Algorithm	**Accuracy (%95 CI)**	**Sensitivity (%95 CI)**	**Specificity (%95 CI)**	**MCC (%95 CI)**	**F1-Score (%95 CI)**
Decision Trees (CART)	81.20 (77.4-84.9)	80.95 (77.17-84.73)	81.46 (77.7-85.2)	62.41 (57.7-67.1)	81.34 (77.6-85.1)
SVM (Radial)	92.77 (90.2-95.2)	96.19 (94.35-98.03)	89.27 (86.2-92.2)	85.72 (82.3-89.1)	93.09 (90.6-95.5)
K-NN	90.60 (87.7-93.4)	100.00 (100-100)	80.98 (77.2-84.7)	82.64 (79-86.3)	91.50 (88.8-94.2)
Random Forest	99.76 (99.2-99.9)	100.00 (100-100)	99.51 (98.8-99.9)	99.52 (98.9-99.9)	99.76 (99.3-99.9)

MCC: Matthews correlation coefficient

TABLE III

THE VARIABLE IMPORTANCE SCORE FOR COVID-19 ACCORDING TO THE RANDOM FOREST

Variables	**Importance Score**
LEU	157.64
PLT	125.95
EOS	87.99
MONO	75.56
RBC	46.09
LYM	44.66
MPV	43.73
HCT	40.17
BASO	39.30
HGB	37.29
MCV	33.10
MCHC	32.75
RDW	32.00
MCH	31.76

Accuracy performance values for CART, SVM, K-NN and Random Forest algorithms used in the classification of COVID-19 disease were found as 81.20, 92.77, 90.60 and 99.76, respectively. Therefore, the algorithm with the best accuracy according to the performance criteria obtained from the algorithms used in classifying COVID-19 disease is the Random Forest algorithm. The importance scores of the complete blood count values for the COVID-19 classification according to the random forest algorithm are given in Table III.

4. CONCLUSION

The studies and reports published by the World Health Organization (WHO) on the new coronavirus, which has spread to the world in a short time since its emergence in Wuhan, China, are followed with interest and concern by the whole world [16]. Therefore, considering the harmful effects of the pandemic on humans, it is crucial to detect COVID-19 positive cases at early stages and to make an immediate and correct intervention. The symptoms of COVID-19 are similar to those of the common cold and flu, making early diagnosis difficult for clinicians.

In this study, different classification algorithms (CART, SVM, K-NN, Random Forest) that classify COVID-19 by using the complete blood count results of COVID-19 positive patients were created. When the experimental results are examined, the accuracy values for each model in the COVID-19 (positive/negative) classification from the person's complete blood count are higher than 81%. Therefore, the classification performance of each model created is remarkable. Finally, when the COVID-19 classification performances of the algorithms used are examined, the random forest algorithm has the best performance.

On the other hand, medical imaging techniques are alternative to the methods used to diagnose COVID-19. Therefore, models that allow early diagnosis and diagnosis with different machine learning algorithms have been developed over images obtained from imaging techniques such as Computed Tomography (CT) and Chest radiography (CXR).

In a study have been used different deep learning models (ResNet18, ResNet34, InceptionV3, InceptionResNetV2, DenseNet161, and their Ensemble) to classify COVID-19 using Chest X-Ray images from infected/non-infected subjects. Among the performance metrics obtained in classifying the COVID-19 of the created models, the average F1-Score varies between 0.66 and 0.875. On the other hand, the average F1-Score value of the model created by ensembling all models was found to be 0.89 [17].

In another study, AlexNet, VGG-16, VGG-19, SqueezeNet, GoogleNet, MobileNet-V2, ResNet-18, ResNet-50, ResNet-101 and Classification models using Xception convolutional neural networks are used. Among all models, ResNet-101 and Xception have the best performance. It can be said that ResNet-101 and Xception convolutional networks are quite successful in distinguishing COVID-19. ($\text{AUC}_{\text{ResNet-101}}$= 0.994, $\text{Sensitivity}_{\text{ResNet-101}}$=100%, $\text{Specificity}_{\text{ResNet-101}}$=99.02%, $\text{Accuracy}_{\text{ResNet-101}}$=99.51%, $\text{AUC}_{\text{Xception}}$ = 0.994, Sensitivity$_{\text{Xception}}$= 98%, 04, Specificity$_{\text{Xception}}$= 100%, Accuracy$_{\text{Xception}}$= 99.02%). Considering all performance metrics, it can be said that the model created using the ResNet-101 convolutional network is successful in COVID-19 classification [18].

In another study, a new DarkNet model was created using the YOLO deep learning architecture, which classifies automated COVID-19 as binary (positive / no finding) and multi-class (positive / no finding/pneumonia) using raw chest x-ray images. The accuracy value of the created model was obtained as 98.08% for binary classification and 87.02% for multi-class classification [19].

A recent study has focused on predicting whether a person is COVID-19 (positive/negative) in the early stage of the disease using ANN, RF and glmnet classifier on solely from 14 different blood counts. Models created in the study using ANN, RF and glmnet classifier can classify COVID-19 disease with accuracy rates of 80%, 86% and 84%, respectively [20].

In another study, using the same data set, an experimental study was conducted to predict patients with negative or positive COVID-19 class using SMOTE and ANN models. While the accuracy value was found to be 86% for the ANN model applied before balancing the data set, the accuracy value was obtained as 90% for the ANN model applied to the data balanced with SMOTE [21].

As a result, there have been many reported studies concerning medical imaging techniques and deep learning models to diagnose the COVID-19. However, the current study proposes a random forest model to classify the COVID-19 based on the complete blood count results. The Random Forest algorithm gives the best predictions in the classification of COVID-19 (99.76% of accuracy). Besides, in the classification of Covid-19, it can be recommended to apply meta-learning algorithms as they can give high predictive results.

REFERENCES

[1] Z. Y. Zu, M. D. Jiang, P. P. Xu, W. Chen, Q. Q. Ni, G. M. Lu, *et al.*, "Coronavirus disease 2019 (COVID-19): a perspective from China," *Radiology,* p. 200490, 2020.

[2] R. Mitchell, J. Michalski, and T. Carbonell, *An artificial intelligence approach*: Springer, 2013.

[3] A. Rajkomar, J. Dean, and I. Kohane, "Machine learning in medicine," *New England Journal of Medicine,* vol. 380, pp. 1347-1358, 2019.

[4] R. Bost, R. A. Popa, S. Tu, and S. Goldwasser, "Machine learning classification over encrypted data," in *NDSS*, 2015, p. 4325.

[5] D. Dietrich, B. Heller, and B. Yang, "Data science & big data analytics: Discovering," *Analyzing, Visualizing and Presenting Data,* 2015.

[6] E. Güldoğan, A. K. Arslan, and J. Yağmur, "Çeşitli Çekirdek Fonksiyonları ile Oluşturulan Destek Vektör Makinesi Modellerinin Performanslarının İncelenmesi: Bir Klinik Uygulama," *Firat Tip Dergisi,* vol. 22, 2017.

[7] D. Kılınç, E. Borandağ, F. Yücalar, V. Tunalı, M. Şimşek, and A. Özçift, "KNN algoritması ve r dili ile metin madenciliği kullanılarak bilimsel makale tasnifi," 2016.

[8] T. M. Oshiro, P. S. Perez, and J. A. Baranauskas, "How many trees in a random forest?," in *International workshop on machine learning and data mining in pattern recognition*, 2012, pp. 154-168.

[9] C.-F. Chien and L.-F. Chen, "Data mining to improve personnel selection and enhance human capital: A case study in high-technology industry," *Expert Systems with applications,* vol. 34, pp. 280-290, 2008.

[10] J. Han and M. Kamber, "Data Mining Concepts and Techniques, Morgan Kaufmann Publishers," *San Francisco, CA,* pp. 335-391, 2001.

[11] W. Y. Loh, "Classification and regression trees," *Wiley Interdisciplinary Reviews: Data Mining and Knowledge Discovery,* vol. 1, pp. 14-23, 2011.

[12] Q. Song, W. Hu, and W. Xie, "Robust support vector machine with bullet hole image classification," *IEEE transactions on systems, man, and cybernetics, part C (applications and reviews),* vol. 32, pp. 440-448, 2002.

[13] J. Han, J. Pei, and M. Kamber, *Data mining: concepts and techniques*: Elsevier, 2011.

[14] L. Breiman, "Random forests," *Machine learning,* vol. 45, pp. 5-32, 2001.

[15] K. J. Archer and R. V. Kimes, "Empirical characterization of random forest variable importance measures," *Computational statistics & data analysis,* vol. 52, pp. 2249-2260, 2008.

[16] H. A. Rothan and S. N. Byrareddy, "The epidemiology and pathogenesis of coronavirus disease (COVID-19) outbreak," *Journal of autoimmunity,* p. 102433, 2020.

[17] S. Chatterjee, F. Saad, C. Sarasaen, S. Ghosh, R. Khatun, P. Radeva, *et al.*, "Exploration of Interpretability Techniques for Deep COVID-19 Classification using Chest X-ray Images," *arXiv preprint arXiv:2006.02570,* 2020.

[18] N. Khadem and A. Mohammadi, "Application of Deep Learning Technique to Manage COVID-19 in Routine Clinical Practice Using CT Images: Results of 10 Convolutional Neural Networks."

[19] T. Ozturk, M. Talo, E. A. Yildirim, U. B. Baloglu, O. Yildirim, and U. R. Acharya, "Automated detection of COVID-19 cases using deep neural networks with X-ray images," *Computers in Biology and Medicine,* p. 103792, 2020.

[20] A. Banerjee, S. Ray, B. Vorselaars, J. Kitson, M. Mamalakis, S. Weeks, *et al.*, "Use of Machine Learning and Artificial Intelligence to predict SARS-CoV-2 infection from Full Blood Counts in a population," *International immunopharmacology,* vol. 86, p. 106705, 2020.

[21] M. Yavaş, A. Güran, and M. Uysal, "Covid-19 Veri Kümesinin SMOTE Tabanlı Örnekleme Yöntemi Uygulanarak Sınıflandırılması," *Avrupa Bilim ve Teknoloji Dergisi,* pp. 258-264.

BIOGRAPHIES

Şeyma YAŞAR obtained her BSc. Degree in mathematics from Gaziosmanpasa University in 2009. She received an MSc. degree in biostatistics and medical informatics from the Inonu University in 2018. She currently continues PhD degrees in biostatistics and medical informatics from the Inonu University. In 2014, she joined the Department of Biostatistics and Medical Informatics at Inonu University as a researcher assistant. Her research interests are cognitive systems, data mining, machine learning, deep learning.

Cemil ÇOLAK obtained his BSc. Degree in statistics from Ondokuz Mayıs University in 1999. He received MSc. degree in Biostatistics from the Inonu University in 2001, and a Ph.D. degree in the Graduate Department of Biostatistics and Medical Informatics of Ankara University in 2005. His research interests are cognitive systems, data mining, reliability, and biomedical system, genetics, and bioengineering. In 2016, he joined the Department of Biostatistics and Medical Informatics at Inonu University as a Professor, where he is presently a professor. He is active in teaching and research in general image processing, artificial intelligence, data mining, analysis.

CLASSIFICATION OF HYPOTHYROID DISEASE WITH EXTREME LEARNING MACHINE MODEL

I. Balikci Cicek, and Z. Kucukakcali

Abstract— *Aim:* In this study, it is aimed to classify hypothyroidism by applying the Extreme Learning Machine model, which is one of the artificial neural network models, on the open access Hypothyroid dataset.

Materials and Methods: In this study, the data set named "Hypothyroid Disease Data Set" was obtained from https://www.kaggle.com/nguyenthilua/hypothyroidcsv. Extreme Learning Machine model, one of the artificial neural network models, was used to classify hypothyroidism. The classification performance of the model was evaluated with classification performance criteria such as accuracy, sensitivity, positive predictive value, negative predictive value and F1-score.

Results: The accuracy obtained from the model was calculated as 0.922, balanced accuracy 0.523, sensitivity 1, positive predictive value 0.922, negative predictive value 1 and F1-score 0.959.

Conclusion: The findings obtained from this study showed that the extreme learning machine model used gave successful predictions in the classification of hypothyroidism.

Keywords— Hypothyroidism, classification, artificial neural networks, extreme learning machine.

1. INTRODUCTION

THE THYROID gland is an endocrine organ located in the front of the neck just in front of the trachea. Its weight is approximately 15-25 grams, each lobe is 2.5-4 cm long, 1.5-2 cm wide, and 1-1.5 cm thick. It is the largest endocrine organ in the human body, its main task is the secretion of thyroid hormones [1]. Thyroid hormones are hormones in protein structure released from the thyroid gland behind the anterior thyroid muscles on the anterior surface of the trachea. They play a fundamental role in growth, development and metabolism [2]. Thyroid hormones have a role in regulating the functions of almost every cell and tissue in our body. A small amount of secretion causes body functions to slow down, and excessive secretion causes body functions to accelerate [3]. Hypothyroidism is a clinical condition characterized by a general slowdown in metabolic events that occurs as a result of thyroid hormone deficiency or when thyroid hormones cannot be effective [4].

The signs and symptoms of hypothyroidism vary according to the rate and severity of the thyroid hormone deficiency and the age at which it occurs. Generally, when thyroid hormone deficiency develops slowly, hypothyroidism has an insidious and slow onset [5]. In adults, it can occur in different pictures ranging from a subclinical course to myxedema coma. Hypothyroidism is a common disease with an incidence of up to 20% [6]. The incidence of hypothyroidism increases with age, especially in women. The prevalence of hypothyroidism is 1.4% in women and 0.1% in men [2]. If left untreated, hypothyroidism can cause more serious complications and even be life-threatening. Serious complications include low metabolism and heart rate, heart failure, severe, life-threatening depression, and coma [7].

Artificial neural networks (ANN) are mathematical models inspired by the functions of biological neural networks. Artificial Neural Networks are computer systems that can learn about events and determine how to react to events coming from the environment by using examples that are products of real brain functions. ANN are information processing systems that have the ability to generate, create and discover new information by way of learning imitating the human brain. ANN can provide nonlinear modeling between input and output variables without the need for any prior knowledge and any assumptions [8]. Today, ANN is effective in many scientific fields due to its superior features such as nonlinearity, learning, generalization, hiding information, producing information about unseen examples, working with incomplete information, classification, adaptability, and error tolerance. At the same time, due to these features of ANN, it has found application place in many scientific fields [9].

Extreme Learning Machine (ELM) is a machine learning technique based on artificial neural networks. ELM algorithm has been developed for machine learning of a single layer feed forward neural network. ELM performs better than traditional feed forward networks in solving many problems [10]. ELM learning algorithm has advantages such as good generalization performance, extremely fast learning ability and low processing complexity [11]. Extreme learning machine using feed-forward neural networks with a single hidden layer have been proposed to solve classification and regression problems. In the basic logic of ELM, input weights and threshold values are assigned randomly and least squares method is used as an algorithm. Randomly assigned input weights and thresholds increase the classification ability of extreme learning machine [10].

✉ **İpek BALIKÇI ÇİÇEK,** Inonu University Department of Biostatistics and Medical Informatics, Faculty of Medicine, Malatya, Turkey, (ipek.balikci@inonu.edu.tr)

Zeynep KÜÇÜKAKÇALI, Inonu University Department of Biostatistics and Medical Informatics, Faculty of Medicine, Malatya, Turkey, (zeynep.tunc@inonu.edu.tr)

Manuscript received Sep 21, 2020; accepted Nov. 19, 2020.
Digital Object Identifier:

2. MATERIAL AND METHODS

2.1. Dataset

In the study, the data set named "Hypothyroid Disease Data Set" was obtained from https://www.kaggle.com/nguyenthilua/hypothyroidcsv to examine the working principle of the Extreme Learning Machine (ELM) method [12]. In the dataset used, there were a total of 3772 patients, 3481 (92.3%) hypothyroid and 291 (7.7%) negative. The variables included in the dataset are given in Table 1.

TABLE I
HYPOTHYROID DATA SET

Variable	**Variable Description**
Age	Integer
Sex	Male(M), Female(F)
On thyroxine	False(f), True(t)
Query on thyroxine	False(f), True(t)
On antythyroid	False(f), True(t)
Sick	False(f), True(t)
Pregnant	False(f), True(t)
Thyroid surgery	False(f), True(t)
T131 treatment	False(f), True(t)
Query Hypothyroid	False(f), True(t)
Query Hyperthyroid	False(f), True(t)
Lithium	False(f), True(t)
Goiter	False(f), True(t)
Tumor	False(f), True(t)
Hypopitutory	False(f), True(t)
Psych	False(f), True(t)
Tsh measured	False(f), True(t)
TSH	Real
T3 measured	False(f), True(t)
T3	Real
TT4 measured	False(f), True(t)
TT4	Real
T4U measured	False(f), True(t)
T4U	Real
FTI Measured	False(f), True(t)
FTI	Real
TBG Measured	False(f), True(t)
TBG	Real
Referal source	SVHC, other, SVI, STMW, SVHD
Class	negative, hypothyroid

Among the variables included in the dataset, the variables used and the explanations of these variables are given in Table 2.

TABLE II
VARIABLES USED IN THE DATASET AND THEIR DESCRIPTIVE PROPERTIES

Variable	**Variable Description**	**Variable Type**	**Variable Role**
Class	Negative, Hypothyroid	Qualitative	Dependent/ Target
Age	Age	Quantitative	Independent/ Predictor
Sex	Male (M), Female (F)	Qualitative	Independent/ Predictor
On thyroxine	False(f), True(t)	Qualitative	Independent/ Predictor
On antythyroid	False(f), True(t)	Qualitative	Independent/ Predictor
Sick	False(f), True(t)	Qualitative	Independent/ Predictor
Pregnant	False(f), True(t)	Qualitative	Independent/ Predictor
Thyroid surgery	False(f), True(t)	Qualitative	Independent/ Predictor
T131 treatment	False(f), True(t)	Qualitative	Independent/ Predictor
Query Hypothyroid	False(f), True(t)	Qualitative	Independent/ Predictor
Query Hyperthyroid	False(f), True(t)	Qualitative	Independent/ Predictor
Lithium	False(f), True(t)	Qualitative	Independent/ Predictor
Goiter	False(f), True(t)	Qualitative	Independent/ Predictor
Tumor	False(f), True(t)	Qualitative	Independent/ Predictor
Hypopitutory	False(f), True(t)	Qualitative	Independent/ Predictor
Psych	False(f), True(t)	Qualitative	Independent/ Predictor
TSH	Real	Quantitative	Independent/ Predictor
T3	Real	Quantitative	Independent/ Predictor
TT4	Real	Quantitative	Independent/ Predictor
T4U	Real	Quantitative	Independent/ Predictor
FTI	Real	Quantitative	Independent/ Predictor
Referal source	SVHC, other, SVI, STMW, SVHD	Qualitative	Independent/ Predictor

3. EXTREME LEARNING MACHINE (ELM)

Artificial neural networks has become very easy today to train just like a human brain. Teaching the knowledge that the human brain has acquired through experience to machines by using artificial neural networks has become the main subject of information technologies. An artificial neural network model consists of input layer, output layer and hidden layer. Artificial neural networks are divided into two as feed forward artificial neural networks and feedback artificial neural networks. In feedforward neural networks, neurons are in layers from input to output. Incoming information is transmitted to the input layer, middle layer and output layer respectively. In feedback networks, it is not only given as input to the layer of the cell that comes after it. It can also be connected as input to any cell in its previous layer or in its own layer [10].

Extreme Learning Machine were first introduced by Huang et al in 2006. Extreme learning machine is a fully connected artificial neural network model consisting of three layers as input layer, one hidden layer and output layer [13]. Extreme learning machine are basically similar to artificial neural networks with one hidden layer. For this reason, the working principle of extreme learning machine is to a certain extent the same with the working principles of artificial neural networks. However, in extreme learning machine, the weights in the hidden layer are randomly assigned and these values do not change in the further stage of education. On the other hand, the weights between the hidden layer and the output layer are determined analytically and quickly with the help of a linear model in one go. While activation functions such as Sigmoid, Sinus and Gauss are used in the hidden layer, linear activation functions are used in the output layer. The weights in the input layer of the feed forward neural network do not affect the performance of the network with one hidden layer [14].

ELM has many advantages over classical ANN networks trained with gradient-based learning algorithms. These advantages can be listed as follows; ELM's learning process is extremely fast. This time is usually in the level of seconds, in some applications even less than a second. ELM has better generalization ability than derivative based backpropagation algorithm in many cases. Classical derivative-based training algorithms and other learning algorithms may be faced with many situations such as stuck to local minimums, inappropriate learning rate, excessive learning and memorization. In order to solve these problems, methods such as early stopping, adding regulation parameters, breaking the weight and using validity sets are used. The ELM learning algorithm tends to reach the solution directly without such intermediate processes and is therefore simpler than the learning algorithms used in classical artificial neural networks [15].

3.1. Performance evaluation metrics

Performance metrics obtained by using the classification matrix (Table 3) given below were used in the performance evaluation of the extreme learning machine model. The classification matrix of performance metrics is given in Table 3.

TABLE III

CLASSIFICATION MATRIX FOR CALCULATING PERFORMANCE METRICS

		Real		
		Positive	**Negative**	**Total**
Predicted	Positive	True positive (TP)	False negative (FN)	TP+FN
	Negative	False positive (FP)	True negative (TN)	FP+TN
	Total	TP+FP	FN+TN	TP+TN+FP+FN

Accuracy = (TP+TN)/(TP+TN+FP+FN)

Sensitivity = TP/(TP+FP)

Positive predictive value = TP/(TP+FN)

Negative predictive value =TN/(TN+FP)

F1-score = (2*TP)/(2*TP+FP+FN)

4. DATA ANALYSIS

Quantitative data are expressed as mean ± standard deviation, median (minimum-maximum), and qualitative data as number (percentage). Conformity to normal distribution was evaluated by the Kolmogorov-Smirnov test. In terms of independent variables, whether there is a statistically significant difference between the "Negative" and " Hypothyroid " groups, which are the categories of the dependent / target variable (class), and whether there is a relationship, Mann-Whitney U test, Pearson chi-square test, Continuity Correction test and Fisher's Exact test it was examined using the chi-square test Values of $p<0.05$ were considered statistically significant. IBM SPSS Statistics 26.0 package program was used for all analyzes.

For the validity of the model, a 10-fold cross-validation method was used. In the 10-fold cross-validation method, all data is divided into 10 equal parts. One part is used as a test set and the remaining 9 parts are used as a training data set and this process is repeated 10 times.

5. RESULTS

Descriptive statistics for the independent variables examined in this study are given in Table 4. According to the findings in Table 4; there is a statistically significant difference between the dependent / target variable groups in terms of TSH, T3, TT4, T4U, FTI variables ($p<0.05$).

According to the findings in Table 5; there is a statistically significant relationship between the sex, onthyroxine, pregnant, queryhypothyroid and referralsource variables and the dependent / target variable (class) groups ($p<0.05$).

TABLE IV

DESCRIPTIVE STATISTICS FOR QUANTITATIVE INDEPENDENT VARIABLES

Variables	Class		**p-value***
	hypothyroid	negative	
	Median(min-max)	**Median(min-max)**	
age	54 (1-455)	55 (1-88)	0.905
TSH	1.2 (0.005-145)	12 (0.015-530)	**<0.001**
T3	2 (0.05-10.6)	1.5 (0.2-4.09)	**<0.001**
TT4	105 (19-430)	77 (2-44076)	**<0.001**
T4U	0.97 (0.25-2.32)	1.01 (0.56-1.65)	**0.005**
FTI	108 (17-395)	77.5 (2-153)	**<0.001**

***: Mann Whitney U test**

TABLE V
DESCRIPTIVE STATISTICS FOR QUALITATIVE INDEPENDENT VARIABLES

Variables		Class		**p-value**
		hypothyroid	negative	
		Number(%)	**Number(%)**	
sex	female	2265 (67.8)	215 (76.8)	**0.002***
	male	1077 (32.2)	65 (23.2)	
onthyroxine	false	3026(86.9)	282 (96.9)	**<0.001****
	true	455 (13.1)	9 (3.1)	
onantithyroidmedication	false	3439 (98.8)	290 (99.7)	0.254***
	true	42 (1.2)	1 (0.3)	
sick	false	3345 (96.1)	280 (96.2)	1**
	true	136 (3.9)	11 (3.8)	
pregnant	false	3428 (98.5)	291 (100.0)	**0.032*****
	true	53 (1.5)	0 (0.0)	
thyroidsurgery	false	3430 (98.5)	289 (99.3)	0.433***
	true	51(1.5)	2 (0.7)	
I131treatment	false	3427 (98.4)	286 (98.3)	0.804***
	true	54 (1.6)	5 (1.7)	
queryhypothyroid	false	3286 (94.4)	252 (86.6)	**<0.001***
	true	195 (5.6)	39 (13.4)	
queryhyperthyroid	false	3259 (93.6)	276 (94.8)	0.484**
	true	222 (6.4)	15 (5.2)	
lithium	false	3464 (99.5)	290 (99.7)	1***
	true	17 (0.5)	1 (0.3)	
goitre	false	3447 (99.02)	291 (100.0)	0.107***
	true	34 (0.98)	0 (0.0)	
tumor	false	3393 (97.5)	283 (97.3)	0.971**
	true	88 (2.5)	8 (2.7)	
hypopituitary	false	3480 (99.97)	291 (100.0)	1***
	true	1 (0.03)	0 (0.0)	
psych	false	3305 (94.9)	283 (97.3)	0.107**
	true	176 (5.1)	8 (2.7)	
referralsource	SVHC	375 (10.8)	11 (3.8)	**0.002***
	other	2028 (58.3)	173(59.5)	
	SVI	937 (26.9)	97 (33.3)	
	STMW	105 (3.0)	7 (2.4)	
	SVHD	36 (1.0)	3 (1.0)	

***: Pearson chi-square test, **: Continuity Correction test, ***: Fisher's Exact test**

The classification matrix for the Extreme Learning Machine model used to classify the hypothyroid dataset in this study is given in Table 6 below.

TABLE VI
CLASSIFICATION MATRIX FOR THE EXTREME LEARNING MACHINE MODEL

Prediction	**Reference**		
	hypothyroid	negative	**Total**
hypothyroid	485	41	526
negative	0	2	2
Total	485	43	528

Values for the metrics of the classification performance in the testing phase of the model are given in Table 7. The accuracy obtained from the model was calculated as 0.922, sensitivity 1, positive predictive value 0.922, negative predictive value 1 and F1-score 0.959.

TABLE VII
VALUES FOR THE METRICS OF THE CLASSIFICATION PERFORMANCE IN THE TESTING PHASE OF THE MODEL

Metric	**Value**
Accuracy	0.922
Sensitivity	1
Positive predictive value	0.922
Negative predictive value	1
F1-score	0.959

6. DISCUSSION

Inability of the thyroid gland to produce as many hormones as necessary is called hypothyroidism. Patients with hypothyroidism experience weight gain, tendency to sleep, reduced exercise capacity and cold intolerance. In more severe patients, constipation, thickening of the voice, hair loss, broken nails, increased cholesterol levels, myxedema, cretinism, skin dryness and goiter are seen. The most common cause of hypothyroidism is iodine deficiency [16]. Cardiovascular, gastrointestinal and metabolic diseases (such as sinus bradycardia, altered gastrointestinal secretion and motility) are the main known clinical symptoms of hypothyroidism. These symptoms can cause anatomical disorders, cardiovascular and cerebrovascular diseases. In addition, it causes delay in skeletal development and mental disorders [17].

ANN can establish a relationship between inputs and outputs determined depending on various parameters of a system by using the properties of biological nervous systems [18]. Feedforward ANNs are widely used for function approach in many fields due to their distinctive features [19]. When evaluating ANN as a classifier, the number of neurons in the hidden layer, the values of the weights between the input layer and the hidden layer and between the hidden layer and the output layer and the selection of the learning algorithm play an important role. The biggest disadvantage of ANN is the use of gradient descent algorithm to adjust weights and parameters during the training process, which makes the performance of the model time consuming and increases the computational burden [20].

ELM is a feed forward ANN algorithm with a single hidden layer. In ELM, unlike conventional feed forward ANNs, the input weights and the latent threshold value are given randomly, and the weights of the neurons in the output layer are calculated. In this approach they suggest, performing all processes without repetition during training enables the learning phase to be completed in a very short time for most applications. Also, studies have shown that ELM has a better classification performance than most gradient-based learning [21].

In this study, an extreme learning machine model, which is an artificial neural network method, was applied to the open source dataset named "Hypothyroid Disease Data Set". Among the performance criteria obtained from the extreme learning machine model, accuracy was calculated as 0.922, sensitivity 1, positive predictive value 0.922, negative predictive value 1 and F1-score 0.959.

As a result, in the study conducted with hypothyroid dataset, the extreme learning machine model used yielded successful results in the classification of hypothyroidism.

REFERENCES

[1] L. E. Braverman and D. Cooper, Werner & Ingbar's the thyroid: a fundamental and clinical text: Lippincott Williams & Wilkins, 2012.

[2] A. S. Fauci, Harrison's principles of internal medicine vol. 2: McGraw-Hill, Medical Publishing Division New York, 2008.

[3] T. R. Kıran, "Hipertiroidili ve hipotiroidili hastalarda oksidatif stres parametreleri ve adenozin deaminaz aktivitesi," 2007.

[4] S. Koloğlu and G. Erdoğan, "Tiroid: Genel Görüş ve Bilgiler," Koloğlu Endokrinoloji, Temel ve Klinik, vol. 2, pp. 155-72.

[5] M. Özata, Tiroid hastalıkları: tanı ve tedavisi: GATA Basımevi, 2003.

[6] D. C. Bauer, B. Ettinger, and W. S. Browner, "Thyroid function and serum lipids in older women: a population-based study," The American journal of medicine, vol. 104, pp. 546-551, 1998.

[7] J. Norman, "Hypothyroidism: Too Little Thyroid Hormone," 2013.

[8] S. Haykin, "Neural Networks, a comprehensive foundation, Prentice-Hall Inc," Upper Saddle River, New Jersey, vol. 7458, pp. 161-175, 1999.

[9] H. H. Dodurgalı, "Karınca Kolonisi Optimizasyonu İle Eğitilmiş Çok Katmanlı Yapay Sinir Ağı İle Sınıflandırma," Fen Bilimleri Enstitüsü, 2010.

[10] C. Karakuzu and A. Bakirci, "Dynamic System Identification Based on an Ensemble of ELMs."

[11] Ö.F. Alçin, "Aşırı öğrenme makinelerinin seyrek geri çatma algoritmaları ile optimizasyonu/Optimization of extreme learning machine with sparse recovery algorithms," 2015.

[12] Available: https://www.kaggle.com/nguyenthilua/hypothyroidcsv.

[13] G.-B. Huang, Q.-Y. Zhu, and C.-K. Siew, "Extreme learning machine: theory and applications," Neurocomputing, vol. 70, pp. 489-501, 2006.

[14] M. Özçalici, "Aşiri Öğrenme Makineleri İle Hisse Senedi Fiyat Tahmini," Hacettepe Üniversitesi İktisadi ve İdari Bilimler Fakültesi Dergisi, vol. 35, pp. 67-88, 2017.

[15] O. Kaynar, H. Arslan, Y. Görmez, and Y. E. Işık, "Makine Öğrenmesi ve Öznitelik Seçim Yöntemleriyle Saldırı Tespiti," Bilişim Teknolojileri Dergisi, vol. 11, pp. 175-185, 2018.

[16] C. Schmid, C. Zwimpfer, M. Brändle, P. A. Krayenbühl, J. Zapf, and P. Wiesli, "Effect of thyroxine replacement on serum IGF-I, IGFBP-3 and the acid-labile subunit in patients with hypothyroidism and hypopituitarism," Clinical endocrinology, vol. 65, pp. 706-711, 2006.

[17] P.-Q. Yuan and H. Yang, "Hypothyroidism increases Fos immunoreactivity in cholinergic neurons of brain medullary dorsal vagal complex in rats," American Journal of Physiology-Endocrinology and Metabolism, vol. 289, pp. E892-E899, 2005.

[18] M. Akın and M. Ceylan, "İyi Huylu Karaciğer Lezyonlarının Sınıflandırılmasında Yapay Sinir Ağı ve Aşırı Öğrenme Makinesi'nin Karşılaştırılması Comparison of Artificial Neural Network and Extreme Learning Machine in Benign Liver Lesions Classification."

[19] K. Hornik, M. Stinchcombe, and H. White, "Multilayer feedforward networks are universal approximators," Neural networks, vol. 2, pp. 359-366, 1989.

[20] W. Sun, C. Wang, and C. Zhang, "Factor analysis and forecasting of CO2 emissions in Hebei, using extreme learning machine based on particle swarm optimization," Journal of cleaner production, vol. 162, pp. 1095-1101, 2017.

[21] N.-Y. Liang, G.-B. Huang, P. Saratchandran, and N. Sundararajan, "A fast and accurate online sequential learning algorithm for feedforward networks," IEEE Transactions on neural networks, vol. 17, pp. 1411-1423, 2006.

BIOGRAPHIES

İpek BALIKÇI ÇİÇEK obtained her BSc. degree in mathematics from Çukurova University in 2010. She received MSc. degree in biostatistics and medical informatics from the Inonu University in 2018. She currently continues Ph.D. degrees in biostatistics and medical informatics from the Inonu University. In 2014, she joined the Department of Biostatistics and Medical Informatics at Inonu University as a researcher assistant. Her research interests are cognitive systems, data mining, machine learning, deep learning.

Zeynep KÜÇÜKAKÇALI obtained her BSc. degree in mathematics from Çukurova University in 2010. She received MSc. degree in biostatistics and medical informatics from the Inonu University in 2018. She currently continues Ph.D. degrees in biostatistics and medical informatics from the Inonu University. In 2014, she joined the Department of Biostatistics and Medical Informatics at Inonu University as a researcher assistant. Her research interests are cognitive systems, data mining, machine learning, deep learning.

CLASSIFICATION OF CORONARY ARTERY DISEASES USING STACKING ENSEMBLE LEARNING METHOD

A. Doganer, and M. Kirisci

Abstract*— *Aim: Coronary artery disease is one of the most fatal diseases in the the society. Early diagnosis and treatment of coronary artery disease plays an important role in reducing the number of deaths. In this study, it is aimed to classify coronary artery disease by Stacking based ensemble learning methods.

Material and Methods: The study was obtained from the data of 244 patients with coronary artery disease and 116 individuals without coronary artery disease who were treated in Kahramanmaras Sutcu Imam University Health Practice and Research Hospital. The data were obtained retrospectively. The data set consists of 15 predictor variables and 1 dependent variable. In the classification process, Naive Bayes, Sequential Minimal Optimization, Random Forest classifiers and Stacking ensemble learning method were applied. A 10-fold cross validation method was applied to the model. Accuracy, sensitivity, specificity, F-measure and AUC metrics were applied to evaluate the performance of classifiers. The most essential variables in predicting coronary artery disease have been determined.

Results: ACC = 0.774, SEN = 0.888, SPE = 0.719, F = 0.718 and AUC = 0.913 values were obtained with the Naive Bayes classifier in the study. ACC = 0.883, SEN = 0.733, SPE = 0.955, F = 0.802 and AUC = 0.844 were obtained with the SMO classifier. ACC = 0.908, SEN = 0.853, SPE = 0.934, F = 0.857 and AUC = 0.957 were obtained with Random Forest classifier. ACC = 0.933, SEN = 0.905, SPE = 0.946, F = 0.897 and AUC = 0.959 values were obtained with the stacking ensemble learning method. BUN, MPV, Age, AST and Monocyte variables were determined as the most essential factors in the classification of coronary artery disease, respectively.

Conclusion: Stacking ensemble learning method provided the highest accuracy performance in the classification of coronary artery disease. Stacking ensemble learning method gives successful results in the classification of coronary artery diseases.

***Keywords*—** *Coronary Artery Disease, Stacking Ensemble Learning Methods, Artificial Intelligence, Meta Classifier.*

1. INTRODUCTION

CORONARY artery diseases are among the most fatal diseases in the world [1]. Coronary artery disease is a cardiovascular disease expressed by the occlusion or narrowing of the vessels due to the presence of plaques in the arteries that cause atherosclerosis [2]. Coronary artery disease is stated as the cause of one out of every 7 deaths in the United States [3]. Coronary artery disease is among the diseases with a serious risk of death in both developing and developed countries [4]. Therefore, developments in the early diagnosis and treatment of coronary artery disease are substantial. Early diagnosis of coronary artery disease prevents possible deaths. Coronary angiography and visualisation techniques are used for the detection of coronary artery disease. Coronary angiography is a diagnostic method that is difficult to perform and includes certain risks [5]. Machine learning methods have made a powerful alternative in the diagnosis of coronary artery diseases.

Machine learning methods have been used frequently in the early diagnosis of diseases in recent years. Machine learning methods, which provide a successful classification performance, make an important contribution to the early diagnosis of heart diseases. Machine learning methods and applications related to artificial intelligence in the early diagnosis of coronary artery disease have attracted the attention of many researchers. Researchers have suggested some models for early diagnosis of coronary artery disease. Poss et al. (2018) used machine learning methods to reveal the biomarkers of coronary artery diseases [6]. Verma et al. (2016) proposed a hybrid model consisting of machine learning methods for the diagnosis of coronary artery disease [7]. Zheng et al. (2011) used machine learning methods in their studies to perform coronary artery segmentation [8]. Nikan et al. (2017) compared the performances of 3 different classification models to predict coronary artery risk [9]. Artificial intelligence methods provide successful results for early diagnosis of diseases using ECG signals. Acharya et al. (2017) proposed a model for the automatic detection of coronary artery disease using convolutional neural networks (CNN) and ECG signals [10].

Many researchers have tried to detect coronary artery disease by applying different models. Some classifiers do not provide sufficient performance for classifying coronary artery diseases. In this study, it is aimed to develop a model for the early diagnosis of coronary artery diseases with the Stacking ensemble learning method. With the Stacking ensemble learning method, it is aimed to achieve high accuracy performance in the detection of coronary artery diseases. In the model, the performances of Naive Bayes, Sequential Minimal Optimization, Random Forest classifiers and Stacking ensemble learning methods are compared.

2. MATERIAL AND METHODS

The data set of the study consists of patient records of the Cardiovascular Surgery Service of Kahramanmaras Sutcu

✉ **Adem DOGANER,** Kahramanmaras Sutcu Imam University Department of Biostatistics and Medical Informatics, Faculty of Medicine, Kahramanmaras, Turkey, (adem_doganer@hotmail.com)

Mehmet KIRISCI, Kahramanmaras Sutcu Imam University Department of Cardiovascular Surgery, Faculty of Medicine, Kahramanmaras, Turkey, (drkirisci@yahoo.com)

Manuscript received Sep 9, 2020; accepted Nov. 11, 2020.
Digital Object Identifier:

Imam University Health Practice and Research Hospital. The data were obtained retrospectively. The study includes data on patients with coronary artery disease and individuals without coronary artery disease. Permission was obtained from the clinical research ethics committee of Kahramanmaras Sutcu Imam University to conduct the study (Ethics Committee Approval No: 2019/21 Decision No: 13). The number of samples in the study was determined by power analysis. With α: 0.05 first type error level, B: 0.20 second type error level and 0.80 test power, taking into account the accuracy estimation value 91.2% in the reference study [11], a total of 360 cases were planned to be included in two groups for a 5% deviation level. The data set consists of a total of 365 data, including 249 patients with coronary artery disease and 116 individuals without coronary artery disease. The data set consists of 16 variables, including 1 dependent variable and 15 predictor variables. The variables in the data set are shown in Table 1.

TABLE I
FEATURES OF VARIABLES

Variable	Variable Description	Variable Role
Type of Diseases	CABG/ M. Bridge	Output
Gender	Male/Female	Input
Age (Years)	Quantitative	Input
Hemoglobin (HB)	Quantitative	Input
Hematocrit (HCT)	Quantitative	Input
Neutrophil	Quantitative	Input
Lymphocyte	Quantitative	Input
Monocyte	Quantitative	Input
Platelet (PLT)	Quantitative	Input
Mean Platelet Volume (MPV)	Quantitative	Input
Blood Urea Nitrogen (BUN)	Quantitative	Input
Creatinine	Quantitative	Input
Aspartate Aminotransferase (AST)	Quantitative	Input
Alanine Aminotransferase (ALT)	Quantitative	Input
Total Cholesterol	Quantitative	Input
Low Density Lipoprotein (LDL)	Quantitative	Input

Outliers in the dataset are analyzed with the local outlier factor (LOF) algorithm. Data with outliers were excluded from the study [12]. Standardization (Z-transform) process was applied to the quantitative data for the classification process. In the study, the performances of Random Forest classifier (RF), Naive Bayes (NB), Sequential Minimal Optimization (SMO) and stacking ensemble learning methods were compared for the classification of diseases. Individual classification performances of RF, NB and SMO classifiers were evaluated. In addition to individual classification, the RF, NB and SMO classifiers were included in the ensemble model for Stacking ensemble learning method. The Logistic Classifier has been implemented as a meta classifier for the ensemble model. The K fold-10 cross validation method was applied to train the model. Grid search algorithm was used for hyperparameter optimization of classifiers. Weka 3.9.3 (Waikato Environments for Knowledge Analysis) and R 3.3.2 software were used to perform the classification processes.

2.1. Stacking Ensemble Learning Methods

The ensemble learning method has made a powerful alternative to other machine learning methods for classifying data. Ensemble learning methods are machine learning methods based on the principle that by combining the estimates of more than one classifier, more reliable and more accurate estimates can be procured than the estimates of a single classifier [13]. In order to increase performance in ensemble learning methods, appropriate defragmentation method and appropriate classifiers should be included in the model. Stacking ensemble learning method is one of the ensemble learning methods that can provide high classification performance. The stacking ensemble learning method includes more than one classifier to the model. It transmits the estimate of each classifier in the model to a meta classifier. The meta classifier produces the ensemble estimate by processing the estimates from the classifiers in the model as input [14]. In the study, Random Forest (RF), Naive Bayes (NB) and Sequential Minimal Optimization (SMO) classifiers were included in the Stacking ensemble learning model. Logistic classifier ranks as a meta classifier in the study.

2.2. Random Forest

Random Forest is a classifier consisting of many decision trees. The Random forest classifier developed by Breiman, different subsets created by bootstrap sampling method from the data set are trained with each decision tree [15]. Classification is initiated by determining the number of decision trees and the number of variables that will perform the division in the model. In the random forest classifier, the decision is made by "majority voting" method among the predictions of each decision tree. The random forest classifier, which provides strong and high performance, is also highly resistant to overfitting problem [16-17].

2.3. Naive Bayes

Naive Bayes, a strong and plain classifier, classifies on the basis that all variables are independent [18]. The Naive Bayes classifier, which takes its theoretical basis from the Bayes theorem, is capable of fast learning. The Naive Bayes classifier, which classifies using probabilistic methods and makes use of prior probabilities. It can provide high performance for large data sizes and for classification of categorical data [19-20].

2.4. Sequential Minimal Optimization

The Sequential Minimal Optimization (SMO) classifier which uses analytical quadratic programming techniques, is capable of solving optimization problems quickly. The SMO classifier, which is basically based on the support vector machine classifier, was developed by Platt. The classifier constantly updates the support vector machine with iterative operations [21]. Being a powerful classifier, SMO can provide high performance in classification processes [22]. Accuracy, sensitivity, specificity, F score and ROC Area metrics were obtained to evaluate the performance of Stacking ensemble learning method with RF, NB and SMO classifiers. These metrics are obtained based on the information in Table II.

TABLE II
CONFUSION MATRIX

		Actual		
		Positive	**Negative**	**Total**
Predicted	**Positive**	(TP)True Positive	(FP) False Positive	TP+FP
	Negative	(FN) False Negative	(TN) True Negative	FN+TN
	Total	TP+FN	FP+TN	TP+FP+TN+FN

Sensitivity: True Positive/ True Positive+False Negative
Specificity: True Negative/True Negative+False Positive
Accuracy: TP+TN / TP+FP+TN+FN
F-Score: (2*TP)/ (2*TP+FP+FN)

3. RESULTS

In the study, outlier analyses were carried out within the data set. Outliers were detected in the data of 5 patients in the patient group. Data on 5 patients were excluded from the study. The data of 360 individuals including 244 coronary artery patient data and 116 non-coronary artery disease data were included in the study. 31.8% of the patients were women and 68.2% were men. The mean age of the patients is 58.35 ± 11.31.

The performances of NB, RF, SMO and Stacking ensemble learning methods were compared for classification of individuals with and without coronary artery disease. The performances of the classifiers were evaluated in terms of accuracy (ACC), sensitivity (SEN), specificity (SPE), F-Measure (F), and ROC Area (AUC) metrics. ACC = 0.774, SEN = 0.888, SPE = 0.719, F = 0.718 and AUC = 0.913 values were obtained with the Naive Bayes classifier. ACC = 0.883, SEN = 0.733, SPE = 0.955, F = 0.802 and AUC = 0.844 were obtained with the SMO classifier. ACC = 0.908, SEN = 0.853, SPE = 0.934, F = 0.857 and AUC = 0.957 were obtained with Random Forest classifier. ACC = 0.933, SEN = 0.905, SPE = 0.946, F = 0.897 and AUC = 0.959 values were obtained with the stacking ensemble learning method. The performances of the classifiers are shown in Table III.

TABLE III
PERFORMANCE METRICS OF METHODS

	Accuracy	Sensit.	Specif.	F-Measure	ROC Area
Naive Bayes	0.774	0.888	0.719	0.718	0.913
SMO	0.883	0.733	0.955	0.802	0.844
Random Forest	0.908	0.853	0.934	0.857	0.957
Stacking ELM	0.933	0.905	0.946	0.897	0.959

The SMO classifier provided the highest performance in terms of specificity metric. The Stacking ensemble learning method provided the highest performance in terms of accuracy, sensitivity, F-Measure, and ROC Area metrics. Comparison of the performance of classifiers is shown in Figure 1.

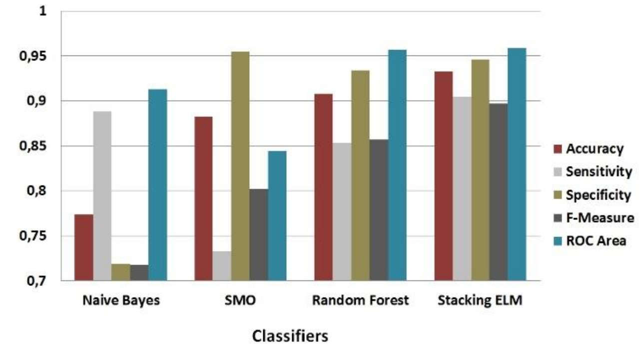


Fig. 1. Performance Metrics of Methods

The significance levels of the predictor variables according to the feature selection methods are shown in figure 2. It has been observed that the variables Bun, Mpv, Age, Ast and Monocyte respectively make the most substantial contribution in the classification of diseases.

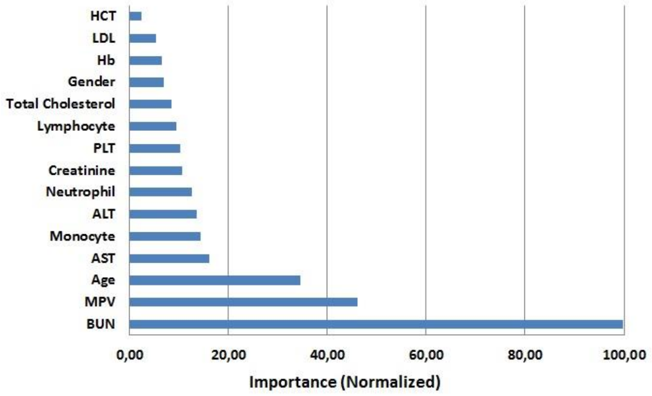


Fig. 2. Importance Value of Predictor Variables

4. DISCUSSION

Machine learning methods can provide successful results for diagnosing and classifying diseases. Early diagnosis of coronary artery diseases is vital in reducing the risk of the disease [23]. Solutions based on clinical applications and imaging methods are available for early diagnosis and classification of coronary artery diseases. However, more practical solutions can be preferred due to the cost and application difficulties of these methods [5]. Machine learning methods can play an effective role in the diagnosis and classification of coronary artery disease and provide practicality for diagnosis.

A great deal of research has been conducted on the early diagnosis of coronary artery diseases with machine learning methods. Different machine learning methods have been applied in many of these studies. Plaiwak (2018) proposed a model in which machine learning methods are used to determine heart disorders with ECG signals. Performance of SVM, KNN, PNN, and RBFNN classifiers were evaluated in the model. SVM classifier provided the highest performance with 90% accuracy [24]. Alizadehsani et al. (2016) applied different functions of some classifiers and SVM classifier in the detection of coronary artery disease in their study. They achieved accuracy performance in the range of approximately 75% to 91% for the 4 functions of SVM proposed in the model [25]. Babaoglu et al. (2010) aimed to increase the performance of SVM classifier in classifying coronary artery diseases with principal component analysis. It was determined that principal components analysis increased the SVM classification performance from 76.67% to 79.71% [26].

Naive Bayes method is one other classifier commonly used for the classification of coronary artery diseases. Gola et al. (2020) compared machine learning methods and different methods in their study to classify coronary artery disease. In their findings, they determined the AUC value of the Naive Bayes classifier as 0.819 [27]. Setiawan et al. (2014) compared the performances of Naive Bayes and J48 classifiers with different dimension reduction methods to increase the classification performance of coronary artery disease in their study [28]. Arjenaki et al. (2015) used the Naive Bayes method and SVM classifiers for the diagnosis of coronary artery diseases. The Naive Bayes classifier provided an accuracy rate of approximately 85% [29]. Random Forest is one of the classifiers that can provide high accuracy performance in diagnosing diseases. Researchers have used the random forest classifier to classify many diseases. Ani et al. (2016) obtained the accuracy performance of the random forest classifier as 89% to predict the risk factors of coronary artery diseases [30]. The random forest classifier have generally provided high accuracy performance in the classification of heart diseases [31-32].

In our study, the classification performances of Random Forest, SMO, Naive Bayes classifiers and Stacking ensemble learning methods were compared for the classification of coronary artery diseases. In our model, the Naive Bayes classifier has an accuracy performance of 0.774. The accuracy performance of the SMO classifier was obtained as 0.883. Finally, the accuracy performance of the Random Forest classifier was obtained as 0.908. The performances of the classifiers in our model are compatible with the findings of the studies in the literature. Stacking ensemble learning method has higher accuracy performance than classifiers. The accuracy value of the stacking ensemble learning method was obtained as 0.933. In our study, the stacking ensemble learning method achieved higher accuracy than the classifiers in the model for the classification of coronary artery diseases. Lo et al. (2016) compared the classification performances of 7 classifiers and ensemble learning methods for the diagnosis of coronary artery diseases in their study. They stated that the ensemble method provided the highest performance [33]. The findings we obtained in our study are consistent with the findings of the literature.

Ensemble learning methods can provide a successful classification for the classification of diseases. Ensemble learning methods can generally provide higher classification performance against single classifiers. The choice of meta classifier in the stacking ensemble learning method significantly affects the classification performance. Choosing the appropriate meta classifier for the model can improve the performance. The compliance of the classifiers included in the ensemble learning model can affect the classification performance. Increasing the number of samples included in the model can contribute to the increase in classification performance.

5. CONCLUSION

In this study, the performances of machine learning methods and ensemble learning method in the classification of coronary artery disease were compared. It was determined that the predictor variables that make the most essential contribution to the classification of coronary artery diseases are BUN, MPV, Age, AST and Monocyte, respectively. In the model, the classification performances of the SMO, Naive Bayes, Random forest classifiers and the Stacking ensemble learning method were compared. In the stacking ensemble learning method, the ensemble model consists of SMO, Naive Bayes and Random Forest classifiers. The meta classifier of the ensemble method is determined as "Logistic". Naive Bayes, SMO and Random forest classifiers achieved an accuracy performance of 0.774, 0.883 and 0.908, respectively. Stacking ensemble learning method successfully classified the diseases with an accuracy of 0.933. Stacking ensemble learning method classified coronary artery diseases with higher accuracy than the other classifiers. The stacking ensemble learning method can be successfully applied for the classification and diagnosis of coronary diseases.

ACKNOWLEDGMENT

The authors declare that they have no conflict of interest. There are no financial supports. Ethical approval no: Clinical research ethics committee of Kahramanmaras Sutcu Imam University, Ethics Committee Approval No: 2019/21 Decision No: 13.

REFERENCES

[1] Q.U.A. Mastoi, T.Y. Wah, R.G. Raj, U. Iqbal, “Automated Diagnosis of Coronary Artery Disease: A Review and Workflow”, Cardiology Research and Practice. Vol.2018, 2018, pp.1-9.

[2] M.G. Tsipouras, T.P. Exarchos, D.I. Fotiadis, A.P. Kotsia, K.V. Vakalis, K.K. Naka, et al, “Automated diagnosis of coronary artery disease based on data mining and fuzzy modeling”, IEEE Trans Inf Technol Biomed, Vol.12, No.4, 2008, pp.447-458.

[3] E.J. Benjamin, P. Muntner, A. Alonso, M.S. Bittencourt, C.W. Callaway, A.P. Carson, et al, “Heart Disease and Stroke Statistics-2019 Update: A Report From the American Heart Association”, Circulation, Vol.139, No.10, 2019, pp.E56-E528.

[4] R. Alizadehsani, M. Roshanzamir, M. Abdar, A. Beykikhoshk, A. Khosravi, M. Panahiazar, et al, “A database for using machine learning and data mining techniques for coronary artery disease diagnosis”, Scientific data, Vol.6, No.1, 2019, pp.1-13.

[5] R. Alizadehsani, M. Abdar, M. Roshanzamir, A. Khosravi, P.M. Kebria, F. Khozeimeh, et al. “Machine learning-based coronary artery disease diagnosis: A comprehensive review”. Computers in Biology and Medicine, Vol.111, 2019. pp.1-14.

[6] A.M. Poss, J.A. Maschek, J.E. Cox, B.J. Hauner, P.N. Hopkins, S.C. Hunt, et al, “Machine learning reveals serum sphingolipids as cholesterol-independent biomarkers of coronary artery disease”, Journal of clinical investigation, Vol.130, No.3, 2020, pp.1363-1376.

[7] L. Verma, S. Srivastava, P.C. Negi, “A Hybrid Data Mining Model to Predict Coronary Artery Disease Cases Using Non-Invasive Clinical Data”, Journal of medical systems. Vol.40, No.7, 178, 2016, pp.1-7.

[8] Y. Zheng, M. Loziczonek, B. Georgescu, S.K. Zhou, F. Vega-Higuera, D. Comaniciu, Machine learning based vesselness measurement for coronary artery segmentation in cardiac CT volumes. In: Medical Imaging 2011, Image Processing. 2011.

[9] S. Nikan, F. Gwadry-Sridhar, M. Bauer, Machine learning application to predict the risk of coronary artery atherosclerosis. In: Proceedings - 2016 International Conference on Computational Science and Computational Intelligence, CSCI 2016, pp.34-39.

[10] U.R. Acharya, H. Fujita, O.S. Lih, M. Adam, J.H. Tan, C.K. Chua, “Automated detection of coronary artery disease using different durations of ECG segments with convolutional neural network”, Knowledge-Based Systems, Vol.132, 2017, pp.62-71.

[11] E.M. Karabulut, T. İbrikçi, “Effective diagnosis of coronary artery disease using the rotation forest ensemble method”, Journal of medical systems. Vol.36, No.5, 2012, pp.3011-3018.

[12] M. Amer, M. Goldstein, “Nearest-Neighbor and Clustering based Anomaly Detection Algorithms for RapidMiner”, Proc 3rd RapidMiner Community Meet Conferernce (RCOMM 2012). 2012.

[13] Z.H. Zhou, Ensemble methods: Foundations and algorithms. Ensemble Methods: Foundations and Algorithms, CRC Press, 2012.

[14] D.H. Wolpert, “Stacked generalization”, Neural Networks, Vol.5, No.2, 1992, pp.241-259.

[15] L. Breiman, “Random forests”, Machine Learning”, Vol.45, No.1, 2001, pp.5-32.

[16] Ö. Akar, O. Güngör “Rastgele orman algoritması kullanılarak çok bantlı görüntülerin sınıflandırılması”, Jeodezi ve Jeoinformasyon Dergisi. Vol.146, 2012, pp.139-146.

[17] K.J. Archer, R. V. “Kimes Empirical characterization of random forest variable importance measures”, Computer Statistics&Data Analysis, Vol.52, No.4, 2008, pp.2249-2260.

[18] I. Rish, “An empirical study of the naive Bayes classifier”, In IJCAI 2001 workshop on empirical methods in artificial intelligence, Vol.3, No.22, 2001, pp.41-46.

[19] P. Bhargavi, S. Jyothi, “Applying Naive Bayes Data Mining Technique for Classification of Agricultural Land Soils”, IJCSNS International journal of computer science and network security, Vol.9, No.8, 2009, pp.117-122.

[20] K.P. Murphy, “Naive Bayes classifiers Generative classifiers”, Bernoulli. 2006.

[21] J. Platt, Fast Training of Support Vector Machines using Sequential Minimal Optimization, In: Advances in Kernel Methods Support Vector Learning, MIT Press, 1999.

[22] S.S. Keerthi, S.K. Shevade, C. Bhattacharyya, K.R.K. Murthy, “Improvements to Platt’s SMO algorithm for SVM classifier design”, Neural computation, Vol.13, No.3, 2001, pp.637-649

[23] D.P. Zipes, P. Libby, R.O. Bonow, D.L. Mann, G. F. Tomaselli, Braunwald's Heart Disease E-Book: A Textbook of Cardiovascular Medicine. Elsevier Health Sciences, 2018.

[24] P. Pławiak, “Novel methodology of cardiac health recognition based on ECG signals and evolutionary-neural system”, Expert Systems with Applications, Vol.92, 2018, pp.334-349.

[25] R. Alizadehsani, M.H. Zangooei, M.J. Hosseini, J. Habibi, A. Khosravi, M. Roshanzamir, et al, “Coronary artery disease detection using computational intelligence methods, Knowledge-Based Systems, Vol.109, 2016, pp.187-197.

[26] I. Babaoğlu, O. Findik ,M. Bayrak, “Effects of principle component analysis on assessment of coronary artery diseases using support vector machine”, Expert Systems with Applications, Vol.37, No.3, 2010, pp.2182-2185.

[27] D. Gola, J. Erdmann,B. Müller-Myhsok, H. Schunkert, I.R. König, “Polygenic risk scores outperform machine learning methods in predicting coronary artery disease status”, Genetic Epidemiology, Vol.44, No.2, 2020, pp.125-138.

[28] N.A. Setiawan, D.W. Prabowo, H.A. Nugroho, *Benchmarking of feature selection techniques for coronary artery disease diagnosis. In: Proceedings - 2014 6th International Conference on Information Technology and Electrical Engineering: Leveraging Research and Technology Through University-Industry Collaboration*, ICITEE 2014, 2014.

[29] H. G. Arjenaki, M. H. N. Shahraki, N. Nourafza, “A low cost model for diagnosing coronary artery disease based on effective feature”, International Journal of Electronics Communication and Computer Engineering, Vol.6, No.1, 2015, pp.93-97.

[30] R. Ani, A. Augustine, N.C. Akhil, O.S. Deepa “Random forest ensemble classifier to predict the coronary heart disease using risk factors”, In Proceedings of the International Conference on Soft Computing Systems, 2016, Springer, New Delhi. pp.701-710.

[31] M.A. Jabbar, B.L. Deekshatulu, P. Chandra, Prediction of heart disease using random forest and feature subset selection. In: Innovations in bio-inspired computing and applications 2016, Springer, Cham, pp. 187-196.

[32] Z. Masetic, A. Subasi. Congestive heart failure detection using random forest classifier. Computer methods and programs in biomedicine, Vol.130, 2016, pp.54-64.

[33] Lo YT, Fujita H, Pai TW. Prediction of coronary artery disease based on ensemble learning approaches and co-expressed observations. Journal of Mechanics in Medicine and Biology, Vol.16, No.1 2016.

BIOGRAPHIES

Adem Doganer obtained his BSc. degree in Statistics from Mugla University in 2008. He received MSc. degree in Statistical Information Systems from the Fırat University in 2011, and Ph.D. degrees in Statistical Information Systems from the Fırat University in 2015. Adem Doganer has a second Ph.D degree on Biostatistics and Medical Informatics from the Inonu University in 2020. He is currently working as an assistant professor the Department of Biostatistics and Medical Informatics in Kahramanmaras Sutcu Imam University His research interests are, Statistical analysis, information systems, cognitive systems, bionformatics, data mining, Modeling and simulation, machine learning, artificial intelligence.

Mehmet Kirisci obtained his BSc. degree in Medicine from Kahramanmaras Sutcu Imam University in 2004. He received Specialist in Medicine degree from the Gazi University in 2012. He is currently working as an assistant professor the Department of Cardiovascular Surgery at Kahramanmaras Sutcu Imam University. His research interests are cardiovascular surgery, ischemia reperfusion and coronary artery diseases.

COMPARISON OF DIFFERENT DECISION TREE MODELS IN CLASSIFICATION OF ANGINA PECTORIS DISEASE

I. Balikci Cicek, Z. Kucukakcali, and E. Guldogan

Abstract— *Aim:* The aim of this study is to classify Angina pectoris disease and compare the estimates of the methods by applying J48 and Random Forest methods, which are among the decision tree models, on the open access angina pectoris data set.

Materials and Methods: In the study, the data set named "Project Angina Data Set" was obtained from https://www.kaggle.com/snehal1409/predict-angina. In the data set, there are a total of 200 patients in whom angina pectoris was evaluated. Decision tree models J48 and Random Forest methods were used to classify angina pectoris disease.

Results: From the applied models, from the performance values obtained from the J48 method, the accuracy was 0.868, balanced accuracy 0.868, sensitivity 0.895, specificity 0.842, positive predictive value 0.85, negative predictive value 0.889 and F1-score 0.872. From the performance values obtained from the Random Forest method, the accuracy was 0.921, balanced accuracy 0.921, sensitivity 0.895, selectivity 0.947, positive predictive value 0.944, negative predictive value 0.9 and F1-score 0.919.

Conclusion: Considering the findings obtained from this study, it has been shown that the decision tree models used give successful predictions in the classification of angina pectoris disease.

Keywords— Classification, decision trees, J48, Random Forest, angina pectoris.

1. INTRODUCTION

CARDIOVASCULAR diseases rank first among the causes of death in developed and developing countries [1]. In a study in which death events in our country between 2009-2016 were examined epidemiologically, it was reported that cardiovascular diseases took the first place among the causes of death in all years [2]. Among the cardiovascular diseases, deaths due to ischemic heart diseases take the first place [3]. The most important symptom of ischemic heart disease is angina pectoris (AP). Angina pectoris is described as a clinical symptom characterized by discomfort or pain in the chest, jaw, shoulder, back and arm [4]. In AP physiopathology, an increase in the oxygen demand of the myocardium at the cellular level or a decrease in the oxygen level presented in the myocardium is the cause of angina.

Although decreased oxygen delivery is often found responsible as a result of narrowness in the coronary arteries, abnormal increases in oxygen demand such as increased heart rate, uncontrolled hypertension, and increased myocardial contractility can also lead to angina [5]. Although it increases with age, angina pectoris occurs between 0.1-20% in the general population between the ages of 45-74. Angina pectoris is thought to be present in 20,000-40,000 people per million populations, especially in most European countries [5]. Decision trees are one of the most used methods in classification problems. Decision trees are easier to construct, understand and interpret compared to other methods. In addition to these, another advantage of decision trees is that they produce successful models. In order to classify in the decision trees method, a tree is created from the data we have and the records in the dataset are applied to this tree, and the classification process of the records takes place according to the result. In other words, a data that we do not know which class belongs to according to the decision trees obtained from the database, is predicted according to the rule set created when we come to it [6]. J48 is a decision tree algorithm based on the very popular C4.5 algorithm developed by J. Ross Quinlan [7]. J48 Algorithm; Based on Information Gain Theory, it has automatic process capability to select relevant properties from data. It is an iterative algorithm that divides samples from where the information gain is best. The J48 can do an effective pruning to cut off not meaningful, in other words, weak branches [8]. The Random Forest (RF) method was proposed by Breiman in 2001 by developing the Bagging method, which envisions combining the decisions of multiple, multivariate trees each trained with different training sets, instead of generating a single decision tree. This method uses the bootstrap technique in the process of creating different sub-training sets and uses random feature selection during the development of trees [9].

The aim of this study is to compare the classification success of angina pectoris by applying J48 and Random Forest methods, which are among the different decision tree methods, on the angina pectoris dataset.

2. MATERIAL AND METHODS

2.1. Dataset

In the study, the dataset named “Project Angina Data Set” was obtained from https://www.kaggle.com/snehal1409/predict-angina in order to examine the working principle of J48 and Random Forest methods [10]. There are 100 (50.0%) no, 100 (50.0%) yes total 200 patients in the dataset used. The variables in the dataset and their descriptive characteristics of the variables are given in Table I.

İpek BALIKCI CICEK, Inonu University Department of Biostatistics and Medical Informatics, Faculty of Medicine, Malatya, Turkey, (ipek.balikci@inonu.edu.tr)

Zeynep KUCUKAKCALI, Inonu University Department of Biostatistics and Medical Informatics, Faculty of Medicine, Malatya, Turkey, (zeynep.tunc@inonu.edu.tr)

🖂 **Emek GULDOGAN,** Inonu University Department of Biostatistics and Medical Informatics, Faculty of Medicine, Malatya, Turkey, (emek.guldogan@inonu.edu.tr)

Manuscript received Sep 10, 2020; accepted Oct 18, 2020.
Digital Object Identifier:

TABLE I
VARIABLES IN PROJECT ANGINA DATASET AND DESCRIPTIVE PROPERTIES OF THESE VARIABLES

Variable	Variable Description	Variable Type	Variable Role
Status	Whether the woman has angina (no/yes)	Qualitative	Dependent/ Target
Age	Age	Quantitative	Independent/ Predictor
Smoke	Smoking status (current, ex-, non-smoker)	Qualitative	Independent/ Predictor
Cig	Average number of cigarettes smoked per day	Quantitative	Independent/ Predictor
Hyper	Hypertensive condition (absent, mild, moderate)	Qualitative	Independent/ Predictor
Angfam	Family history of angina (no/yes)	Qualitative	Independent/ Predictor
Myofam	Family history of myocardial infarction (no/yes)	Qualitative	Independent/ Predictor
Strokefam	A family history of stroke (no/yes)	Qualitative	Independent/ Predictor
Diabetes	Whether the woman has diabetes (no/yes)	Qualitative	Independent/ Predictor

3. DECISION TREES

Decision trees, one of the prediction methods, is one of the popular and powerful methods of information discovery and data mining. Decision trees are a hierarchical and ordered way of displaying the rules in the data. Decision trees are a visual modeling method that shows the mass of information about the problem faced by the decision maker in a more understandable way, and presents the decision options and probabilistic situations in a certain order by sorting. In this context, it can be said that decision trees represent a hierarchical model that includes decisions and results. Thanks to its easy-to-understand graphical structure and rules, it is widely used in many areas [11]. Decision trees model, which is among the classification models in data mining, is a model with predictive value. Decision trees ask questions starting from the first stage to the final decision options and form their structure with the answers they receive to these questions, and rules (if-then rule) can be written with this tree structure [12].

3.1. J48

J48 developed by *Quinalan* is a C4.5 decision tree developed for the classification process of nonlinear and small size data. J48 is a decision tree that uses entropy concept knowledge to classify. It applies Quinlan's C4.5 algorithm to generate a pruned C4.5 tree. Decision making is done by dividing each attribute dataset into subsets to examine entropy differences. The highest normalized information gains, the attributes are selected [13].

3.2. Random Forest

In this algorithm, developed by *Breiman* in 2001, the purpose for the classifier is to combine the decisions of multiple trees, each trained in different training sets, rather than generating a single decision tree. Random feature selection with the same distribution is used for different training sets. While creating decision trees, when determining the attribute at each level, firstly, some calculations are made in all trees and the attribute is determined, then the attributes in other trees are combined and the most used attribute is selected. After the selected attribute is included in the tree, the same processes are repeated at other levels. To start the algorithm, the number of variables used in each node and the number of trees to be developed must be determined by the user. Random Forest uses the CART (Classification and Regression Tree) algorithm to generate a tree. Nodes and branches are created in accordance with the features of this algorithm [14].

3.3. Performance Evaluation of Models

Performance criteria obtained by using the classification matrix given below were used in the performance evaluation of J48 and Random Forest methods.

TABLE II
CLASSIFICATION MATRIX FOR CALCULATING PERFORMANCE CRITERIA

		Real		
		Positive	**Negative**	**Total**
Predicted	**Positive**	True positive (TP)	False negative (FN)	TP+FN
	Negative	False positive (FP)	True negative (TN)	FP+TN
	Total	TP+FP	FN+TN	TP+TN+FP+FN

The performance criteria to be used in the performance evaluation of the models in this study are given below.

$\text{Accuracy} = (TP+TN)/(TP+TN+FP+FN)$

$\text{Balanced accuracy} = [[TP/(TP+FP))]+[TN/(TN+FN)]]/2$

$\text{Sensitivity} = TP/(TP+FP)$

$\text{Specificity} = TN/(TN+FN)$

$\text{Positive predictive value} = TP/(TP+FN)$

$\text{Negative predictive value} = TN/(TN+FP)$

$\text{F1-score} = (2*TP)/(2*TP+FP+FN)$

4. DATA ANALYSIS

Quantitative data are expressed as mean ± standard deviation, median (minimum-maximum), and qualitative data as number (percentage). Conformity to normal distribution was evaluated by the Kolmogorov-Smirnov test. In terms of independent variables, whether there is a statistically significant difference between the "no" and "yes" groups, which are the categories of the dependent / target variable (status), and whether there is a relationship, Mann-Whitney U test, Pearson chi-square test, Continuity Correction test and Fisher's Exact test. It was examined using the chi-square test values of $p<0.05$ were considered statistically significant. IBM SPSS Statistics 26.0 package program was used for all analyzes.

For the validity of the model, a 10-fold cross-validation method was used. In the 10-fold cross-validation method, all data is divided into 10 equal parts. One part is used as a test set and the remaining 9 parts are used as a training dataset and this process is repeated 10 times.

5. RESULTS

Descriptive statistics of independent variables examined in this study are given in Table 3. According to the findings in Table 3; there is a statistically significant difference between the dependent / target variable groups in terms of age and cig variables (p<0.05).

According to the findings in Table 4; there is a statistically significant relationship between the smoke, hyper and myofam variables and the dependent / target variable (status) groups (p<0.05).

TABLE III
DESCRIPTIVE STATISTICS FOR QUANTITATIVE INDEPENDENT VARIABLES

Variables	Status		**p-value***
	no	yes	
	Median(min-max)	**Median(min-max)**	
age	49 (29-74)	57 (29-73)	**<0.001**
cig	0 (0-30)	12 (0-40)	**<0.001**

***: Mann Whitney U test**

TABLE IV
DESCRIPTIVE STATISTICS FOR QUALITATIVE INDEPENDENT VARIABLES

Variables		Status		**p-value**
		no	yes	
		Number (%)	**Number (%)**	
smoke	current	22 (22)	61 (61)	**<0.001***
	ex	14 (14)	26 (26)	
	non-smoker	64 (64)	13 (13)	
hyper	absent	83(83)	67 (67)	**0.022***
	mild	14 (14)	23 (23)	
	moderate	3 (3)	10 (10)	
angfam	no	94 (94)	85 (85)	0.065**
	yes	6 (6)	15 (15)	
myofam	no	88 (88)	47 (47)	**<0.001****
	yes	12 (12)	53 (53	
strokefam	no	94 (94)	94 (94)	1**
	yes	6 (6)	6 (6)	
diabetes	no	97 (97. 97)	94 (94.9)	0.445***
	yes	2 (2.02)	5 (5.1)	

***: Pearson chi-square test, **: Continuity Correction test,**
*****: Fisher's Exact test**

Classification matrix of J48 and Random Forest models are given in Table 5 and Table 6, respectively.

TABLE V
CLASSIFICATION MATRIX OF THE J48 MODEL

Prediction	**Reference**		
	no	yes	**Total**
no	17	3	20
yes	2	16	18
Total	19	19	38

TABLE VI
CLASSIFICATION MATRIX OF THE RANDOM FOREST MODEL

Prediction	**Reference**		
	no	yes	**Total**
no	17	1	18
yes	2	18	20
Total	19	19	38

Table 7, the values of performance criteria calculated from models created to classify Angina pectoris disease in the test stage are given below.

TABLE VII
PERFORMANCE CRITERIA VALUES CALCULATED FROM CREATED MODELS IN THE TESTING PHASE

Model / **Performance Metrics**	**J48**	**Random Forest**
	Value	**Value**
Accuracy (%)	86.8	92.1
Balanced accuracy (%)	86.8	92.1
Specificity (%)	84.2	94.7
Sensitivity (%)	89.5	89.5
Positive predictive value (%)	85.0	94.4
Negative predictive value (%)	88.9	90.0
F1-score (%)	87.2	91.9

In Figure 1, the values of the performance criteria obtained from J48 and Random Forest methods are plotted.

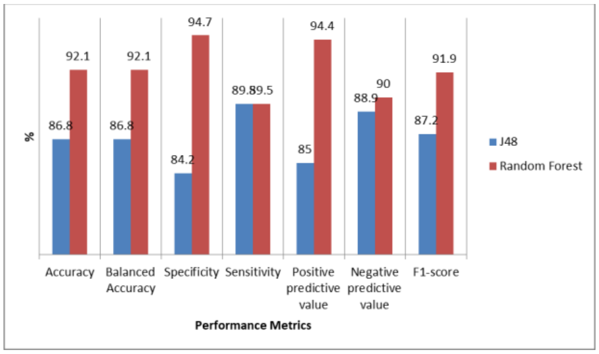


Fig.1. Performance metric values obtained from J48 and Random Forest methods in the testing stage

6. DISCUSSION

Angina pectoris (AP) was first described by Heberden in 1772, and the most common cause is coronary atherosclerosis [15]. Discomfort in the chest caused by ischemic atherosclerotic coronary artery disease associated with impaired coronary blood flow is called angina pectoris. It mostly spreads to the neck, left shoulder, and medial side of the left arm and lasts no longer than 10-15 minutes. Angina pectoris responds immediately to short-acting nitrates [16, 17]. The prevalence of angina in both sexes increases rapidly with age. From 0.1-1% in women aged 45-54, to 10-15% in women aged 65-74; It increases in 2-5% for men aged 45-54 and to 10-20% for men aged 65-74. Accordingly, it can be calculated that in most European countries, 20 000-40 000 people per million of the general population have angina [18]. Decision trees are a method that is frequently used in classification because it is easier to structure and understand when compared to other classification methods [19]. Decision trees are the most widely used methods among classification models because of their easy interpretation, easy integration with database systems, and good reliability. These methods have predictive and descriptive properties [20]. Decision trees decide which class the new data belongs to, based on the old data, by subtracting rules. The decision tree acts in line with the questions asked and the answers received, and creates rules by combining the answers to the questions. We can say that the tree formed is a set of rules consisting of many "if-then" [21]. When creating decision trees, it is important what algorithm is used. Because the shape of the tree created according to the algorithm used can change. Different tree structures give different classification results. The fact that the first node forming the root node is different will change the way to be followed when reaching the farthest leaf, thus the classification [22]. There are many different decision tree algorithms. Of these, J48 is information-based and has automatic process capability to select relevant features from data. In addition, it is the algorithm with the highest classification success according to algorithms such as Naive Bayes, ID3, Logistic Regression [23]. Random Forest is an algorithm that aims to increase the classification value by generating more than one decision tree during the classification process [24].

According to the findings in this study; the performance criteria obtained from the J48 method, accuracy was 0.868, balanced accuracy 0.868, sensitivity 0.895, specificity 0.842, positive predictive value 0.85, negative predictive value 0.889, and F1-score 0.872. The performance criteria obtained from another method, the Random Forest method, the accuracy was 0.921, balanced accuracy 0.921, sensitivity 0.895, specificity 0.947, positive predictive value 0.944, negative predictive value 0.9 and F1-score 0.919. When these classification performances were compared, Random Forest method gave more successful estimation results compared to J48 method.

As a result, the decision trees methods used have produced quite successful results in the study with the angina pectoris dataset.

REFERENCES

[1] A. K. Şimşek and Ş. E. Alpar, "Akut Koroner Sendromlu Hastalarda Sağlıklı Yaşam Davranışlarının Kazandırılması," Turk J Cardiovasc Nurs, vol.11, pp. 31-36, 2020.

[2] K.Y. EmIk and A. E. Önal, "2009-2016 Yıllarında Türkiye'deki Ölümlerin Epidemiyolojik Yönden İncelenmesi ve Ölüm Bildirim Sisteminin Önemi," İstanbul Tıp Fak. Dergisi, vol.82, pp.25-26, 2019.

[3] B. Bayrak, S. Oğuz, S. Arslan, B. Candar, S. Keleş, B. Karagöz, et al., "Miyokard İnfarktüsü Geçirmiş Hastalarda Algılanan Stresin Belirlenmesi," Turk J Cardiovasc Nurs, vol.10, pp.129-137, 2019.

[4] G. A. Diamond, "A clinically relevant classification of chest discomfort," Journal of the American College of Cardiology, vol.1, pp.574-575, 1983.

[5] K. Soylu, "Kararlı angina pektoris," Deneysel ve Klinik Tıp Dergisi, vol. 29, pp. 117-121, 2012.

[6] G. Silahtaroğlu, Veri madenciliği: Kavram ve algoritmaları: Papatya, 2013.

[7] H. Nizam and S. S. Akın, "Sosyal medyada makine öğrenmesi ile duygu analizinde dengeli ve dengesiz veri setlerinin performanslarının karşılaştırılması," XIX. Türkiye'de İnternet Konferansı, 2014.

[8] N. Bhargava, G. Sharma, R. Bhargava, and M. Mathuria, "Decision tree analysis on j48 algorithm for data mining," Proceedings of International Journal of Advanced Research in Computer Science and Software Engineering, vol. 3, 2013..

[9] L. Breiman, "Random forests," Machine learning, vol. 45, pp. 5-32, 2001.

[10] Available: https://www.kaggle.com/snehal1409/predict-angina

[11] S. K. Murthy, "Automatic construction of decision trees from data: A multi-disciplinary survey," Data mining and knowledge discovery, vol.2, pp.345-389, 1998.

[12] R. Lior, Data mining with decision trees: theory and applications, World scientific, vol. 81, 2014.

[13] G.I. Salama, M. Abdelhalim, and M.A.-e. Zeid, "Experimental comparison of classifiers for breast cancer diagnosis," in 2012 Seventh International Conference on Computer Engineering & Systems (ICCES), 2012, pp.180-185.

[14] B. Daş and İ. Türkoğlu, "DNA dizilimlerinin sınıflandırılmasında karar ağacı algoritmalarının karşılaştırılması," Elektrik-Elektronik-Bilgisayar ve Biyomedikal Müh. Semp. (ELECO 2014), pp.381-383, 2014.

[15] C. Sliub, "Angina pectoris and coronary heart disease. RO Brandenburg (ed.) Cardiology: Fundamentals and Practice. New York," Year Book Medical Publ, 1987.

[16] C. Man, Z. Dai, and Y. Fan, "Dazhu hongjingtian preparation as adjuvant therapy for unstable angina pectoris: A meta-analysis of randomized controlled trials," Frontiers in Pharmacology, vol.11, p.213, 2020.

[17] T. F. Imran, R. Malapero, A. H. Qavi, Z. Hasan, B. de la Torre, Y.R. Patel, et al., "Efficacy of spinal cord stimulation as an adjunct therapy for chronic refractory angina pectoris," International journal of cardiology, vol.227, pp.535-542, 2017.

[18] D.K. Hancı, "Perkütan Koroner Girişim Yapılan Hastalarda Stent Balonu Dilatasyonu Sırasında Angina Gelişme Sıklığı, Yeri, Karakteri ve Angina Lokalizasyonunun Koroner Anatomi İle İlişkisi," Kardiyoloji Uzmanlık Tezi, Kocaeli Üniversitesi, Tıp Fakültesi, 2020.

[19] E. Atılgan, "Karayollarında meydana gelen trafik kazalarının karar ağaçları ve birliktelik analizi ile incelenmesi," Hacettepe Üniversitesi Istatistik Anabilimdalı,Yüksek Lisans Tezi, 2011.

[20] J. Han, J. Pei, and M. Kamber, Data mining: concepts and techniques: Elsevier, 2011.

[21] M. Akman, "Veri madenciliğine genel bakış ve random forests yönteminin incelenmesi: sağlık alanında bir uygulama," master's thesis, Ankara University, Ankara, 2010.

[22] G. Silahtaroğlu, "Veri madenciliği," Papatya Yayınları, İstanbul, 2008.

[23] A. Altınkardeş, H. Erdal, F. Baba, and A.S. Fak, "ABPM Ölçümü Olmaksızın Karar Ağaçları Algoritması ile Non-Dipper/Dipper Öngörüsü," IX. Ulusal Tıp Bilişimi Kongresi, Antalya, Türkiye, 15-17 Kasım 2012.

[24] Ö. Akar and O. Güngör, "Classification of multispectral images using Random Forest algorithm," Journal of Geodesy and Geoinformation, vol. 1, pp. 105-112, 2012.

BIOGRAPHIES

İpek BALIKÇI ÇİÇEK obtained her BSc. degree in mathematics from Çukurova University in 2010. She received MSc. degree in biostatistics and medical informatics from the Inonu University in 2018. She currently continues Ph.D. degrees in biostatistics and medical informatics from the Inonu University. In 2014, she joined the Department of Biostatistics and Medical Informatics at Inonu University as a researcher assistant. Her research interests are cognitive systems, data mining, machine learning, deep learning.

Zeynep KUCUKAKCALI obtained her BSc. degree in mathematics from Çukurova University in 2010. She received MSc. degree in biostatistics and medical informatics from the Inonu University in 2018. She currently continues Ph.D. degrees in biostatistics and medical informatics from the Inonu University. In 2014, she joined the Department of Biostatistics and Medical Informatics at Inonu University as a researcher assistant. Her research interests are cognitive systems, data mining, machine learning, deep learning.

Emek GÜLDOĞAN obtained his BSc. degree in Computer Engineering from Middle East Technical University in 2001. He received MSc. degree in biostatistics and medical informatics from the Inonu University in 2005, and Ph.D. degrees in biostatistics and medical informatics from the Inonu University in 2017. He is currently working as an assistant professor of the Department of Biostatistics and Medical Informatics at Inonu University and as the information processing manager at Turgut Özal Medical Center. His research interests are cognitive systems, data mining, machine learning, deep learning.

ARTIFICIAL NEURAL NETWORKS BASED-PREDICTION OF AUTISM SPECTRUM DISORDER

I. Ucuz, and A. Uzun Cicek

Abstract— *Aim:* Autism Spectrum Disorders (ASD) is one of the important neurodevelopmental disorders. This study aimed to perform artificial-intelligence-based modelling based on the prenatal-perinatal factors, family history, and developmental characteristics, which are emphasized as risk factors for ASD in the literature. *Materials and Methods:* The study was designed with retrospective management and data from 136 children with ASD and 143 healthy children were included. *Results:* According to the findings of the MLP model, the five most important factors were the mean age of first words (months), the mean age of head control (months), the mean age of sitting without support (months), history of autism in the family, and the mean paternal age at pregnancy (years), respectively. Overall percentages of the training and testing samples were 91.4% and 88.0%. AUC for the model was 0.922 for the separation of the autism and control groups. *Conclusion:* The proposed model is able to successfully differentiate patients with autism spectrum disorders from healthy individuals and identify factors associated with the disease.

Keywords— Artificial neural networks, autism, prenatal risk factors, perinatal risk factors

1. INTRODUCTION

AUTISM Spectrum Disorders (ASD) is one of the important neurodevelopmental disorders that manifest itself with deficits in social communication-interaction and limited-repetitive behavioural patterns, whose symptoms start at the early developmental period and cause clinically significant impairment in social or other important areas of functioning [1]. Although symptoms can be seen even in the 6-12 month period, the diagnosis age is usually around 24 months [2]. Studies on the prevalence of ASD have indicated that its prevalence varies according to the countries and the diagnostic methods used; however, the main emphasis is that the incidence is gradually increasing [3]. In general, the prevalence in developed countries is reported to be 1-5% [4]. A study comparing the prevalence of ASD in China and western countries found a 1% prevalence rate [5].

ASD is an important disorder that has many burdens for both the individual and society [1]. Given that the dramatic increase in its prevalence rates, it is important to know the situations that pose a risk for ASD. Several factors have been suggested in the etiology. In the meta-analysis study conducted by *Tick et al.*, it was emphasized that ASD is hereditary at a rate of 64-91%, and environmental factors, as well as genetic features, have an important effect on the increase in ASD prevalence [6]. Among environmental factors, prenatal characteristics are shown as one of the most important factors emphasized in ASD etiology. A meta-analysis study by *Gardener et al.* has determined that advanced parental age at pregnancy, mother's use of medication in the prenatal period, presence of gestational diabetes, history of bleeding in pregnancy, being the first-child are the risk factors associated with ASD [7].

Artificial neural networks (ANNs) are computing systems created on the computer to automatically perform data samples regarding the abilities and features of the human brain, such as generating new information through learning and discovering new information [8]. Artificial intelligence-based systems have found wide use in many fields of medical sciences as well as other fields [8-13]. This study aimed to perform artificial-intelligence-based modelling based on the prenatal-perinatal factors, family history, and developmental characteristics, which are emphasized as risk factors for ASD in the literature. In this way, it is purposed to determine the contribution of possible factors on ASD development by addressing many risk factors and developmental variables together.

2. MATERIAL AND METHODS

2.1. Study Procedure

This study was designed retrospectively and conducted on the data of 136 patients diagnosed with ASD (case group) at the Department of Child and Adolescent Psychiatry in Inonu University Faculty of Medicine between October 2015-September 2020, and the data of 143 healthy children (control group) matched with the case group for age and gender. ASD was diagnosed according to the Diagnostic and Statistical Manual of Mental Disorders (DSM-V) criteria [14]. Children with known chronic medical illness(es) and non-psychiatric drug use were excluded from the study.

Developmental characteristics, prenatal-perinatal period risk factors, and family history of children included in both the case and control groups were used as input parameters in artificial intelligence modelling. The developmental characteristics considered in the study are the age of onset of the first words (months), age of providing head control (months), age of starting to walk (months), and age of starting to sit without support (months). Variables of the maternal and paternal age at pregnancy, family history of autism, number and order of siblings, socioeconomic level of the family, maternal and paternal education level, place of residence, and parental

✉ **İlknur UCUZ,** Inonu University Faculty of Medicine, Department of Child and Adolescent Psychiatry, Malatya, Turkey, (ilknur.ucuz@inonu.edu.tr)

Ayla UZUN CİCEK, Cumhuriyet University, Faculty of Medicine, Department of Child and Adolescent Psychiatry, Sivas, Turkey, (aylauzun@cumhuriyet.edu.tr)

Manuscript received Sep 9, 2020; accepted Oct 30, 2020.
Digital Object Identifier:

consanguinity were employed as family history. The presence of the mother's history of regular drug use, excessive tea and coffee consumption, smoking, hypertension, gestational diabetes, an endocrine disease other than diabetes mellitus, radiation exposure, threatened miscarriage, severe infection requiring inpatient treatment such as respiratory tract or urogenital during pregnancy, and vitamin supplements status and nutritional problems in pregnancy were included in the category of prenatal risk factors. Variables including the month and season of the child's birth, birth time (weeks), birth weight, birth type, birthplace, neonatal jaundice, history of blood incompatibility and blood transfusion, birth complication(s) and seizure, and nulliparity were assigned to the model as perinatal factors. Artificial intelligence-based modelling was made for the diagnosis of ASD by giving the selected variables as input parameters to the system.

3. DATA ANALYSIS

IBM SPSS Statistics 25.0 program was used for data analysis. Qualitative data were expressed as number (n) and percentage (%), quantitative data were given as mean ± standard deviation (SD). Normality was analyzed using the one-sample Kolmogorov-Smirnov test. Statistical comparisons were performed with the chi-square test, Mann Whitney U test, independent t-test as appropriate. A p-value of less than 0.05 was accepted as statistically significant.

4. MODELLING AND PERFORMANCE EVALUATION

Nearly 70 percent and 30 percent of the entire dataset were used for training and testing stages in the development of MLP ANN models, respectively. For the model, the rescaling procedure for quantitative variables was standardized, the number of the hidden layer was 1, the number of units in the hidden layer was 6, the hidden layer activation function was a hyperbolic tangent, the number of units in the output layer was 2, the activation function of the output layer was softmax and error function was cross-entropy. The gradient descent approach optimized the models' hyperparameters. The performance of the model was evaluated with the accuracy, cross-entropy error, and area under the receiver operating curve metrics.

5. RESULTS

The mean age of the children was 45.9±11.1 (months) in the ASD group and 46.5±11.3 (months) in the control group (p=0.677). Of the participants, 99 (72.8%) children in the ASD group, and 108 (75.5%) children in the control group were male (p=0.602). While the mean maternal age at pregnancy was 29.1±7.4 (years) in the ASD group and 28.9±5.6 (years) in the control group (p=0.752), the mean paternal age at pregnancy was 32.7±7.8 (years) in the ASD group 31.4 ± 5.7 (years) (p= 0.324). The mean age of providing head control of children in the ASD group was 3.8±0.9 (months), the mean age of sitting unsupported was 9.0±1.7 (months), the mean age of walking was 14.4±2.4 (months), and the mean age of the first words was 22.8±7.6 (months). In the control group, these values were 2.6±0.8 (months) for head control, 7.9±1.4 (months) for sitting unsupported, 12.7±1.6 (months) for walking, 15.6±3.3 (months) for first words, respectively (all p-values <0.001). While children with ASD were born mostly in July (15.4%) and summer (39.7%), children in the control group were born mostly in May (14.7%) and spring (28.0%) (p=0.060 and 0.023, for the month at birth and the birth season, respectively).

Sociodemographic characteristics such as place of residence, parental education level, family income level, number and order of siblings, parental consanguinity, and some variables in pregnancy including threatened miscarriage, hypertension, presence of endocrine disease other than diabetes mellitus, regular drug use, radiation exposure, nutritional problems, vitamin supplement intake, excessive tea, and coffee consumption during pregnancy were similar between the two groups (all p-values >0.05). Also, there was no statistically significant difference between the case and control groups for some data on birth (birthplace, birth time and weight, birth complication(s), history of blood incompatibility, and blood transfusion) (all p-values> 0.05).

While the rate of positive family history for autism was 44.1% in the case group, it was 5.6% in the control group (p <0.001). 48.5% of the children with ASD and 34.3% of the control group children were the first children (p=0.016). While the nulliparity rate in the mothers of those in the ASD group was 49.3%, this rate was 34.3% in the control group (p=0.011). The smoking rate of the mothers during pregnancy was 28.7% in the ASD group and 16.8% in the control group (p=0.018). Gestational diabetes history was 8.8% in the ASD group and 2.8% in the control group (p=0.030). While 23.5% of the ASD group had a history of severe infection requiring inpatient treatment during pregnancy, this value was 11.2% in the control group (p=0.006). 66.2% of the children with ASD and 82.5% of the children in the control group were born by normal vaginal delivery (p=0.002). 72 (57.1%) of the children with ASD and 59 (41.3%) of the children in the control group had a history of neonatal jaundice (p=0.009). Besides, 44 (34.9%) of children with ASD had a history of seizure/convulsion in 6 (4.2%) of healthy children (p <0.001).

Independent variable importance values of the MLP model is demonstrated in Table 1. According to the findings of the MLP model, the five most important factors were the mean age of first words (months), the mean age of head control (months), the mean age of sitting without support (months), history of autism in the family, and the mean paternal age at pregnancy (years), respectively. Importance and normalized importance values for each variable are presented in descending order (Table 1).

The classification results of the training and testing samples for the model are summarized in Table 2. Overall percentages of the training and testing samples were 91.4% and 88.0%.

TABLE I
INDEPENDENT VARIABLE IMPORTANCE VALUES OF THE MLP MODEL

Variable	Importance	Normalized Importance
Age of first words (months)	0.189	100.0%
Age of head control (months)	0.140	74.3%
Age of sitting without support (months)	0.060	31.8%
Paternal age at pregnancy (years)	0.057	30.1%
Positive family history of ASD	0.055	29.0%
Maternal age at pregnancy (years)	0.047	25.1%
Month at birth	0.029	15.5%
Regular drug use during pregnancy	0.028	14.9%
Number of sibling(s)	0.025	13.1%
Threatened miscarriage	0.022	11.4%
Family income level	0.021	11.2%
Excessive tea and/or coffee consumption during pregnancy	0.021	11.2%
Birth type	0.021	10.9%
Birth place	0.020	10.8%
Smoking during pregnancy	0.020	10.8%
Order of sibling(s)	0.020	10.5%
Other endocrine problems other than diabetes during pregnancy	0.019	10.2%
Gestational age at birth (years)	0.019	10.1%
First-born	0.019	10.1%
Paternal education level	0.018	9.5%
History of neonatal jaundice	0.017	9.1%
Vitamins, minerals, and supplements in pregnancy	0.017	8.8%
Gender	0.015	7.7%
Birth season	0.014	7.2%
Seizure/convulsion history	0.012	6.6%
Blood incompatibility	0.011	5.8%
Birth complication(s)	0.010	5.2%
Severe infection in pregnancy	0.010	5.1%
Birth weight	0.010	5.1%
Place of residence	0.009	4.9%
Hypertension in pregnancy	0.008	4.3%
Nulliparity	0.004	2.4%
Gestational diabetes	0.003	1.8%
Age of starting to walk (months)	0.002	1.3%
Maternal education level	0.002	1.3%
Nutritional problems during pregnancy	0.002	1.2%
Parental consanguinity	0.002	1.0%
Blood transfusion	0.001	0.5%
Radiation exposure in pregnancy	0.000	0.1%

TABLE II
THE CLASSIFICATION RESULTS OF THE TRAINING AND TESTING SAMPLES FOR THE MODEL

Sample	Observed	Predicted		
		Autism	Control	Percent Correct
Training	**Autism**	76	15	83.5%
	Control	1	94	98.9%
	Overall Percent	41.4%	58.6%	91.4%
Testing	**Autism**	28	7	80.0%
	Control	3	45	93.8%
	Overall Percent	37.3%	62.7%	88.0%

Table 3 indicates the model summary information for both training and testing stages. Cross entropy errors were 54.226 for the training sample and 26.631 for the testing sample. Besides, AUC for the model was 0.922 for the separation of the autism and control groups.

TABLE III
THE MODEL SUMMARY INFORMATION FOR BOTH TRAINING AND TESTING STAGES

Model Summary		
Training	Cross-Entropy Error	54.226
	Percent Incorrect Predictions	8.6%
Testing	Cross-Entropy Error	26.631
	Percent Incorrect Predictions	12.0%

6. DISCUSSION

In this study, 39 variables were included in the modelling for ASD prediction in MLP modelling, and variables contributing to the diagnosis of ASD were determined according to their contribution percentages. Thus, among the variables selected for ASD diagnosis, it was found that the age of onset of the first words, age of providing head control, age of starting to sit without support, history of autism in the family, and age of the father at pregnancy was the most important five variables. On the other hand, it was determined that the presence of nutritional problems during pregnancy, parental consanguinity, the history of blood transfusion during the neonatal period, and the history of radiation exposure in pregnancy affected the least effect on the diagnosis of ASD.

Machine learning is a field of research that creates prediction models within the framework of accuracy and originality from data sets by integrating artificial intelligence, mathematics, and other branches of science. In this way, it is ensured that the data give the best estimation outcome without too much human intervention [15]. Regarding the relevant literature in ASD, it is seen that artificial intelligence-based machine learning systems are used for different purposes in ASD [15-20]. In a meta-analysis study aiming to determine the accuracy of machine learning algorithms in+ diagnosing ASD from brain magnetic resonance imaging (MRI) studies, 40 studies were included. This study has found that the sensitivity is 0.83 and the specificity is 0.84 for the structural MRI (sMRI) subgroup meta-analysis, and researchers have emphasized that further research is required to determine the potential benefits of machine learning in the clinical setting [16]. In a study conducted on the diagnosis of ASD using machine learning methods from 3-minute home videos, 30 behavioural characteristics including eye contact, expressive language, emotion expression, echolalia, joint attention were evaluated, and eight different machine learning methods were used for ASD prediction. Based on the results of the study, it has been suggested that modelling can be used in the diagnosis of ASD [17]. In addition, it is seen that artificial intelligence-based systems can be used for ASD in different areas such as determining the best distinguishing features in assessment tools used in the diagnosis of ASD [21], the prediction of stereotypical movements, which are one of the basic symptoms in ASD, with different techniques [22], analysis of neonatal leukocyte epigenomic markers for ASD prediction [23], developing a deep convolutional neural network that can help children with ASD recognize facial expressions accurately using mobile devices [24], autism classification based on logistic regression analysis [25].

In particular, clinical features and neuroimaging techniques have been used in artificial intelligence-based models for the diagnosis of ASD. Studies examining models based on different environmental features are extremely limited. In our study, modelling was performed based on various developmental characteristics and prenatal-perinatal risk factors. When the relevant literature in this field is examined; in a meta-analysis study conducted with 37,634 children with ASD, several factors such as mother and father's age above 35, some ethnic origins, history of hypertension and diabetes during pregnancy, birth complications, cesarean delivery, low birth weight were associated with ASD. In contrast, this study has indicated that umbilical cord entanglement, premature rupture of membranes, 5th minute Apgar score of less than 7, and respiratory tract infection are not among the factors that increase the risk in ASD [26]. In the study of *Modabbernia et al.*, it was found that vaccination, maternal smoking, thimerosal exposure, and assisted reproductive techniques were not associated with ASD risk. In parallel with our study, advanced parental age is associated with an increased risk of ASD. This study has also shown that while birth complications related to trauma or ischemia and hypoxia are strongly related to ASD, other factors related to pregnancy, such as maternal obesity, maternal diabetes, and cesarean delivery, are less associated with ASD [27]. In our study, among many prenatal-perinatal variables, the paternal age at pregnancy was one of the most effective factors in diagnosing ASD. In addition, the results of our study demonstrated that the age of the first words, the age of obtaining head control, the age of sitting without support, and the family history of autism are more dominant characteristics for the diagnosis of ASD than other prenatal-perinatal factors. Among these variables, the two most predictive factors are the age of onset of the first words and the time to provide head control.

ASD is one of the important neurodevelopmental disorders that significantly impair the quality of life of individuals in different social environments such as school [1, 28]. Therefore, making the diagnosis process as early as possible also provides significant benefits for the treatment and follow-up process [29]. However, one of the most important obstacles to early diagnosis is the lack of any definitive markers for diagnosis. This problem suggests the importance of determining the risk factors that would bring ASD to mind and lead to the diagnosis. The strength of this study is that it reveals the importance of several developmental characteristics for diagnosis before clinical signs fully appear. Despite our study has many strengths, it also has some limitations. Since the data included in the modelling are recruited retrospectively, some data may be missing or subject to recall bias. Future longitudinal studies with a larger sample size are of great importance in supporting the findings of this study.

REFERENCES

[1] C. Lord, M. Elsabbagh, G. Baird, and J. Veenstra-Vanderweele, "Autism spectrum disorder," The Lancet, vol. 392, pp. 508-520, 2018.

[2] K. Sanchack and C. A. Thomas, "Autism spectrum disorder: Primary care principles," American family physician, vol. 94, pp. 972-979, 2016.

[3] E. Fombonne, "The rising prevalence of autism," Journal of Child Psychology and Psychiatry, vol. 59, pp. 717-720, 2018.

[4] K. Lyall, L. Croen, J. Daniels, M. D. Fallin, C. Ladd-Acosta, B. K. Lee, et al., "The changing epidemiology of autism spectrum disorders," Annual review of public health, vol. 38, pp. 81-102, 2017.

[5] V. Courchesne, A.-A. S. Meilleur, M.-P. Poulin-Lord, M. Dawson, and I. Soulières, "Autistic children at risk of being underestimated: school-based pilot study of a strength-informed assessment," Molecular Autism, vol. 6, p. 12, 2015.

[6] C. A. Labarrere, J. Woods, J. Hardin, G. Campana, M. Ortiz, B. Jaeger, et al., "Early prediction of cardiac allograft vasculopathy and heart transplant failure," American Journal of Transplantation, vol. 11, pp. 528-535, 2011.

[7] H. Gardener, D. Spiegelman, and S. L. Buka, "Prenatal risk factors for autism: comprehensive meta-analysis," The British journal of psychiatry, vol. 195, pp. 7-14, 2009.

[8] D. Rahul, "Machine learning in medicine," Circulation, vol. 132, pp. 1920-1930, 2015.

[9] A. Ari and D. Hanbay, "Deep learning based brain tumor classification and detection system," Turkish Journal of Electrical Engineering & Computer Sciences, vol. 26, pp. 2275-2286, 2018.

[10] M. C. Çolak, C. Çolak, H. Kocatürk, S. Sagiroglu, and İ. Barutçu, "Predicting coronary artery disease using different artificial neural network models/Koroner arter hastaliginin degisik yapay sinir agi modelleri ile tahmini," Anadulu Kardiyoloji Dergisi: AKD, vol. 8, p. 249, 2008.

[11] E. Güldoğan, T. Zeynep, and C. Çolak, "Classification of Breast Cancer and Determination of Related Factors with Deep Learning Approach," The Journal of Cognitive Systems, vol. 5, pp. 10-14.

[12] E. Güldoğan, T. Zeynep, A. Ayça, and C. Çolak, "Performance Evaluation of Different Artificial Neural Network Models in the Classification of Type 2 Diabetes Mellitus," The Journal of Cognitive Systems, vol. 5, pp. 23-32.

[13] J. Stewart, P. Sprivulis, and G. Dwivedi, "Artificial intelligence and machine learning in emergency medicine," Emergency Medicine Australasia, vol. 30, pp. 870-874, 2018.

[14] A. Association, "American Psychiatric Association's Diagnostic and statistical manual of mental disorders (DSM-V)," 2013.

[15] F. Thabtah, "Machine learning in autistic spectrum disorder behavioral research: A review and ways forward," Informatics for Health and Social Care, vol. 44, pp. 278-297, 2019.

[16] S. J. Moon, J. Hwang, R. Kana, J. Torous, and J. W. Kim, "Accuracy of Machine Learning Algorithms for the Diagnosis of Autism Spectrum Disorder: Systematic Review and Meta-Analysis of Brain Magnetic Resonance Imaging Studies," JMIR mental health, vol. 6, p. e14108, 2019.

[17] Q. Tariq, J. Daniels, J. N. Schwartz, P. Washington, H. Kalantarian, and D. P. Wall, "Mobile detection of autism through machine learning on home video: A development and prospective validation study," PLoS medicine, vol. 15, p. e1002705, 2018.

[18] A. Pratap and C. Kanimozhiselvi, "Predictive assessment of autism using unsupervised machine learning models," International Journal of Advanced Intelligence Paradigms, vol. 6, pp. 113-121, 2014.

[19] M. Duda, R. Ma, N. Haber, and D. Wall, "Use of machine learning for behavioral distinction of autism and ADHD," Translational psychiatry, vol. 6, pp. e732-e732, 2016.

[20] S. H. Lee, M. J. Maenner, and C. M. Heilig, "A comparison of machine learning algorithms for the surveillance of autism spectrum disorder," PloS one, vol. 14, p. e0222907, 2019.

[21] D. P. Wall, J. Kosmicki, T. Deluca, E. Harstad, and V. A. Fusaro, "Use of machine learning to shorten observation-based screening and diagnosis of autism," Translational psychiatry, vol. 2, pp. e100-e100, 2012.

[22] N. M. Rad, S. M. Kia, C. Zarbo, T. van Laarhoven, G. Jurman, P. Venuti, et al., "Deep learning for automatic stereotypical motor movement detection using wearable sensors in autism spectrum disorders," Signal Processing, vol. 144, pp. 180-191, 2018.

[23] R. O. Bahado-Singh, S. Vishweswaraiah, B. Aydas, N. K. Mishra, A. Yilmaz, C. Guda, et al., "Artificial intelligence analysis of newborn leucocyte epigenomic markers for the prediction of autism," Brain research, vol. 1724, p. 146457, 2019.

[24] M. I. U. Haque, "A Facial Expression Recognition Application Development Using Deep Convolutional Neural Network For Children With Autism Spectrum Disorder To Help Identify Human Emotions," 2019.

[25] F. Thabtah, N. Abdelhamid, and D. Peebles, "A machine learning autism classification based on logistic regression analysis," Health information science and systems, vol. 7, p. 12, 2019.

[26] C. Wang, H. Geng, W. Liu, and G. Zhang, "Prenatal, perinatal, and postnatal factors associated with autism: a meta-analysis," Medicine, vol. 96, 2017.

[27] A. Modabbernia, E. Velthorst, and A. Reichenberg, "Environmental risk factors for autism: an evidence-based review of systematic reviews and meta-analyses," Molecular autism, vol. 8, p. 13, 2017.

[28] P. Titelman, Differentiation of self: Bowen family systems theory perspectives: Routledge, 2014.

[29] S. Mishra, D. Joshi, R. Ribeiro, and S. Anand, "Kinematics-coordinated walking pattern based on embedded controls," Journal of medical engineering & technology, vol. 34, pp. 329-334, 2010.

BIOGRAPHIES

İIknur UCUZ, graduated from Çukurova University Faculty of Medicine in 2005 and received her speciality education in the Department of Child and Adolescent Psychiatry at Atatürk University between 2009-2014. Since 2017, she has been working as a faculty member at İnönü University Faculty of Medicine, Department of Child and Adolescent Psychiatry. She is interested in anxiety disorders, autism spectrum disorders, attention deficit and hyperactivity disorder, artificial intelligence and machine learning in children and adolescents.

Ayla UZUN ÇİÇEK, who graduated from Cumhuriyet University Faculty of Medicine in 2007, received her speciality education in the Department of Child and Adolescent Psychiatry at Akdeniz University between 2009-2013. Between 2013-2015, he served as a specialist in Rize Recep Tayyip Erdoğan Training and Research Hospital. Since 2015, she has been working as a faculty member at Cumhuriyet University Faculty of Medicine, Department of Child and Adolescent Psychiatry. She deals with mood disorders, anxiety disorders, autism spectrum disorders, attention deficit and hyperactivity disorder in children and adolescents.

DETERMINATION OF *CRYPTOSPORIDIUM* SPP. RISK FACTORS USING MULTILAYER PERCEPTRON NEURAL NETWORK AND RADIAL BASED FUNCTIONAL ARTIFICIAL NEURAL NETWORK METHOD

U. Karaman, and I. Balikci Cicek

Abstract— *Aim:* In the study, it is aimed to compare the estimates of Multilayer artificial neural network (MLPNN) and radial based function artificial neural network (RBFNN) methods, which are among the artificial neural network models in the presence and absence of *Cryptosporidium* spp., and to determine the factors associated with parasite.

Materials and Methods: In the study, "*Cryptosporidium* spp. Dataset," the data set named was obtained from Ordu University. In order to classify the presence and absence of *Cryptosporidium* spp, MLPNN, and RBFNN methods, which are among the artificial neural network models, were used. The classification performance of the models was evaluated with accuracy from the classification performance criteria.

Results: The accuracy, which is the performance criterion obtained with MLPNN, was obtained as 75% of the applied models. The accuracy, which is the performance criterion obtained with the RBFNN model, was achieved as 71.4%. When the effects of variables in the data set in this study on the presence and absence of *Cryptosporidium* spp. are examined, the three most important variables for the MLPNN model were nausea-vomiting, General Puriri, and sex, respectively. For the RBFNN model, age was obtained as cancer and General Puriri.

Conclusion: It was seen that MLPNN and RBFNN models used in this study gave successful predictions in classifying the presence and absence of *Cryptosporidium* spp.

Keywords— Multilayer perceptron neural network, Radial-based function neural network, classification, *Cryptosporidium* spp., risk factors.

1. INTRODUCTION

Intestinal parasites are a significant public health problem worldwide, including in developing countries [1]. Among the intestinal parasites, the coccidian parasites cryptosporidium is one of the obligate intracellular parasites that cause diarrhea in all age groups and individuals with normal immunity [2].

Among the Criptosporidiums (*Cryptosporidium* spp.), C. parvum, the most diseased species in humans, is located in the microvilli of intestinal epithelial cells, causing short-term (about two weeks) spontaneous diarrhea in people with sufficient immunity, and maybe life-threatening in the host whose immune system is suppressed [3]. In immunocompromised individuals, the parasite can spread from the intestinal tract to the bile ducts, pancreas, stomach, respiratory system, and kidneys through the hematogenous pathway. Cryptosporidiums can be transmitted by contaminated water and food, from person to person or from animal to person [4].

Artificial neural networks (ANN) are parallel, distributed information processing models that are developed using the physiology of the human brain and are connected by weighted connections, each consisting of processing elements with their own memory, and are computer programs that imitate biological neural networks [5]. The most essential task of an artificial neural network is to determine an output set that can correspond to an input set shown to it. In order to do this, the network is trained with examples of the relevant event and gained the ability to generalize [6].

Multilayer Artificial Neural Networks (MLPNN) model has been the most used neural network model, especially in medical and engineering applications. This model is widely used because many learning algorithms can be easily used in the training of this network. Multilayer networks consist of an input layer, one or more hidden layers, and an output layer. The information flow is forward, and there is no feedback [7]. The purpose of this method is; It is to make the error between the desired output of the network and the output it produces to a minimum [8].

Feedforward neural networks are widely used in many areas, such as controlling nonlinear systems in addition to modeling. One of the feed-forward neural networks is Radial Based Function Artificial Neural Networks (RBFNN) [9]. RBFNN is a particular case of a multilayer feed-forward artificial neural network and has two distinctive characteristics. The first is that it has only one hidden layer. The second feature is that radial based functions are used as activation functions in the hidden layer. Another essential feature of radial-based artificial neural networks is the transfer of information from input neurons to hidden layer neurons without change [10].

In this study, *Cryptosporidium* spp. by applying MLPNN and RBFNN methods to the data set, it aimed to classify the presence and absence of *Cryptosporidium* spp. and determine the risk factors.

İpek BALIKCI CICEK, Inonu University Department of Biostatistics and Medical Informatics, Faculty of Medicine, Malatya, Turkey, (ipek.balikci@inonu.edu.tr)

Ülkü KARAMAN, Ordu University Department of Medical Parasitology Dep., Faculty of Medicine, Ordu, Turkey, (ukaraman@ordu.edu.tr)

Manuscript received Oct 4, 2020; accepted Oct 29, 2020.
Digital Object Identifier:

2. MATERIAL AND METHODS

2.1. Dataset

In this study, *Cryptosporidium* spp. classification process was performed by applying MLPNN and RBFNN methods to the "*Cryptosporidium* spp" data set obtained from Ordu University for the presence-absence situation. There are a total of 497 patients in this dataset. There were 142 (28.6%) people with *Cryptosporidium* spp. and 355 (71.4%) without the parasite. The variables and the descriptive properties of the variables in the relevant data set are given in Table 1.

TABLE I
VARIABLES IN THE DATA SET AND DESCRIPTIVE PROPERTIES OF VARIABLES

Variables	**Variable Explanation**	**Variable type**	**Variable role**
Cryptosporidium spp.	Parasite (0 = absence, 1 = presence)	Qualitative	Dependent/ Target
Age	Age	Quantitative	Independent/ Predictor
Sex	1=male, 2=female	Qualitative	Independent/ Predictor
nausea / vomiting	0= no nausea or vomiting 1= there is nausea and vomiting	Qualitative	Independent/ Predictor
immunosuppressive	0 = absence, 1 = presence	Qualitative	Independent/ Predictor
eosinophilia	0 = absence, 1 = presence	Qualitative	Independent/ Predictor
diabetes	0 = absence, 1 = presence	Qualitative	Independent/ Predictor
cancer	0 = absence, 1 = presence r	Qualitative	Independent/ Predictor
urine syphilis	0 = absence, 1 = presence	Qualitative	Independent/ Predictor
diarrhea	0 = absence, 1 = presence	Qualitative	Independent/ Predictor
neutropenia	0 = absence, 1 = presence	Qualitative	Independent/ Predictor
obesity	0 = absence, 1 = presence	Qualitative	Independent/ Predictor
anemia	0 = absence, 1 = presence	Qualitative	Independent/ Predictor
Aurtiker	0 = absence, 1 = presence	Qualitative	Independent/ Predictor
General puriri	0 = absence, 1 = presence	Qualitative	Independent/ Predictor
Kürtiker	0 = absence, 1 = presence	Qualitative	Independent/ Predictor
urticaria	0 = absence, 1 = presence	Qualitative	Independent/ Predictor
ucolite	0 = absence, 1 = presence	Qualitative	Independent/ Predictor

3. MULTILAYER PERCEPTRON NEURAL NETWORK (MLPNN)

This model, which was developed by Rumelhart in 1986, is also called the error propagation model. Multilayer Perceptron neural networks have multiple layers between the input and output layers. It consists of the input layer, output layer, and intermediate layers. The processor elements in the input layer act as a buffer that distributes the input signals to the processor elements in the middle layer. The information flow runs from the input layer to the middle layer and then to the output layer [11].

Training of the multi-layered artificial neural network is carried out according to the "generalized delta rule". In multilayer neural networks, first, an example of the network is introduced. As a result of the example, what kind of result will be obtained is revealed. The examples are applied to the input layer, processed in the intermediate layers, and the error between the desired output is spread back over the weights, changing the weights until the error is minimized. The multi-layer perceptron network is a forward feed network, and the most general result is obtained from the output layer [12].

3.1. Radial Based Function Neural Network (RBFNN)

Radial Based Artificial Neural network first emerged in a multivariate real interpolation solution [13]. RBFNN is a curve-fitting approach in multidimensional space. Training of RBFNN can be called the problem of finding the best surface suitable for data in multidimensional space. RBFNN is feed-forward networks consisting of three layers as input, hidden, and output layer.

The input layer consists of a source of artificial nerve cells. The hidden layer is the hidden layer whose number of artificial nerve cells can be changed and uses a radian-based function as an activation function. The output layer is the part where the output of the network is produced according to the input values. There is a non-linear transformation from the input layer to the hidden layer and a linear transformation from the hidden layer to the output layer [14].

RBFNN working principle; to determine the radial-based function with the appropriate width and center values in the intermediate layer according to the input values in the input layer and to create the linear combinations of the outputs of the radial-based functions with the appropriate weight values and to determine the relationship between the input values and the output values [15].

3.2. Performance Evaluation of Models

In the performance evaluation of the radial-based artificial neural network and multilayer artificial neural network models, which were created to predict the factors that may be associated with the presence or absence of *Cryptosporidium* spp., the performance criteria obtained by using the classification matrix given below were used.

The performance criteria used in the performance evaluation of the models in this study are given below.

Accuracy = (TP+TN)/(TP+TN+FP+FN)

TABLE II
CLASSIFICATION MATRIX FOR CALCULATING PERFORMANCE CRITERIA

		Real		
		Positive	**Negative**	**Total**
Predicted	**Positive**	True positive (TP)	False negative (FN)	TP+FN
	Negative	False positive (FP)	True negative (TN)	FP+TN
	Total	TP+FP	FN+TN	TP+TN+ FP+FN

4. DATA ANALYSIS

Quantitative data are expressed as median (minimum-maximum) and qualitative data as number (percentage). The Kolmogorov-Smirnov test evaluated conformity to normal distribution.

In terms of independent variables, whether there is a statistically significant difference between the "parasite presence " and "parasite absence " groups, which are the categories of the dependent/target variable (*Cryptosporidium* spp.), and whether there is a relationship, Mann-Whitney U test, Pearson chi-square test, Continuity Correction test, and Fisher's Exact test. It was examined using the chi-square test values of p<0.05 were considered statistically significant. IBM SPSS Statistics 26.0 package program was used for all analyzes.

For the validity of the model, a 10-fold cross-validation method was used. In the 10-fold cross-validation method, all data is divided into ten equal parts. One part is used as a test set, and the remaining nine parts are used as a training dataset, and this process is repeated ten times.

5. RESULTS

Descriptive statistics for quantitative independent variables examined in this study are given in Table 3, and descriptive statistics for qualitative independent variables are given in Table 4. There is a statistically significant relationship between the dependent/target variable groups (p<0.05) in terms of cancer variable.

TABLE III
DESCRIPTIVE STATISTICS FOR QUANTITATIVE INDEPENDENT VARIABLES

Variable	Cryptosporidium		**p-value***
	Parasite absence	Parasite presence	
	Median(min-max)	**Median(min-max)**	
age	22 (1-84)	21 (2-72)	0.619

***: Mann Whitney U test**

TABLE IV
DESCRIPTIVE STATISTICS FOR QUALITATIVE INDEPENDENT VARIABLES

Variables		Cryptosporidium		**p-value**
		Parasite absence	Parasite presence	
		Number **(%)**	**Number** **(%)**	
sex	male	184 (51.8)	76 (53.5)	0.733^{*}
	female	171 (48.2)	66 (46.5)	
nausea / vomiting	absence	341 (96.1)	135 (95.1)	0.805^{**}
	presence	14 (3.9)	7 (4.9)	
immunosuppressive	absence	344 (96.9)	136 (95.8)	0.586^{***}
	presence	11 (3.1)	6 (4.2)	
eosinophilia	absence	352 (99.2)	142 (100)	0.561^{***}
	presence	3 (0.8)	0 (0)	
diabetes	absence	353 (99.4)	141 (99.3)	1^{***}
	presence	2 (0.6)	1 (0.7)	
cancer	absence	336 (94.6)	124 (87.3)	**0.009^{**}**
	presence	19 (5.4)	18 (12.7)	
urine syphilis	absence	350 (98.6)	142 (100)	0.328^{***}
	presence	5 (1.4)	0 (0)	
diarrhea	absence	280 (78.9)	108 (76.1)	0.493^{*}
	presence	75 (21.1)	34 (23.9)	
neutropenia	absence	350 (98.6)	142 (100)	0.328^{***}
	presence	5 (1.4)	0 (0)	
obesity	absence	353 (99.4)	142 (100)	1^{***}
	presence	2 (0.6)	0 (0)	
anemia	absence	344 (96.9)	134 (94.4)	0.283^{**}
	presence	11 (3.1)	8 (5.6)	
General puriri	absence	336 (94.6)	132 (93.0)	0.607^{**}
	presence	19 (5.4)	10 (7.0)	
Aurtiker	absence	352 (99.2)	141 (99.3)	1^{***}
	presence	3 (0.8)	1 (0.7)	
Kürtiker	absence	349 (98.3)	139 (97.9)	0.719^{***}
	presence	6 (1.7)	3 (2.1)	
urticaria	absence	344 (96.9)	137 (96.5)	0.783^{***}
	presence	11 (3.1)	5 (3.5)	
ucolite	absence	351 (98.9)	138 (97.2)	0.233^{***}
	presence	4 (1.1)	4 (2.8)	

***: Pearson chi-square test, **: Continuity Correction test,**
*****: Fisher's Exact test**

Classification martix of MLPNN and RBFNN models are given in Table 5 and Table 6, respectively.

TABLE V
CLASSIFICATION MATRIX OF MLPNN MODEL

Real / **Predicted**	presence	absence	**Total**
presence	2	38	40
absence	0	112	112
Total	2	150	152

TABLE VI
CLASSIFICATION MATRIX OF RBFNN MODEL

Real **Predicted**	presence	absence	**Total**
presence	0	44	44
absence	0	110	110
Total	0	154	154

Table 7, shows the values of the performance criteria calculated from the models created to classify the *Cryptosporidium* spp.

TABLE VII
PERFORMANCE CRITERIA VALUES CALCULATED FROM CREATED MODELS IN THE TESTING PHASE

Model **Performance Metric**	**MLPNN**	**RBFNN**
	Value	**Value**
Accuracy (%)	75.0	71.4
AUC	0.515	0.547

AUC: Area under the ROC curve; MLPNN: Multilayer Perceptron Neural Network; RBFNN: Radial Based Function Neural Network

In this study, the importance values of the factors associated with the *Cryptosporidium* spp. are given in Table 8, while the values for these importance percentages are shown in Figure 1.

TABLE VIII
IMPORTANCE VALUES OF EXPLANATORY VARIABLES ACCORDING TO MLPNN AND RBFNN MODELS

Explanatory Variables	**MLPNN**	**RBFNN**
sex	0.085	0.007
nausea / vomiting	0.185	0.069
immunosuppressive	0.032	0.067
eosinophilia	0.062	0.056
diabetes	0.019	0.055
cancer	0.025	0.076
urine syphilis	0.029	0.053
diarrhea	0.030	0.009
neutropenia	0.031	0.056
obesity	0.051	0.056
anemi	0.072	0.063
General puriri	0.157	0.072
Aurtiker	0.038	0.059
Kürtiker	0.019	0.063
urticaria	0.070	0.065
ucolite	0.075	0.061
age	0.018	0.114
Total	1	1

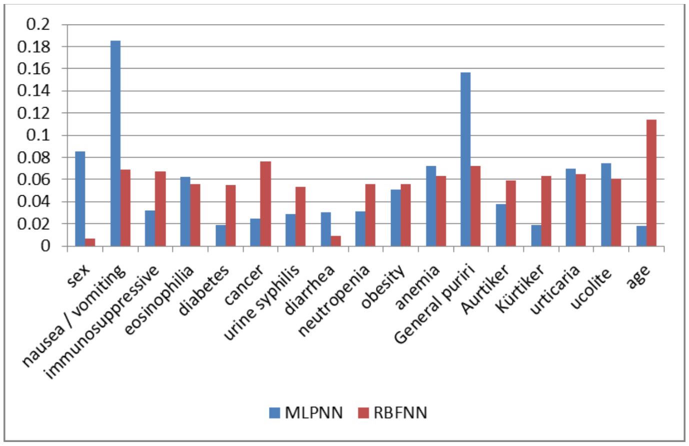


Fig.1. The importance values for possible risk factors

6. DISCUSSION

Cryptosporidium spp. oocysts, which are obligate intracellular parasites, are 4-6 µm in size and spread among living things as a result of consuming water and food contaminated with feces [4, 16, 17].

The parasite has a high prevalence rate in some occupational groups (animal husbandry, veterinarians, laboratory staff, nursery staff), people who travel to endemic areas, those who live in places where hygienic conditions are inadequate, children, the elderly, and those who come into close contact with infected people [18]. *Cryptosporidium* outbreaks have been reported from public swimming pools, common meals, well water, and unhygienic drinking water. It has also been stated that there may be a transition from animals to humans in rural areas [18]. In the spread of the parasite, lack of clean water and sanitation facilities, crowded home environment, and close contact of reservoir animals to individuals are potentially effective [19].Symptoms of cryptosporidiosis differ depending on the type of infected host, the state of the immune system, and age [20].

Artificial neural networks are a method used to estimate the relationship between dependent and independent variables. Artificial Neural Networks in general; It is accepted as a powerful method es such as parameter estimation, classification, and the structure of existing data in many statistical processes parameter estimation, classification, and optimization. Artificial neural networks can reveal complex relationships between predictive variables and make inferences [21].

In this study, the multilayer artificial neural network and radial-based artificial neural network models, which are among the artificial neural network models, were obtained from *Cryptosporidium* spp. It was applied to the data set and aimed to compare the classification estimates of these two models. In this context, *Cryptosporidium* spp. The factors that may be associated with the positivity-negativity (dependent variable) were estimated by multilayer artificial neural network and radial based artificial neural network models. Thus, it has been shown that artificial neural network models can be used in the classification problem.

The imbalanced class problem is one of the important topics in machine learning. In the data set being studied, one of the classes formed by the observations is higher in number than the class or classes formed by other observations, revealing the imbalanced class problem. The problem of bias arises in the classification of sick and non-sick individuals in a two-class data set with an imbalance in the distribution between classes. Because machine learning models used in classification and sensitive to unbalanced class distributions are under the influence of large class (s) and the existence of small classes disappears [22, 23]. There were 142 (28.6%) people with *Cryptosporidium* spp. and 355 (71.4%) people without parasite in this data set. This situation causes the classification algorithms to give biased results and the results to be interpreted incorrectly. Therefore, the accuracy value obtained in the MLPNN and RBFNN models was 75.0 and 71.4, respectively.

In this study, among the performance criteria used to compare classification performances according to the accuracy result, the MLPNN model gave better predictive results than the RBFNN model in the classification of presence-absence of *Cryptosporidium* spp. According to the MLPNN model, the three most important risk factors that may be associated with the presence and absence of *Cryptosporidium* spp. are; nausea-vomiting, general puriri, and sex have been obtained. The RBFNN model estimated as age, cancer, and general puriri.

The presence of the parasite according to the artificial neural network models used according to the findings obtained in the study; They differ in nausea-vomiting, general puriri, age, cancer, and sex variables. Accordingly, it was concluded that *Cryptosporidium* spp. positivity should be investigated in line with the complaints of general puriri, nausea-vomiting, and cancer patients.

REFERENCES

[1] B. Bhattachan, J. B. Sherchand, S. Tandukar, B. G. Dhoubhadel, L. Gauchan, and G. Rai, "Detection of Cryptosporidium parvum and Cyclospora cayetanensis infections among people living in a slum area in Kathmandu valley, Nepal," BMC research notes, vol. 10, pp. 1-5, 2017.

[2] P. Kumar, O. Vats, D. Kumar, and S. Singh, "Coccidian intestinal parasites among immunocompetent children presenting with diarrhea: Are we missing them?," Tropical Parasitology, vol. 7, p. 37, 2017.

[3] D. Dirim Erdoğan, N. Turgay, and M. Z. Alkan, "Bir Cryptosporidiosis olgusunun kinyoun Asit-fast boyası ve polimeraz zincir reaksiyonu (PZR) ile takibi," Türkiye Parazitoloji Dergisi, vol. 27, pp. 237-239, 2003.

[4] D. P. Clark, "New insights into human cryptosporidiosis," Clinical Microbiology Reviews, vol. 12, pp. 554-563, 1999.

[5] [M. Çuhadar, "Türkiye'ye yönelik diş turizm talebinin MLP, RBF ve TDNN yapay sinir aği mimarileri ile modellenmesi ve tahmini: karşilaştirmali bir analiz," Journal of Yasar University, vol. 8, 2013.

[6] S. Tuna, "Şablon eşleme ve çok katmanlı algılayıcı kullanılarak yüz tanıma sisteminin gerçeklenmesi," 2008.

[7] A. Auclair, "Feed-forward neural networks applied to the estimation of magnetic field distributions," 2004.

[8] Ç. Çatal and L. Özyilmaz, "Çok katmanli algilayici ile multiple myeloma hastaliğinin gen ekspresiyon veri çözümlenmesi."

[9] C.-N. Ko, "Identification of non-linear systems using radial basis function neural networks with time-varying learning algorithm," IET signal processing, vol. 6, pp. 91-98, 2012.

[10] O. Akbilgiç, Ğ. Danışman, and H. Bozdoğan, "Hibrit radyal tabanli fonksiyon ağlari ile değişken seçimi ve tahminleme: menkul kiymet yatirim kararlarina ilişkin bir uygulama."

[11] B. K. Wong, T. A. Bodnovich, and Y. Selvi, "Neural network applications in business: A review and analysis of the literature (1988–1995)," Decision Support Systems, vol. 19, pp. 301-320, 1997.

[12] E. Öztemel and Y. S. Ağları, "Papatya yayıncılık," İstanbul, 2003.

[13] M. D. Buhmann, Radial basis functions: theory and implementations vol. 12: Cambridge university press, 2003.

[14] U. Okkan and H. yıldırım Dalkiliç, "Radyal tabanlı yapay sinir ağları ile Kemer Barajı aylık akımlarının modellenmesi," Teknik Dergi, vol. 23, pp. 5957-5966, 2012.

[15] C. Cetinkaya, "Retina Görüntülerinde Radyal Tabanlı Fonksiyon Sinir Ağları İle Damar Tipik Noktalarının Tespit Edilmesi," Ege Üniversitesi Uluslararası Bilgisayar Enstitüsü, YL Tezi, 2011.

[16] S. Özçelik, Ö. Poyraz, K. Kalkan, E. Malatyalı, and S. Değerli, "The investigation of Cryptosporidium spp. prevalence in cattle and farmers by ELISA," Kafkas Üniversitesi Veteriner Fakültesi Dergisi, vol. 18, 2012.

[17] J.K. Griffiths, "Human cryptosporidiosis: epidemiology, transmission, clinical disease, treatment, and diagnosis," in Advances in parasitology. vol. 40, ed: Elsevier, 1998, pp. 37-85.

[18] D. Miron, J. Kenes, and R. Dagan, "Calves as a source of an outbreak of cryptosporidiosis among young children in an agricultural closed community," The Pediatric infectious disease journal, vol. 10, pp. 438-441, 1991.

[19] G. Börekçi, F. Otağ, and G. Emekdaş, "Mersin'de bir gecekondu mahallesinde yaşayan ailelerde Cryptosporidium prevalansı," İnfeksiyon Derg, vol. 19, pp. 39-46, 2005.

[20] Z. Egyed, T. Sreter, Z. Szell, and I. Varga, "Characterization of Cryptosporidium spp.—recent developments and future needs," Veterinary parasitology, vol. 111, pp. 103-114, 2003.

[21] M. Kayri and Ö. Çokluk, "Examining Factors of Academic Procrastination Tendency of University Students by using Artificial Neural Network," International Journal of Computer Trends and Technology, vol. 34, pp. 1-8, 2016.

[22] L. Nanni, C. Fantozzi, and N. Lazzarini, "Coupling different methods for overcoming the class imbalance problem," Neurocomputing, vol. 158, pp. 48-61, 2015.

[23] A. Sarmanova and S. Albayrak, "Alleviating class imbalance problem in data mining," in 2013 21st Signal Processing and Communications Applications Conference (SIU), 2013, pp. 1-4.

.

BIOGRAPHIES

İpek BALIKÇI ÇİÇEK obtained her BSc. degree in mathematics from Çukurova University in 2010. She received MSc. degree in biostatistics and medical informatics from the Inonu University in 2018. She currently continues Ph.D. degrees in biostatistics and medical informatics from the Inonu University. In 2014, she joined the Department of Biostatistics and Medical Informatics at Inonu University as a researcher assistant. Her research interests are cognitive systems, data mining, machine learning, deep learning.

Ülkü KARAMAN she is Associate professor Ordu University Faculty of Medicine, Department of Medical Parasitology, Ordu, Turkey. She has copleted he Ph.D. from Department of Medical Parasitology. Her area of research interest is medical parasitology and microbiolog. He has about 22 years working in health science.

ESTIMATION OF FACTORS RELATED TO PREMENSTRUAL SYNDROME IN FEMALE STUDENTS BY USING ARTIFICIAL NEURAL NETWORK MODEL

Y. Aksoy Derya, C. Çolak, B. Kayhan Tetik, A.N. Yılmaz, and Z. Özşahin

Abstract— **Objective:** Premenstrual syndrome is a disorder with psychological and physical symptoms in almost every cycle, and it concerns almost 90% of women of reproductive age. In recent years, studies investigating the relationship between premenstrual syndrome and vitamin D, trace elements, and lymphocyte/platelet ratios have been conducted. In this study, the factors associated with PMS in female students were tried to be determined by using an artificial neural network (ANN) model. **Materials and Methods:** This study was conducted on female students at Inonu University Faculty of Medicine and Health Sciences, between 01 May and 30 June 2019. Demographic characteristics and menstruation histories of 860 female students were collected and recorded for the study. A multi-layer perceptron artificial neural network model was used to determine the factors associated with premenstrual syndrome. The performance of the model is determined by the accuracy rate and the area under the process characteristic curve. **Results:** Correct classification rates of the created multi-layer perceptual neural network model for premenstrual syndrome were calculated to be 63.2% in the training data set and 63.0% in the test data set. Considering the importance values of the variables; it was found that the duration of active internet use (phone, tablet, computer) was the most influential factor on premenstrual syndrome and the economic status of a student was the least ,influential factor. **Conclusion:** According to the findings of the designed artificial neural network model, the three most important factors related to premenstrual syndrome were determined to be the duration of active Internet use, present age and age of menarche. Given the high prevalence of PMS, the uncertainty of etiology, and its potential to affect a woman's lifestyle; the use of artificial intelligence models with larger sample size and including different factors is recommended.

***Keywords*—** *Artificial neural network model, Associated factors, Female student, Premenstrual syndrome.*

Yeşim AKSOY DERYA, Inonu University Dep. of Midwifery, Faculty of Health Sciences, Malatya, Turkey, (yesim.aksoy@inonu.edu.tr)

Cemil ÇOLAK, Inonu University Department of Biostatistics and Medical Informatics, Faculty of Medicine, Malatya, Turkey, (cemil.colak@inonu.edu.tr)

Burcu KAYHAN TETİK, Inonu University Department of Family Medicine, Inonu University Medical Faculty, Malatya, Turkey, (burcu.tetik@inonu.edu.tr)

Ayşe Nur YILMAZ, Firat University Department of Midwifery, Faculty of Health Sciences, Elazıg, Turkey, (anucar@firat.edu.tr)

Zeliha ÖZŞAHİN, Inonu University Department of Midwifery, Faculty of Health Sciences, Malatya, Turkey, (zeliha.ozsahin@inonu.edu.tr)

Manuscript received Sep 17, 2020; accepted Nov 11, 2020.
Digital Object Identifier:

1. INTRODUCTION

PREMENSTRUAL syndrome (PMS) is a condition characterized by physical, cognitive, emotional and behavioral symptoms, which develop in the luteal phase of the menstrual cycle, decrease or end with the onset of menstruation [1,2]. Premenstrual syndrome is a common condition that concerns women's health and is seen in four out of ten women [3]. It has been reported that the prevalence of PMS is 47.8% worldwide and can vary between countries [4].

Although many studies have been conducted in the literature to determine the factors that play a role in the development of PMS, its etiology is still uncertain. The most accepted hypothesis is that PMS develops due to the effects of estrogens, and progesterone is related to neurotransmitters such as serotonin, opioids, GABA, and catecholamines [5,6]. In addition, the risk factors for PMS include genetic factors [7]. The uncertainty of the etiology of PMS indicates the need for studies evaluating PMS from different angles. Artificial Intelligence Technologies can also be used for such evaluations. Artificial Neural Network prediction technique simply imitates the human brain and has been used widely in many areas [8]. Artificial Neural Networks methodology has many essential features such as working with an unlimited number of variables, learning from data, and making generalizations. Thanks to these features, it provides very important benefits. ANN has been effectively used in many applications such as detecting previously unnoticed patterns in medical research data, classifying, determining the characteristics of medical images, and controlling medical devices [8]. It can be suggested that the ANN technique can potentially reduce the costs and the need for unnecessary researches [9]. Considering the achievements and potentiality of ANN research, there is no doubt that it will facilitate many innovations in medical science and will be a guide for more advanced diagnosis and treatment methods [10]. This study aimed to determine the factors associated with PMS in female students by using (ANN) model.

2. MATERIAL - METHOD

This study was designed as a cross-sectional study and conducted on female students studying in the Faculty of Medicine and Health Sciences of a public university, in the east of Turkey, between 01 May and 30 June 2019. The universe of the study consisted of all the students studying at the Faculty of Medicine and Health Sciences of the university (N = 1837). In this context, 860 students who accepted to participate in the study constituted the sample of the study. Students who accepted to participate in the study were selected from the relevant population by simple random sampling method. A personal information form and Premenstrual Syndrome Scale

(PMSS) created by the researchers in line with the literature review were applied to the participants.

2.1. Data Collection Tools

2.1.1. Personal Introduction Form

The Personal Information Form includes 17 questions regarding the sociodemographic information (age, department, class, family structure, etc.) and menstruation characteristics (age of menarche, frequency of menstruation (days), menstruation pattern) of the participants.

2.1.2. Premenstrual Syndrome Scale (PMSS)

The Premenstrual Syndrome Scale is a 44-item, five-point Likert-type scale developed by *Gençdoğan* in 2006 to measure premenstrual symptoms and determine the severity of these symptoms. The scale questions the symptoms in the last three months. The scale consists of 9 sub-dimensions, including depressive affection, anxiety, fatigue, irritability, depressive thoughts, pain, appetite changes, sleep changes, and distension. The sum of the scores obtained from 9 sub-dimensions constitutes the total score of PMSS. The lowest total score is 44, and the highest score is 220. The presence of PMS was defined as a score exceeding 50% of the highest score that can be obtained from the total scale and sub-dimensions. Cronbach alpha reliability of the scale was found to be 0.75 [11]. In this study, the Cronbach alpha reliability of the scale was found to be 0.96.

2.2. Data analysis

The data were analyzed by using SPSS 25.0 statistical package program. In the evaluation of the data percentage distribution and arithmetic mean were used. A value of $p<0.05$ was accepted as statistically significant. The factors associated with PMS were determined by using the multi-layer perceptron (MLP) ANN model. A single layer perceptron can only predict linear functions. The feed-forward multi-layer perceptrons with hidden layers between the input and output layers have no limitations in comparison to the single-layer perceptrons. In this study, 70% of the data was used to train the MLP ANN model, and the remaining 30% was used to test the model [12]. The performance of the model is determined by the accuracy rate and the area under the curve.

2.3. Ethical Approval

Ethical approval for the study was obtained from the Health Sciences Scientific Research and Publication Ethics Committee of Inonu University (No. 2020/1182). Before starting the research, the students were informed about the study, and it was warranted that their personal information would be protected and the volunteers were included in the study.

3. RESULTS

The distribution of the students participating in the study according to their socio-demographic characteristics is given in Table 1. It was found that the average age of the participants was 21.11 ± 1.97, the average height was 163.52 cm, and the average weight was 57.03 kg. It was determined that 75.6% of the students studied in health sciences, 24.4% were in the medical school, and 31.4% were 2nd graders.

TABLE I.
DISTRIBUTION OF SOCIO-DEMOGRAPHIC CHARECTERISTIC OF PARTICIPANTS (N=860)

Descriptive Characteristics	$\overline{X}$ ± SS	
Age (years)	21.11± 1.97 (min:17, max:35)	
Height (cm)	163.52± 5.65 (min:140, max:178)	
Weight (kg)	57.03± 9.05 (min:28, max:100)	
	n	**%**
Faculty		
Faculty of Health Sciences	650	75.6
Faculty of Medicine	210	24.4
Department		
Child Development	148	17.2
Midwifery	200	23.3
Physical therapy and rehabilitation	150	17.4
Audiology	152	17.7
Medicine	210	24.4
Class		
1st Class	202	23.5
2nd Class	270	31.4
3rd grade	177	20.6
4th Grade	177	20.6
5th grade	25	2.9
6th grade	9	1.0
Economic Status		
Low (0-500 TL)	234	27.2
Medium (501-1000 tl)	438	50.9
High (1001 TL and above)	188	21.9
Family structure		
Nuclear family	710	82.6
Extended family	127	14.8
Broken Family	23	2.7
Place of Residence		
Family	339	39.4
Student house	59	6.9
Dormitory	462	53.7
Geographical Region where she was born and raised		
Eastern Anatolia	394	45.8
Southeastern Anatolia	216	25.1
Aegean	25	2.9
Mediterranean	125	14.5
Central Anatolia	65	7.6
Black Sea	19	2.2
Marmara	16	1.9
Total	**860**	**100.0**

It was determined that 82.6% of the students had a nuclear family structure, 50.9% had a medium economic status, 53.7% resided in a dormitory, and 45.8% were born and raised in the Eastern Anatolia region (Table 1).

The distribution of the students, according to their menstrual characteristics, is given in Table 2. The average age of menarche was found to be 13.39 ± 1.23.

It was determined that the menstruation period was between 6-8 days in 59.4% of the students, the menstruation frequency of was between 21-34 days 82.4%, and menstruation was regular in 71.5%. It was found that 84.2% of the students did not have intermittent bleeding other than menstruation, 59.0% experienced pain from time to time during menstruation, and 84.4% did not have polycystic ovary syndrome (Table 2).

TABLE II.
DISTRIBUTION OF THE STUDENTS PARTICIPATING IN THE STUDY ACCORDING TO THEIR MENSTRUAL CHARECTERISTICS (n= 860)

	$\overline{\mathrm{X}}$ ± SS	
Menarş Age (years)	13.39± 1.23 (min:9, max:18)	
	n	**%**
Menstruation time (days)		
3-5 days	322	37.4
6-8 days	511	59.4
9 ≥ days	27	3.2
Menstruation frequency (days)		
Less than 21 days	44	5.2
21-34 days	709	82.4
35 days and more	107	12.4
Menstruation Order		
Yes	615	71.5
No	245	28.5
Presence of Intermediate Bleeding Other Than Menstruation		
No way	724	84.2
It happens from time to time	83	9.7
Happens occasionally	53	6.1
Presence of Pain During Menstruation		
No way	81	9.4
It happens from time to time	507	59.0
Happens occasionally	272	31.6
Presence of Polycystic Ovary Syndrome		
Yes	134	15.6
No	726	84.4
Total	**860**	**100.0**

The distribution of the participants, according to their PMS experience, is shown in Table 3. It was found that 68.1% of the students experienced PMS, and 31.9% did not experience PMS (Table 3).

The distribution of the mean lowest and highest total scores of PMSS and subscales is given in Table 4. We found that the mean score of depressive affection was 19.02 ± 6.12, irritability 14.87 ± 5.27, depressive thoughts 18.26 ± 7.17, pain 8.93 ± 3.24, appetite changes 9.34 ± 3.54, sleep changes 8.43 ± 3.45, and distension was 9.30 ± 3.67. The mean total PMSS score was found to be 126.69 ± 35.8 (Table 4).

TABLE III.
DISTRIBUTION OF PARTICIPANTS BY PMS LIVING STATUS (N=860)

PMSS	n	%
Yes	586	68.1
No	274	31.9
Total	860	100.0

TABLE IV.
DISTRIBUTION OF THE LOWEST-HIGHEST SCORES THAT CAN BE OBTAINED FROM PMSS TOTAL AND SUB-DIMENSIONS, AND THE LOWEST-HIGHEST SCORES AND AVERAGE SCORES OF THE PARTICIPANTS (N=860)

		Obtainable		Received	
PMSS	X ± SS	Min-Max		Min- Max	
Depressed affect	21.46±6.93	7	35	7	35
Anxiety	17.04 ±6.65	7	35	7	35
Tiredness	19.02±6.12	6	30	6	30
Irritability	14.87±5.27	5	25	5	25
Depressive thoughts	18.26±7.17	7	35	7	35
pain	8.93±3.24	3	15	3	15
Change in appetite	9.34±3.54	3	15	3	15
Sleep change	8.43±3.45	3	15	3	15
Swelling	9.30±3.67	3	15	3	15
PMSS Total	**126.69±35.8**	**44**	**220**	**44**	**220**

The significance values of the factors affecting PMS are given in Table 5. Correct classification rates of the created MCA ANN model for PMS were 63.2% in the training data set and 63.0% in the test data set. When the significance values of the variables are examined; it was found that the duration of active Internet use (phone, tablet, computer, etc.) was the most influential factor on PMS and the economic status of the student was the least influential factor (Table 5).

4. DISCUSSION

Premenstrual syndrome is a common and essential condition affecting most women of the reproductive age, and the number of women suffering from PMS is increasing [13,14]. It has been reported that PMS negatively affects women's self-confidence and social relations, causes workforce loss and decrease in work productivity, increases the tendency to crime and domestic fights, causes mothers to behave negatively towards their children and negatively affects the girls especially in terms of attendance to the classes. The etiology of PMS is not fully understood yet; however, it has been associated with diet, lifestyle, and hormonal and neuronal imbalance [15]. In this study, which was conducted to predict the factors associated with premenstrual syndrome, by using the ANN model, it was found that the duration (years) of active use of the Internet (phone, tablet, computer, etc.) was the most influential factor. Similarly, *Ko et al.,* found that the probability of having Internet use disorder was approximately five times higher in women with the premenstrual dysphoric disorder compared to those without [16].

TABLE V.
SIGNIFICANCE VALUES OF FACTORS AFFECTING PMS

Factors affecting PMS	**Significance Value**
The duration of active use of the internet (phone, tablet, computer, etc.) (years)	0.112
Age (years)	0.108
Age of menarche (years)	0.107
Height (cm)	0.100
Weight (kg)	0.100
The number of pads he changed on the busiest day of menstruation (pieces)	0.077
How much of his day he spends on the internet (phone, tablet, computer) (hours)	0.050
Frequency of menstruation	0.046
Menstruation time	0.045
Polycystic ovary diagnosis	0.044
The geographical region where he was born and raised	0.044
Menstruation scheme	0.033
Family type	0.032
Place of residence	0.028
Pain during menstruation	0.025
Breakthrough bleeding other than menstruation	0.025
Economical situation	0.023

In this study, we found that age was the second factor affecting PMS (Table 5). In the literature, studies are reporting that the age affected PMS symptoms, similar to our study [17,18]. In some studies, it has been found that the prevalence of PMS increased as the age increased [19,20]. In a study, conducted by *Chayachinda et al.*, in Thailand, it was found that the prevalence of PMS differs according to age, and it was seen in 70.8% of the women aged 30 and younger [21]. Similarly, in a study conducted in Manisa, it was found that PMS symptoms are more common in young people [22]. In the study by *Tschudin et al.*, the prevalence of PMS was found to be higher in women aged 35-44 [23]. On the other hand, in the study conducted in Virginia, it was found that women between the ages of 35-44 had PMS less frequently compared to younger women, and the highest prevalence was between the ages of 25 and 34 [24]. In addition, in the literature, there are studies suggesting that there is no significant relationship between PMS and age [14, 25-30].

In our study, the age of menarche was found to be the third factor affecting PMS (Table 5). In the literature, it has been reported that there was a relationship between PMS and the age of menarche and the possibility of PMS increases as the age of menarche decreases [24,25, 31-36]. In the study of *Selçuk et al.,* the average PMS score was found to be higher in those with menarche age 14 and above, although it was not statistically significant [37]. On the other hand, studies are reporting that there was no relationship between the age of menarche and experiencing PMS [28,30,38-42].

Obesity, which is an important problem today, is also an essential major risk factor for PMS. *Masho et al.*, reported that obese women (BMI≥30) had approximately three times the risk of PMS compared to non-obese women [43]. Similarly, in the study of *Özmermer*, it was observed that the prevalence of PMS increases as the weight increases [19]. On the other hand, in *Şahin's* study, the frequency of PMS was found to be lower in overweight / obese students [44]. The risk of having PMS is lower for women with a BMI in the normal range. The increase in BMI causes an increase of PMS symptoms, and this situation continues to increase as a vicious circle [29]. In another study, it was found that there is a strong relationship between BMI and risk of PMS, and 1 kg / m increase in BMI is associated with a significant increase of 3% in PMS risk [45]. When relevant studies are evaluated, as BMI values increase, the frequency of PMS symptoms also increases [14,29,46]. On the other hand, studies are suggesting that BMI does not affect PMS [41,47-49]. In our study, by using the ANN model, we found that weight and height were associated factors in BMS (Table 5).

We found that the time spent on the internet during the day (hours) is another factor influencing PMS (Table 5). Similar to our results, Yoshimi et al., in their study, reported that women with PMS spent a significantly longer time on the internet compared to women without PMS [36].

In our study, features such as the number of pads changed on the heaviest day of menstruation, frequency, duration, the pattern of menstruation, presence of polycystic ovary, and non-menstrual intermittent bleeding were found to be other factors affecting PMS. In the study of *Aşçı et al.*, No significant difference was found between the students with and without PMS in terms of menstruation frequency and duration [41]. In the study of *Çitil and Kaya*, it was reported that there was no statistically significant relationship between menstruation period and PMS [14]. In this study, the pain was one of the most common symptoms experienced during the menstruation. In our study, we also found that pain during menstruation is one of the factors affecting PMS (Table 5). In the literature, it was reported that PMS was significantly higher in those who had pain during menstruation [21,39,41,47,50,51].

The geographical region where the students were born and raised, family type and place of residence are other factors that have smaller effects on PMS (Table 5). In the study of *Aşçı et al.*, no significant difference was found in terms of the place where the student resided during university education. In the study of *Kırcan et al*., it was found that 62.9% of the students living in the city center and 76.2% of the students living in the rural regions had PMS symptoms, but the difference was not statistically significant. In the study of *Selçuk et al.,* it was found that students living in dormitories are at 4.19 times higher risk of PMS compared to students living at home [37].

In our study, it was determined that the economic status of the student was the least effective factor in PMS. This finding suggests that in training aimed at preventing PMS, all women should be addressed regardless of their income level differences. In the literature, some studies did not find a statistically significant relationship between economic status and PMS, similar to our study [7,30,41,49]. On the other hand, there are studies that found that women experienced PMS symptoms more intensely as their income decreased [14,21].

5. CONCLUSION

In this study, we determined the factors associated with PMS in female students and the level of significance of these factors by using the proposed ANN model. Thus, due to the limited number of studies using ANN models in PMS, it is anticipated that the results of this study will create an infrastructure for other similar studies. In the literature, it has been reported that factors such as staying in a dormitory, gynecological diseases, fast-food consumption, age, weight, income level, smoking, and history of PMS in the mother and sister are associated with PMS [19,37,52]. According to the findings of the ANN model designed in this study, the three most important factors associated with PMS were found to be the duration of active internet use, present age and age of menarche were. Considering the high prevalence of PMS, the uncertain etiology, and its effects on the lifestyle of the woman, we recommend using artificial intelligence models that can include further factors, in larger samples.

ACKNOWLEDGEMENT

We would like to thank the female students who participated and completed this questionnaire.

This research won the oral presentation first prize Award in the 1st International / 4th Traditional Health Professionals Days (14-16 November 2019).

REFERENCES

[1] M. Delara, F. Ghofranipour,P. Azadfallah , S.S. Tavafian, A. Kazemnejad, A.Montazeri, “Health Relates Quality of Life Among Adolescents with Premenstrual disorders: A Cross Sectional Study”, Health and Quality of Life Outcomes,Vol.10,No.1, 2012.

[2] P. Palas Karaca, and N. Kızılkaya Beji, “Premenstrual Sendromunun Tanı ve Tedavisinde Kanıt Temelli Yaklaşımlar ve Hemşirelik Bakımı”, Balıkesir Sağlık Bilimleri Dergisi, Vol.4,No.3, 2015, pp.178-186.

[3] Royal College of Obstetricians and Gynaecologists Green-Top Guideline No 48. In: Panay N, ed. The Management Of Premenstrual Syndrome 2016, DOI: 10.1111/1471-0528.14260.

[4] M.A. Direkvand, K. Sayehmiri, A. Delpisheh, S. Kaikhavandi, “Epidemiology of Premenstrual Syndrome (PMS)- A Systematic Review and MetaAnalysis Study”, Journal of Clinical and Diagnostic Research. Vol.8,No.2,2014,pp.106-109.

[5] B. Haydardedeoğlu, and G. Çoban, “Premenstrüel Sendrom, Premenstrüel Disforik Bozukluk, Dismenore”, A. Ayhan, G. Bozdağ, Ö. Deren, T. Durukan, İ. Esinler, S. Günalp, T. Gürgan, Ö. Harmanlı, L.S. Önderoğlu, Ö. Özyüncü, C. Salman, H. Yaralı, K. Yüce, editörler.Temel Kadın Hastalıkları ve Doğum Bilgisi. 3.Baskı. Ankara: Güneş Tıp Kitabevleri, 2014, pp.216-217.

[6] S. Öztürk, A. Karaca, “Premenstrual Sendrom ve Sağlıklı Yaşam Biçimi Davranışlarına İlişkin Ebe ve Hemşirenin Rolü”, Balıkesir Sağlık Bilimleri Dergisi, Vol.8, No.2, 2019, pp.105-110.

[7] Z. Ölçer, N. Bakır, E. Aslan, “Üniversite Öğrencilerinde Kişilik Özelliklerinin Premenstrual Şikâyetlere Etkisi”, Düzce Üniversitesi Sağlık Bilimleri Enstitüsü Dergisi, Vol.7,No.1, 2017, pp.30-37.

[8] B. Ataseven, “Yapay Sinir Ağları ile Öngörü Modellemesi”, Öneri Dergisi, Vol.10,No.39, 2013, pp.101-115.

[9] B. Yılmaz, “Maliyet Fonksiyonun Belirlenmesinde Yapay Sinir Ağı Modellerinin Kullanımı”, The Journal of Accounting and Finance, Ağustos Özel Sayı ,2019,pp.329-344.

[10] Y. Gönül, Ş. Ulu, A. Bucak, A. Bilir, “Yapay Sinir Ağları ve Klinik Araştırmalarda Kullanımı”, Genel Tıp Dergisi, Vol. 25, 2015, pp.104-111.

[11] B. Gençdoğan, “Premenstrual sendrom için yeni bir ölçek”, Türkiye’de Psikiyatri, Vol.8,No.2, 2006, pp.82-87.

[12] K.K. Çevik, E.Dandıl, “Yapay Sinir Ağları İçin Net Platformunda Görsel Bir Eğitim Yazılımının Geliştirilmesi”, Bilişim Teknolojileri Dergisi, Vol.5, No.1, 2012, pp.19-28.

[13] N. Buddhabunyakan, S. Kaewrudee, C. Chongsomchai, S. Soontrapa, W. Somboonporn, J. Sothornwit, “Premenstrual syndrome (PMS) among high school students”, Int J Womens Health, Vol.9, 2017, pp.501-505.

[14] E.T. Çitil, N. Kaya, “Ebelik Öğrencilerinin Bazı Değişkenler Yönünden Premenstruel Belirtileri”, Düzce Üniversitesi Sağlık Bilimleri Enstitüsü Dergisi, Vol.8,No.3,2018, pp. 133-141.

[15] M.Döll M. (2009). Das prämenstruelle Syndrom. Wirksamkeit von Vitex agnus castus [The premenstrual syndrome: effectiveness of Vitex agnus castus]. Medizinische Monatsschrift fur Pharmazeuten, 32(5), 186–191.

[16] C.H. Ko,C.F. Yen, C.Y. Long, C.S. Chen, T.H. Huang, J.Y.Yen, “The association between premenstrual dysphoric disorder and internet use disorder”, Women Health, Vol.54,No.3, 2014, pp.245-261.

[17] S.H. Hamaideh, S.A. Al-Ashram, H. Al-Modallal, “Premenstrual Syndrome and Premenstrual Dysphoric Disorder among Jordian Women”, Journal of Psychiatric and Mental Health Nursing, Vol.21,No.1, 2014, pp.60-68.

[18] M. Tadakawa, T. Takeda, Y. Monma, S. Koga, N. Yaegashi, “The Prevalence and Risk Factors of School Absenteeism Due to Premenstrual Disorders in Japanese High School Students”, Biopsychosocial Medicine, Vol.10,No.13,2016, pp.2-7.

[19] T. Özmermer, “Şanlıurfa'da Üniversite Öğrencilerinde Premenstrual Sendrom Sıklığı, Baş Etme Yolları ve Etkileyen Faktörler”, Harran Üniversitesi Sağlık Bilimleri Enstitüsü,Yüksek Lisans Tezi, Şanlı Urfa,2017.

[20] S. Oruç Koltan,Y. Uyar, “Premenstrüel disforik bozuklukta semptomatolojinin adet döngüsüyle ilişkisi”, Anadolu Psikiyatri Dergisi, Vol.2,No.1,2001, pp. 15 - 20.

[21] C. Chayachinda, M. Rattanachaiyanont, S. Phattharayuttawat, S. Kooptiwoot, “Premenstrual syndrome in Thai nurses”, Journal of psychosomatic obstetrics and gynaecology, Vol.29,No.3,2008, pp.199–205.

[22] H. Adıgüzel, E.O. Taşkın, A. Esen Danacı, “Manisa İlinde Premenstruel Sendrom Belirti Örüntüsü ve Belirti Yaygınlığının Araştırılması”, Türk Psikiyatri Dergisi, Vol.18,2007, pp.215-222

[23] S.Tschudin, P.C. Bertea, E. Zemp, “Prevalence and predictors of premenstrual syndrome and premenstrual dysphoric disorder in a population based sample”, Arch Womens Ment Health, Vol.13,No.6, 2010, pp.485-94.

[24] P.A. Deuster, T. Adera, J. South-Paul, “Biological, social, and behavioral factors associated with premenstrual syndrome”, Arch Fam Med, Vol.8, 1999, pp.122–128.

[25] S. Poyrazoğlu, “Kayseri İl Merkezi’nde Yaşayan 15-49 Yaş Arası Kadınlarda Premenstruel Sendrom Prevalansı ve Etkileyen Faktörler”,

Erciyes Üniversitesi Tıp Fakültesi Halk Sağlığı Anabilim Dalı, Tıpta Uzmanlık Tezi, Kayseri, 2010.
[26] B. Önal, "Premenstrüel sendromda risk faktörleri ve tedavi arama davranışının araştırılması", Dokuz Eylül Üniversitesi Tıp Fakültesi Aile Hekimliği Anabilim Dalı, Uzmanlık Tezi, İzmir, 2011.
[27] İ. Arpacı, "Bir Vakıf Üniversitesindeki Kız Öğrencilerin Premenstrual Sendrom Yaşama Durumlarının ve Etki Eden Faktörlerin Değerlendirilmesi", Hemşirelik Anabilim Dalı Tezli Yüksek Lisans Programı, Yüksek Lisans Tezi, Gaziantep, 2018.
[28] N. Akmalı, N. Özerdoğan, E. Gürsoy, "Bir devlet hastanesi'nde çalışan üreme çağındaki kadınlarda premenstrual sendrom prevalansı, ilişkili faktörler ve yaşam kalitesine etkisi", Mersin Üniv Sağlık Bilim Derg , Vol.13,No.1,2020, pp.63-74
[29] M. F. Albayrak, "20-45 Yaş Arası Kadınlarda Premenstruel Sendromun Beslenme Durumuna Etkisinin Saptanması", Marmara Üniversitesi Sağlık Bilimleri Enstitüsü, Beslenme ve Diyetetik Anabilim Dalı,Yüksek Lisans Tezi, İstanbul, 2019.
[30] Z. Aydınay, "Üniversite Öğrencisi Genç Kızlarda Premenstrüel Sendrom ile Mental İyi Oluşluk Düzeyi Arasındaki İlişki", Cmumhuriyet Üniversitesi, Doğum ve Kadın Hastalıkları Hemşireliği Anabilim Dalı, Yüksek Lisans Tezi, Sivas, 2019.
[31] B. Demir, L.Y. Algül, E.S.G. Güven, "Sağlık çalışanlarında premenstrüel sendromun insidansı ve etkileyen faktörlerin araştırılması", Türk Jinekoloji ve Obstetrik Derneği Dergisi,Vol.3,2006,pp.262-270.
[32] C.M. Silva, D.P. Gigante, G.C. Minten, "Premenstrual symptoms and syndrome according to age at menarche in a 1982 birth cohort in southern Brazil", Cad Saude Publica, Vol.24, 2008, pp.835-44.
[33] M.H. Balaha,M.A. Amr, M .Saleh Al Moghannum, N. Saab Al Muhaidab, "The phenomenology of premenstrual syndrome in female medical students: a cross sectional study", Pan Afr Med J, Vol.5,No.4,2010.
[34] A.C. Eke, J.C. Akabuike,K. Maduekwe, "Predictors of premenstrual syndrome among Nigerian university students", Int J Gynaecol Obstet, Vol.112,No.1, 2011, pp.63-64.
[35] A. Göker, A.B. Ülkümen, F. Aktenk, N. İkiz, "Premenstrual Syndrome in Turkish Medical Students and Their Quality of Life", Journal of Obstetrics and Gynaecology, Vol.35, No.3, 2015, pp. 275-278.
[36] K. Yoshimi, M. Shiina, T. Takeda, "Lifestyle Factors Associated with Premenstrual Syndrome: A Cross-sectional Study of Japanese High School Students", J Pediatr Adolesc Gynecol, Vol.32,No.6, 2019, pp.590-595.
[37] K. Tarı Selçuk, D. Avcı, F. Alp Yılmaz, "Hemşirelik Öğrencilerinde Premenstrual Sendrom Prevalansı ve Etkileyen Etmenler", Psikiyatri Hemşireliği Dergisi,Vol.5,No.2, 2014, pp.98-103.
[38] M. Karavuş, D. Cebeci, M. Bakırcı, O. Hayran, "Üniversite öğrencilerinde premenstrual sendrom",Türkiye Klinikleri Tıp Bilimleri,Vol.17, 1997, pp.184-90
[39] N. Erbil, N. Bölükbaş, S. Tolan, F. Uysal, "Evli kadınlarda premenstrual sendrom görülme durumu ve etkileyen faktörlerin belirlenmesi" Uluslararası İnsan Bilimleri Dergisi, Vol.8,No.1,2011, pp.827-838
[40] N.K. Chandraratn, N.S. Gunawardena, "Premenstrual syndrome: the experience from a sample of Sri Lankan adolescents", J Pediatr Adolesc Gynecol, Vol.24, 2011, pp.304-310.
[41] Ö. Aşcı, H. Süt, F. Gökdemir, "Üniversite Öğrencilerinde Premenstrüel Sendrom Prevalansı ve Risk Faktörleri" DEUHFED, Vol.9,No.3,2016, pp.79-87.
[42] N. Elkin, "Sağlık Bilimleri Yüksekokulu Öğrencilerinde Premenstrual Sendrom Görülme Sıklığı", Adıyaman Üniv Sağlık Bilim Derg, Vol.1,No.2,2015, pp.94-110
[43] S.W. Masho, T. Adera, J. South-Paul, "Obesity as a risk factor for premenstrual syndrome", Journal of psychosomatic obstetrics and gynaecology, Vol.26,No.1, 2005, pp.33–39.
[44] S. Şahin, K. Ozdemir, A. Unsal, "Evaluation of premenstrual syndrome and quality of life in university students" The Journal of the Pakistan Medical Association, Vol.64,No.8,2014, 915–922.
[45] E.R Bertone-Johnson, S.E. Hankinson, W.C Willett, S.R. Johnson, J.E. Manson, "Adiposity and the development of premenstrual syndrome" J Womens Health, Vol.19,2010, pp.1955–1962
[46] K. Işgın, "Premenstrual Sendromda Beslenme Durumu ve Yeme Tutumunun Değerlendirilmesi", H.Ü. Sağlık Bilimleri Enstitüsü, Yüksek Lisans Tezi, Ankara , 2014.
[47] G. Pınar, M. Colak, E. Oksuz, "Premenstrual Syndrome in Turkish college students and its effects on life quality", Sexual & reproductive healthcare : official journal of the Swedish Association of Midwives, Vol.2, No.1, 2011, 21–27.
[48] A.H. Alpaslan, K. Avcı, N. Soylu, H.L. Taş, "Association between Premenstrual Syndrome and Alexithymia among Turkish University Students" Gynecological Endocrinology, Vol.30,No.5, 2014, pp. 377-80.
[49] H. Doğan, "Üniversite Öğrencilerinde Premenstrual Sendrom Görülme Sıklığı Ve Etkileyen Faktörler" Karabük Üniversitesi Sağlık Bilimleri Enstitüsü Ebelik Anabilim Dalı, Yüksek Lisans Tezi, Karabük, 2018.
[50] N. Kırcan, F. Ergin, F. Adana, H. Arslantaş, "Hemşirelik Öğrencilerinde Premenstrüel Sendrom Prevalansı ve Yaşam Kalitesi İle İlişkisi" ADÜ Tıp Fakültesi Dergisi, Vol.13,No.1,2012, pp. 19-25.
[51] S. Kısa, S. Zeyneloğlu, N. Güler, "Üniversite Öğrencelerinde Premenstrual Sendrom Görülme Sıklığı ve Etkileyen Faktörler" Gümüşhane Üniversitesi Sağlık Bilimleri Dergisi,Vol.1,No.4, 2012.
[52] M.S.Hashim, A.A. Obaideen, H.A. Jahrami, H. Radwan, H.J. Hamad, A.A. Owais, L.G. Alardah,S. Qiblawi, N. Al-Yateem, M.A.E. Faris, "Premenstrual Syndrome Is Associated with Dietary and Lifestyle Behaviors among University Students: A Cross-Sectional Study from Sharjah, UAE", Nutrients, Vol.18,No.8, 2019.

BIOGRAPHIES

Yeşim Aksoy Derya obtained her BSc degree in midwifery from Inonu University in 2005. She received MSc. Diploma in midwifery from the Sivas University in 2007, and PhD degrees in the Graduate Department of Obstetrics and Gynecology Nursing of Atatürk University in 2012. She joined İnönü University Midwifery Department as a lecturer in 2007. She is currently an Associate Professor PhD of Inonu University Faculty of Health Sciences, Department of Midwifery, Malatya, Turkey. She is active in teaching and researching about midwifery.

Cemil Çolak obtained his BSc.Degree in statistics from Ondokuz Mayıs University in 1999. He received MSc. Diploma in statistics from the Inonu University in 2001, and PhD degrees in the Graduate Department of Biostatistics and Medical Informatics of Ankara University in 2005. He accepted as a Postdoctoral Researcher by the department of biostatistics and medical informatics of Inonu University in 2007. His research interests are cognitive systems, data mining, reliability, and biomedical system, and genetics, and bioengineering. In 2016, he joined the Department of Biostatistics and Medical Informatics at Inonu University as a professor, where he is presently a professor. He is active in teaching and research in the general image processing and data mining modelling, analysis.

Burcu Kayhan Tetik obtained her BSc degree in a doctor from Erciyes University in 2002 and Ass. Prof. Dr. Department of Family Medicine from Ankara University in 2013. She is currently an Associate Professor PhD of Inonu University School of Medicine, Department of Family Medicine, Malatya, Turkey. She is active in teaching and researching about family medicine.

Ayşe Nur Yılmaz obtained his BSc degree in midwifery from Fırat University in 2013. She received MSc. Diploma in midwifery from the Inonu University in 2018, and continues her Ph.D education at İnönü University midwifery department. She joined Fırat University Midwifery Department in 2014 as a research assistant and continues at the same place. She is active in teaching and researching about midwifery.

Zeliha Özşahin obtained her BSc degree in midwifery from Adıyaaman University in 2008. She received MSc. Diploma in midwifery from the Inonu University in 2017, and Ph.D. degrees in the Graduate Department of midwifery of Inonu University in 2020. She joined Siirt University Midwifery Department in 2016 as a research assistant. She is currently an Assistant Professor PhD of Inonu University School of Medcine, Department of Family Medicine, Malatya, Turkey. She is active in teaching and researching about midwifery.

EFFECTS OF REAL AND IMAGINARY HAND MOVEMENTS ON EEG

O. Akgun

Abstract— Our brain is one of the most complex structures in the known universe, and no matter how we want to make it simple, the way of work of it is quite complicated. In this study, the brain's structure and function that enable our movements were briefly discussed. Electroencephalography (EEG) is a non-invasive method of examination in which spontaneous electrical activity of the brain is recorded through electrodes. This examination reflects the current functional state of the brain rather than its structural properties. This study is based on an analysis of EEG signals obtained through a subject raising his left and right arm forward and then imagining him raising his left and right arm forward. First, the change of signals by the time was obtained, then the amplitude spectra of EEG signals were reached by applying the Fast Fourier Transform. Finally, the power spectrum analysis of signals was performed using the Welch method.

Keywords— *Hand movements, EEG, Histogram, FFT, Welch method*

1. INTRODUCTION

ALMOST every relationship we have with the outside world involves more or less movement. Mimic movements emerging by contraction and relaxation of the facial muscles in a certain order, our conversation occurring with the work of the tongue, larynx, and mouth muscles, and making gestures with the work of the hand and arm muscles are often quite complex movements that we are not even aware of.

In medicine, the ability to move is called motor function. Even reaching out to a cup and taking it requires making a large number of calculations in our brain, creating appropriate commands by integrating different data, and then constantly fine-tuning of commands with sensory feedbacks.

The main part of the brain associated with motor function is called the primary motor cortex, or M1 area. M1 is located in the frontal lobe, along the brain curl called the precentral gyrus. Its task is to produce neural impulses that control the performing of movement. A large cortical area is required to control the complex movements of the hands and fingers, while the torso and legs, whose movements are simpler, are represented in a smaller area. In the brain, other movement-related parts are secondary motor areas. Movement signals from the brain reach the motor neurons in the part of the spinal cord called the anterior horn. The anterior horn motor neurons can receive thousands of impulses from the cortical and subcortical motor regions and intermediate neurons in the spinal cord. The impulse coming out of the front horn motor neuron goes directly to the muscle fibers and ensures them to contract [1-4]. Recording of bioelectrical activity of the brain by electroencephalography (EEG) with electrodes inserted into the scalp was first discovered by Hans Berger in 1929. For this examination, which has been also accompanied with IT technical developments over the years, amplification of very low-amplitude waves at the microvolt level and application of the filter systems are required. While all traces have been printed out to paper output, digital EEG devices are becoming increasingly common today. Digital EEG has many advantages, such as assessment of the trace more detailed and removal of artifacts [5-8]. The postsynaptic potentials that constitutes the source of the EEG are collected in the cortex, and by spreading to the scalp through the structures that surround the brain, they are recorded from the scalp with metal electrodes. The location of each electrode covered with a conductive substance is determined by standard measurements made from nasion, inion, and right and left preauricular points, and placed according to the International 10-20 system (Figure 1) [9,10]. The EEG shows potentials at various frequencies and amplitudes. Although the basic activity varies depending on age, an activity at a frequency of 8-12 Hz is observed in the parieto-occipital regions when a normal adult is awake and eyes are closed; this activity is called alpha activity. The alpha activity disappears or is suppressed when the eyes open. Beta activity is a rhythm that is evident in frontal and central regions at a frequency of 13-25 Hz. Pathological findings that may be encountered in EEG are divided into two main groups as nonspecific slow waves and epileptiform activity. The slow wave activity is grouped as theta (4-7Hz) and delta (1-3Hz) [9, 11-13].

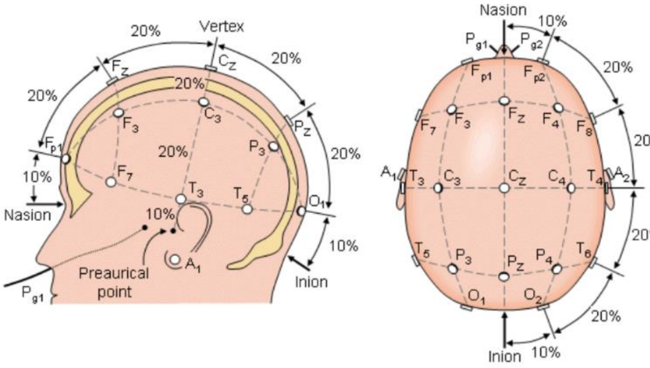


Fig.2. Map for head placement of EEG electrodes

2. APPLICATION AND MATHEMATICAL BACKROUND

The EEG signals used in this study are taken from a site called github [14]. The data was recorded from a 21-year-old right-

Omer Akgun, is with Technology Faculty Marmara University, Istanbul, Turkey, (e-mail: oakgun@marmara.edu.tr).

Manuscript received Aug 19, 2020; accepted Oct 11, 2020.
Digital Object Identifier:

handed male subject by using 19 electrodes while his eyes were closed.

In this study, the signals received from the F4 electrode (Figure 1) were used. Recording was performed first by asking the subject to raise his left hand in the air, then it was carried out by asking the subject to imagine he was raising his left hand. The same process was also repeated for his right hand.
Signal analysis was started through obtaining the change of the signal by time. Since the change of EEG signals by time is quite complex, histograms of signals were obtained and evaluated together.

By applying Fast Fourier Transform (FFT) to the signals, their amplitude spectra were taken out. Then, using Welch from nonparametric methods, power spectral density estimation of signals was made.

In order to understand this method, the Barlett method on which it is based must be expressed first. The Barlett method is defined as the method of receiving the average periodogram. In contrast to standard periodograms, in the Barlett method, the variance of the periodogram is decreased by reducing its resolution. With this method, which is widely used in physics, engineering and mathematics, the power spectrum of multiple signal windows of "n" length is taken without the overlap process. By taking the averages of the power components corresponding to the same frequency on the entire signal, inference of spectral density is performed. The Discrete-Time Fourier Transformation of each window is taken. The Welch method, on the other hand, emerged as a result of the development of the Barlett method. Here, in addition to windowing, a certain amount of overlapping process is performed on each window. This overlapping process is performed to eliminate the losses caused by the window function while moving from the center of the signal to its edges [15, 16].

The Welch method is given in the following equations.

$$x_i(n) = x(n + iD),\ n = 0,1,2,\dots,(M-1) \qquad (1)$$

In Equation 1, x_i shows the data sequence, iD shows the starting point of the sequence, and M shows the generated data partitions. Equation 2 is used to calculate periodograms. Here, U is the normalization of power, and it is given in Equation 3.

$$P_x(f) = \frac{1}{MU}\left(\sum_{n=0}^{M-1} x_i\,(n) w(n) e^{-j2\pi f n}\right)^2 \qquad (2)$$

$$U = \frac{1}{M}\sum_{n=0}^{M-1} w^2(n) \qquad (3)$$

$w(n)$ is the window function. The mean of the modified periodograms gives Welch's power spectrum. The power spectrum is shown in Equation 4.

$$P_x^w = \frac{1}{L}\sum_{i=0}^{L-1} P_x\,(f) \qquad (4)$$

where L shows the length [17, 18].

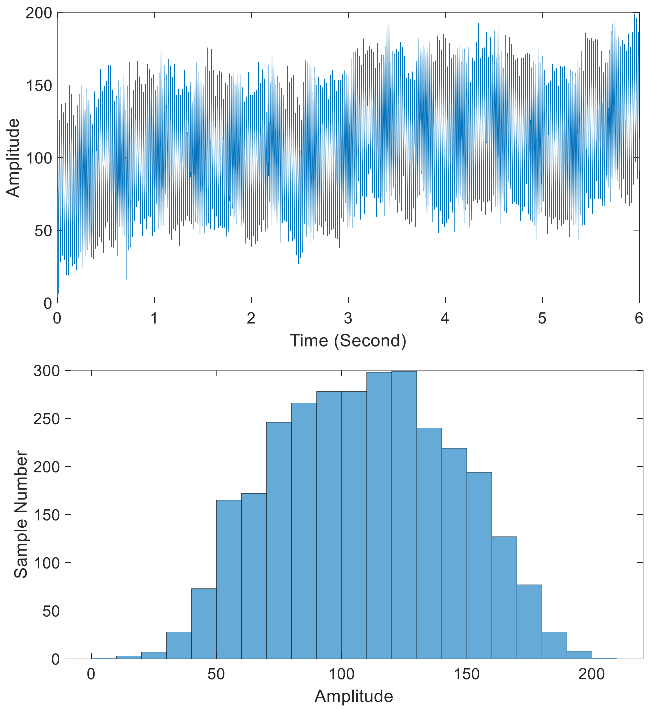


Fig.2. EEG time signal (top) and histogram (bottom) in case the left hand is raised forward.

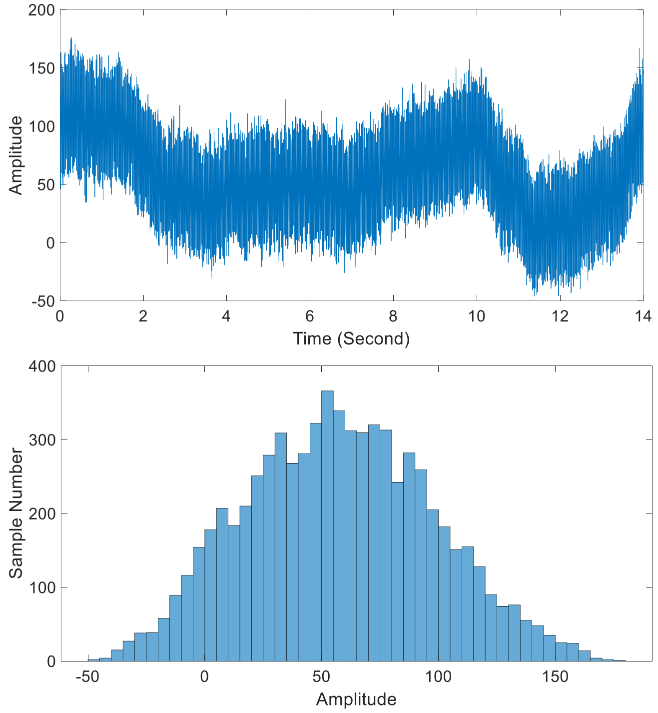


Fig.3. EEG time signal (top) and histogram (bottom) in case the left hand is imagined to be raised forward.

Although the time and histogram graphs of real and imaginary signals are similar in terms of character, the real signal receives values at positive alternant, while negative values are also observed in imaginary signal. In this sense, the positive high amplitude is more numerous in the actual signal (Figure 2, 3).

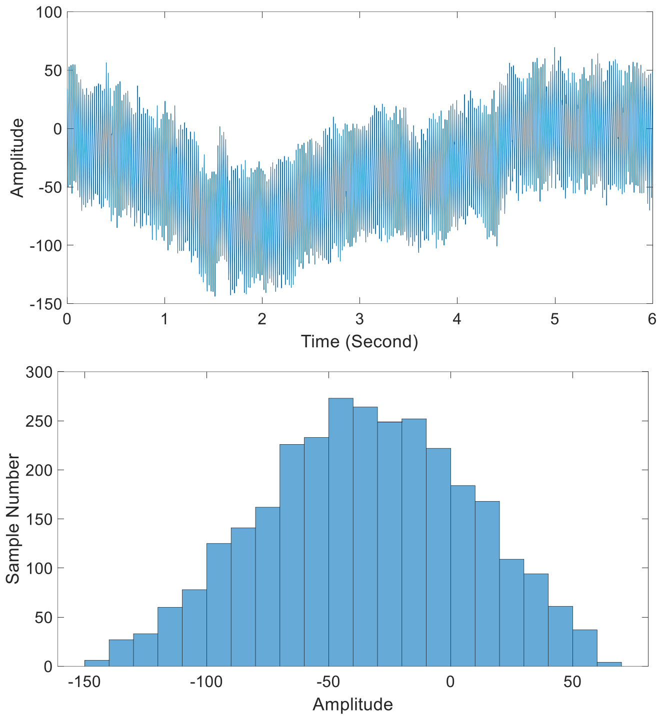


Fig.4 EEG time signal (top) and histogram (bottom) in case the right hand is raised forward.

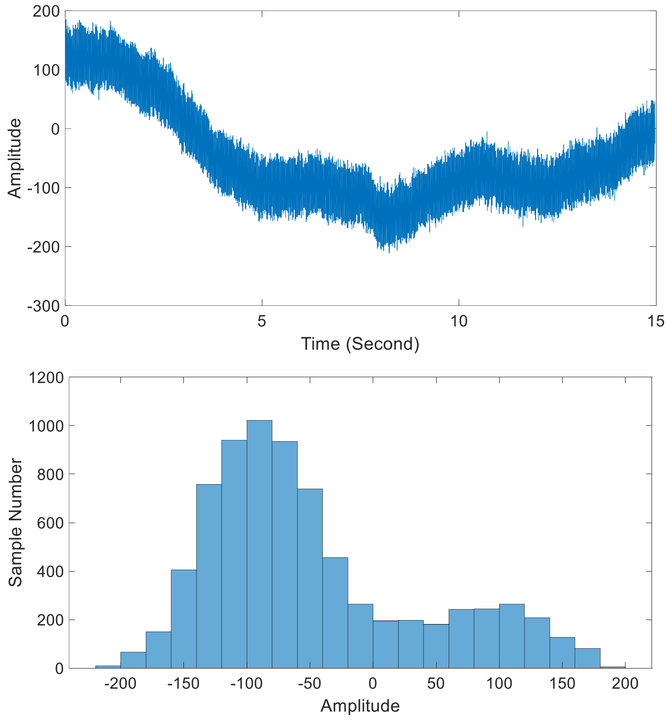


Fig.5 EEG time signal (top) and histogram (bottom) in case the right hand is imagined to be raised forward.

Considering that he is a right-handed subject, it attracts attention that there are more negative amplitudes both in reality and imagination. Especially in the imaginary right hand time change, the graph shows a decrease towards negative values (Figure 4, 5).

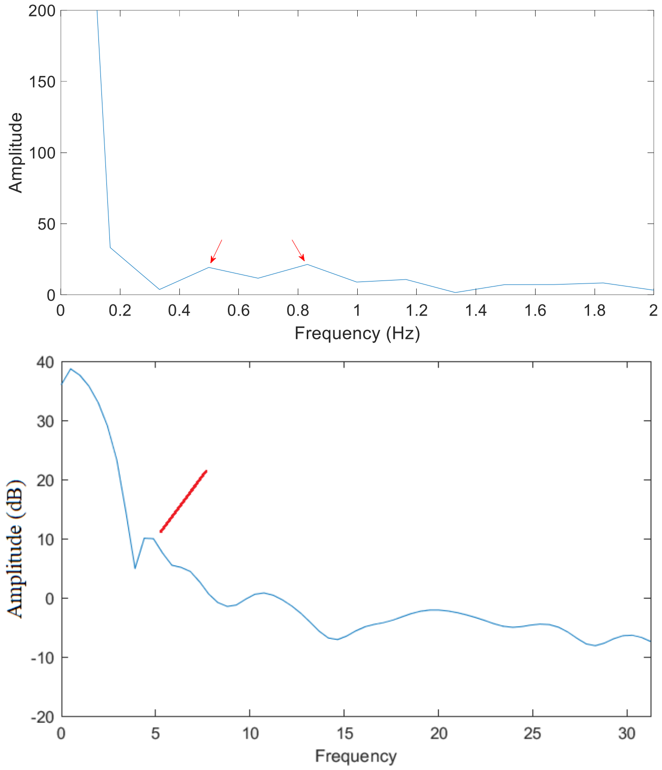


Fig.6 FFT graph (top) and power spectrum (bottom) of EGG in case the left hand is raised forward.

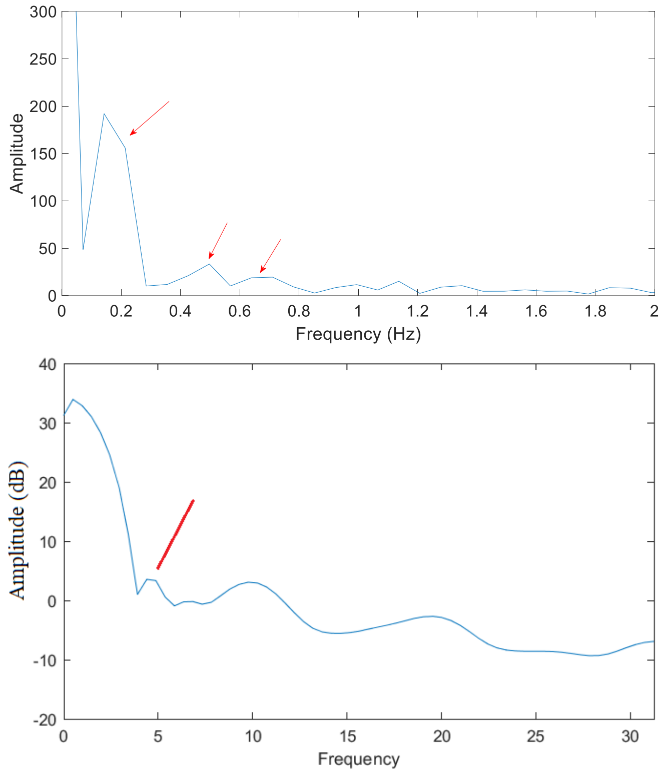


Fig.7 FFT graph (top) and power spectrum (bottom) of EGG in case the left hand is imagined to be raised forward.

In the real signal of the left hand, delta waves are attracting attention while the FFT moves towards 1Hz. In the power spectrum, the delta wave is manifesting itself with a peak of about 10 dB. Although the same waves are observed in the imaginary case, the amplitude of the delta wave in the power spectrum has decreased to 2dB.

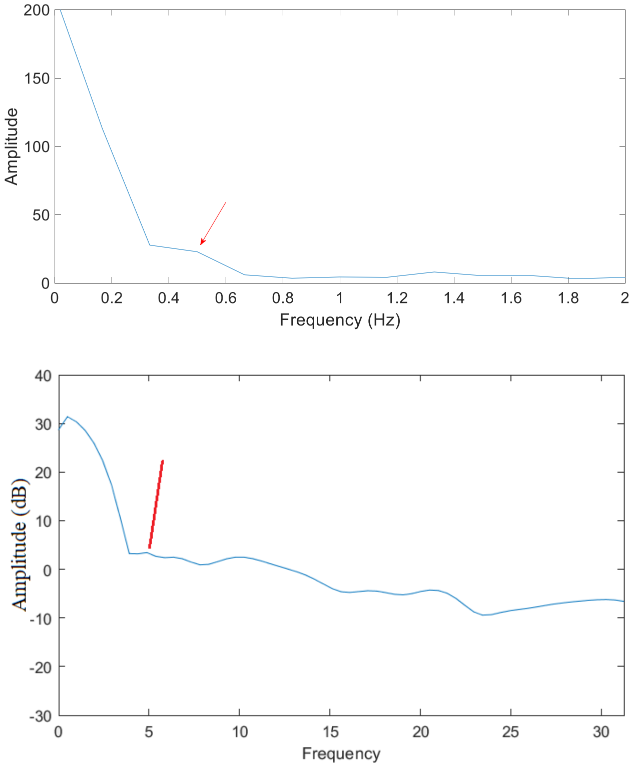


Fig.8 FFT graph (top) and power spectrum (bottom) of EGG in case the right hand is raised forward.

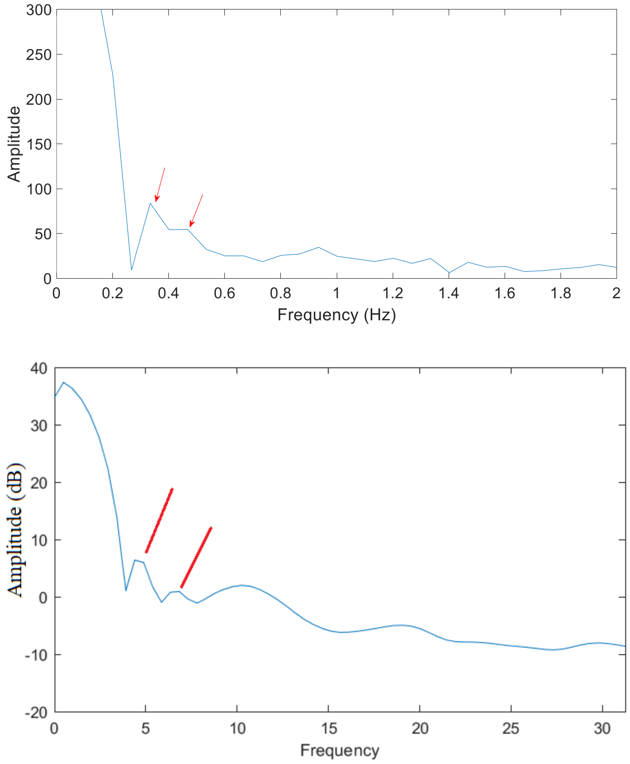


Fig.9 FFT graph (top) and power spectrum (bottom) of EGG in case the right hand is imagined to be raised forward.

The real signal of the right hand is noticed through delta waves that are at similar levels to the left hand but do not show much fluctuation. A fluctuation raising up to 75 units is observed in the FFT of the imaginary signal of the right hand. Compared to reality, there has been a rise. The same situation is seen clearly in the power spectrum.

3. CONCLUSIONS

The main part of the brain associated with motor function is called the primary motor cortex, or M1 area. This part is responsible for producing signals that ensures the performing of movement. Electroencephalography (EEG) is a non-invasive examination method in which spontaneous electrical activity of the brain is recorded through electrodes. This examination reflects the current functional state of the brain rather than its structural properties. In the EEG, the alpha wave oscillates at frequencies of 8-12 Hz, beta activity at 13-25 Hz, theta wave at 4-7 Hz, and delta wave at 1-3 Hz. Since the graph of the change of EEG by time showed a fairly intense signal density, it was more efficient to evaluate it together with the histogram. Because he was a right-handed subject, negative amplitudes were much more intense in both real and imagination signals related to the right hand compared to the left hand. In the frequency zone analysis, the primarily noticeable thing is the changes in delta activity. In the real case of the left hand, the delta wave in the power spectrum is 10 dB, while in the imaginary sign it falls to 2 dB. Because the subject was right-handed, the real signals of the right hand did not show much fluctuation though they were at a similar level. However, in the imaginary case, both a raising in the amplitude level and an increase in the fluctuation were observed. In conclusion, it can be said that we first design the movements we will make in our brain, and then perform them. This situation was revealed in this study by similar characters in the graphs between real and imaginary conditions.

REFERENCES

[1] https://www.doktorfizik.com/sinir-sistemi/norobilim/beyinde-hareket-merkezi-neresidir/, 12 July 2019
[2] A.M. Badakva, N.V. Miller, L.N. Zobova, “New Concepts on the Implementation of Motor and Cognitive Functions in the Brain: Facts and Hypotheses”, Human Physiology, Vol.46, No.3, 2020, pp.343-350.
[3] F. Zhou, Y. Liuye, Q. Zhiyu, “Research on Multi-parameter Visualization Technology of Brain Function Based on EEG”, 2019 International Conference on Medical Imaging Physics and Engineering (ICMIPE), 1-4 Nov, 2019.
[4] N.N. Dygalo, “Investigation of Brain Functions Using Genetically Encoded Tools”, Neuroscience and Behavioral Physiology, Vol.50, No.8, 2020, pp.1051-1056.
[5] E.N. Vanlı Yavuz, N. Bebek, “Place of Electroencephalography (EEG) in the Diagnosis and Treatment of Epilepsy”, Klinik Gelişim, pp. 35-38.
[6] M. Duraıraj, R. Jansı Ranı, “Mining and Extracting Emotional Information from Human-Based EEG: An Overview”, Productivity, Vol. 59, No. 4, 2019, pp.343-351.
[7] S. Sanei, J.A. Chambers, *EEG Signal Processing*, Wiley-Interscience; 1st edition, 2013.
[8] S. Yan, L.Y. Zhang, *EEG Signal Analysis and Classification*, Springer Nature Switzerland, 2020.
[9] B. Baykan, E. Altındağ, A.D. Elmalı, *Elektroensefalografi*, http://www.itfnoroloji.org/semi2/eeg.htm
[10] 10. www.researchgate.net/publication/321094865
[11] A. Riyadi, Munawar, T. Prakoso, ; F.O. Whaillan, ; D.W. Marcelinus, A. Hidayatno, “Classification of EEG-based Brain Waves for Motor Imagery using Support Vector Machine”, 2019 International Conference on

Electrical Engineering and Computer Science (ICECOS), 422-425 Oct, 2019.

[12] O.S. Sushkova, A.A. Morozov, A.V. Gabova, “A method of analysis of EEG wave trains in early stages of Parkinson's disease”, 2016 International Conference on Bioinformatics and Systems Biology (BSB),

[13] A. Quintero-Rincón, V. Muro, C. D’Giano, J. Prendes, H. Batatia, “Statistical Model-Based Classification to Detect Patient-Specific Spike-and-Wave in EEG Signals”, Computers, Vol 9, No 85, pp. 85.

[14] https://github.com/meagmohit/EEG-Datasets

[15] www.baskent.edu.tr › ~tbayrak › LAB7

[16] F. Castanié (Editor), *Digital Spectral Analysis: Parametric, Non-Parametric and Advanced Methods*, Wiley, 2011.

[17] M. Tosun, M. Erginli, O. Kasım, B. Uğraş, S. Tanrıverdi, T. Kavak, “By Using EEG data, Classification of the Physical Hand Movements and the Imagination of these movements via Artificial Neural Networks”, Sakarya Unıversıty Journal Of Computer And Informatıon Scıences, Vol.1, No.2, 2018.

[18] K.A. Budunova, V.F. Kravchenko, D. Churikov, “Application of the Family of Kravchenko-Rvachev Atomic Weight Functions (Windows) in Welch Method EEG Power Spectral Density Estimation”, 2019 PhotonIcs & Electromagnetics Research Symposium - Spring (PIERS-Spring). 2019, pp.500-506.

BIOGRAPHIES

Omer AKGUN, is an assistant professor at the, Department of Computer Engineering, Technology Faculty, Marmara University. He received his first Ph.D. in the Communication Engineering in 2009 from Yildiz Technical University and the second Ph.D. in the Electronic and Communication Education Department in 2011 from Marmara University. His current research interests are signal processing, biomedical signal processing, signal modelling and communication systems.

PERFORMANCE EVALUATION OF THE DEEP LEARNING MODELS IN THE CLASSIFICATION OF HEART ATTACK AND DETERMINATION OF RELATED FACTORS

Z. Kucukakcali, I. Balikci Cicek, and E. Guldogan

Abstract*— *Aim: The aim of this study is to classify the condition of having a heart attack and determine the related factors by applying the deep learning method, one of the machine learning methods, on the open-access data set.

Materials and Methods: In this study, deep learning method was applied to an open-access data set named “Health care: Data set on Heart attack possibility”. The performance of the method used was evaluated with accuracy, sensitivity, selectivity, positive predictive value, negative predictive value. The factors associated with having a heart attack were determined by deep learning methods and the most important factors were identified.

Results: Accuracy, sensitivity, specificity, positive predictive value and negative predictive value obtained from the model were 0.814, 0.804, 0.823, 0.809 and 0.834 respectively. The most important 3 factors that may be associated with having a heart attack were obtained as thal, age, ca.

Conclusion: The findings obtained from this study showed that successful predictions were obtained in the classification of having a heart attack by the deep learning method used. In addition, the importance values of the factors associated with the model used were estimated.

Keywords— Heart attack, machine learning, deep learning, classification, variable importance

1. INTRODUCTION

CARDIOVASCULAR diseases are the most common cause of mortality and morbidity in western countries today. Studies show that the mortality rate seen as a result of cardiovascular diseases worldwide will increase from 28.9% to 36.3% between 1990 and 2020 [1]. The most common cardiovascular disease is myocardial infarction (MI), which is usually called a heart attack [2]. MI is the condition of damage to the heart muscle cells that do not receive enough oxygen due to the deterioration of the blood supply of a part of the heart. In addition, if the heart muscle is without oxygen for a long time, death may occur. 50% of deaths from MI occur within the first hour, and this rate reaches 80% within the first 24 hours [3]. MI is an important public health problem that is frequently seen in the productive age group of the population, causes serious problems due to post-acute period complications and can result in death. Despite the advances in diagnosis and treatment in recent years, it is one of the most important causes of morbidity and mortality in our country and industrialized societies [4].

Deep learning is a field of study that has emerged with reference to the way the human brain works and deals with computer systems evaluating a problem based on existing data and producing outputs to solve the problem. Especially in recent years, its popularity has increased with the development of technology and the breakthroughs made in the field of software. Deep learning is considered within the fields of machine learning and artificial intelligence. Machine learning is concerned with computer systems learning certain operations that humans can do and automatically performing these operations.

Artificial intelligence, which includes machine learning, is concerned with the realization of intelligence-based personal evaluations and thoughts by computer systems. Therefore, deep learning artificial intelligence is a wide field of the study evaluated within the scope of machine learning fields. The deep learning field, which has been developing gradually over the years, has achieved successful and important results in problems that cannot be solved with machine learning [5, 6].

In this study, it is aimed to classify the situation of having a heart attack and to determine the related factors by applying the deep learning method on the open-access heart attack data set.

2. MATERIAL AND METHODS

2.1. Dataset

In order to examine the working principle of the deep learning method and to evaluate the model, an open-access data set named “Health care: Data set on Heart attack possibility” is available at https://www.kaggle.com/nareshbhat/health-care-data-set-on- Obtained from heart-attack-possibility. There are 303 patients in the data set used. While 138 (45.5%) of these patients had a low risk of heart attack, 165 (54.5%) of them had a high risk of heart attack.

Explanations about the variables in the data set and their properties are given in Table 1.

Zeynep KUCUKAKCALI, Inonu University Department of Biostatistics and Medical Informatics, Faculty of Medicine, Malatya, Turkey, (zeynep.tunc@inonu.edu.tr)

İpek BALIKCI CICEK, Inonu University Department of Biostatistics and Medical Informatics, Faculty of Medicine, Malatya, Turkey, (ipek.balikci@inonu.edu.tr)

🖂 **Emek GULDOGAN,** Inonu University Department of Biostatistics and Medical Informatics, Faculty of Medicine, Malatya, Turkey, (emek.guldogan@inonu.edu.tr)

Manuscript received Oct 14, 2020; accepted Nov 26, 2020.
Digital Object Identifier:

TABLE I
EXPLANATIONS ABOUT THE VARIABLES IN THE DATASET AND THEIR PROPERTIES

Variable	Variable Description	Variable Type	Variable Role
age	age	Quantitative	Predictor
sex	sex	Qualitative	Predictor
cp	chest pain type (4 values)	Qualitative	Predictor
trestbps	resting blood pressure	Quantitative	Predictor
chol	serum cholestoral in mg/dl	Quantitative	Predictor
fbs	fasting blood sugar > 120 mg/dl	Qualitative	Predictor
restecg	resting electrocardiographic results (values 0,1,2)	Qualitative	Predictor
thalach	maximum heart rate achieved	Quantitative	Predictor
exang	exercise induced angina	Qualitative	Predictor
oldpeak	oldpeak = ST depression induced by exercise relative to rest	Quantitative	Predictor
slope	the slope of the peak exercise ST segment	Qualitative	Predictor
ca	number of major vessels (0-3) colored by flourosopy	Qualitative	Predictor
thal	thal: 0 = normal; 1 = fixed defect; 2 = reversable defect	Qualitative	Predictor
target	target: 0= less chance of heart attack 1= more chance of heart attack	Qualitative	Output

3. DEEP LEARNING

Deep learning is the methods that learn the representations at different levels on data with complex relationships and are included in the machine learning sub-field. Since 2006, researches in the field of deep structured learning or deep learning in its short form have opened a new page in the field of machine learning [6]. When the definitions in the literature are examined, deep learning, in its most general form, is a machine learning technique that is used to solve problems and perform actions such as analysis, inference, observation, and learning using large amounts of data. Unlike traditional machine learning algorithms, they can be in different hierarchical structures [7].

Deep learning is a sub-field of artificial neural networks (ANN). It is the type of ANN that works to get a certain output value from pure data with nonlinear transformations. Deep learning explores the complex structure of multidimensional data sets using the backpropagation algorithm. It does this by comparing the values of the parameters calculated in each layer with the values obtained in the previous layer and deciding what change should be made. Deep learning is based on multi-layered neural networks. Multi-layer neural networks consist of the input layer where the inputs are represented, hidden layers where the information from the input layer is processed into an output, and the output layer where the results from the last hidden layer are converted into output values. Thanks to deep learning methods, the success rate has increased significantly in the fields of natural language processing, image processing, visual object detection, and drug discovery [8].

3.1. Performance evaluation criteria

The classification matrix for the calculation of performance metrics is given in Table 2.

TABLE II
THE METRICS OF THE MODEL'S CLASSIFICATION PERFORMANCE

		Real		
		Positive	**Negative**	**Total**
Predicted	Positive	True positive (TP)	(FN) False negative	TP+FN
	Negative	False positive (FP)	(TN) True negative	FP+TN
	Total	TP+FP	FN+TN	TP+TN+FP+FN

Accuracy = (TP+TN)/(TP+TN+FP+FN)

Sensitivity = TP/(TP+FP)

Specificity = TN/(TN+FN)

Positive predictive value = TP/(TP+FN)

Negative predictive value =TN/(TN+FP)

4. DATA ANALYSIS

Quantitative data are summarized by median (minimum-maximum) and qualitative variables are given by number and percentage. Normal distribution was evaluated with the Kolmogorov-Smirnov test. In terms of input variables, the existence of a statistically significant difference and relationship between the categories of output variable, " less chance of heart attack " and " more chance of heart attack " groups, was examined using Mann-Whitney U, Pearson Chi-square test and Yates's correction chi-square test. $p<0.05$ values were considered statistically significant. In all analyzes, IBM SPSS Statistics 26.0 for the Windows package program was used.

For the validity of the model, a 10-fold cross-validation method was used. In the 10-fold cross-validation method, all data is divided into 10 equal parts. One part is used as a test set and the remaining 9 parts are used as a training data set and this process is repeated 10 times. Hyperparameters related to the deep learning model were selected as activation function (Rectifier), hidden layer sizes (50), the number of revolutions (10), epsilon (1.0 e8) and rho (0.99). RapidMiner Studio software was used in all modeling and analysis [9].

5. RESULTS

Descriptive statistics related to the target variable examined in this study are presented in Table 3 and Table 4. There is a statistically significant difference between the dependent variable classes in terms of other variables other than the "fbs" variable.

Table III
Descriptive statistics for Quantitative Input variables

Variables	Predicted Class		P* value
	more chance of heart attack **Median (min-max)**	less chance of heart attack **Median (min-max)**	
age	52(29-76)	58(35-77)	<0,001
trestbps	130(94-180)	130(100-200)	0,035
chol	234(126-564)	249(131-409)	<0,036
thalach	161(96-202)	142(71-195)	<0,001
oldpeak	0,2(0-4,2)	1,4(0-6,2)	<0,001

*: Mann Whitney U test

Table IV
Descriptive statistics for Qualitative Input variables

Variables		Predicted Class		**P value**
		more chance of heart attack	less chance of heart attack	
sex	0	72(43,6%)	24(17,4%)	<0,001**
	1	93(56,4%)	114(82,6%)	
cp	0	39(23,6%)	104(75,4%)	<0,001*
	1	41(24,8%)	9(6,5%)	
	2	69(41,8%)	18(13,0%)	
	3	16(9,7%)	7(5,1%)	
fbs	0	142(86,1%)	116(84,1%)	0,744
	1	23(13,9%)	22(15,9%)	
restecg	0	68(41,2%)	79(57,2%)	0,007*
	1	96(58,2%)	56(40,6%)	
	2	1(0,6%)	3(2,2%)	
exang	0	142(86,1%)	62(44,9%)	<0,001**
	1	23(13,9%)	76(55,1%)	
slope	0	9(5,5%)	12(8,7%)	<0,001*
	1	49(29,7%)	91(65,9%)	
	2	107(64,8%)	35(25,4%)	
ca	0	130(78,8%)	45(32,6%)	<0,001*
	1	21(12,7%)	44(31,9%)	
	2	7(4,2%)	31(22,5%)	
	3	3(1,8%)	17(12,3%)	
	4	4(2,4%)	1(0,7%)	
thal	0	1(0,6%)	1(0,7%)	<0,001*
	1	6(3,6%)	12(8,7%)	
	2	130(78,8%)	36(26,1%)	
	3	28(17,0%)	89(64,5%)	

*: Pearson's chi-square test;** Yates's correction chi-square test

The classification matrix of the deep learning model used to classify the situation of having a heart attack in this study is given in Table 5 below.

Table V
Classification matrices of Deep learning models

Prediction	**Reference**		
	more chance of heart attack	less chance of heart attack	**Total**
more chance of heart attack	140	31	171
less chance of heart attack	25	107	132
Total	165	138	303

The values for the metrics of the classification performance of the model are given in Table 6. Accuracy, sensitivity, specificity, positive predictive value and negative predictive value obtained from the model were 0.814, 0.804, 0.823, 0.809 and 0.834 respectively.

Table VI
Values for the metrics of the classification performance of deep learning models

Metric	**Value(%)**
Accuracy	0.814
Sensitivity	0.804
Specificity	0.823
Positive predictive value	0.809
Negative predictive value	0.834

In Figure 1, values related to performance criteria are given graphically.

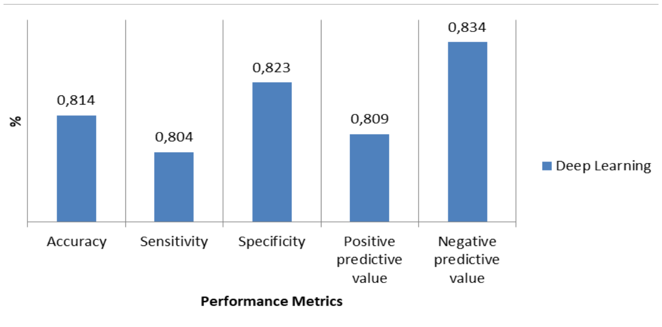


Fig. 1. Values related to performance criteria

The values showing the importance of the variables are shown in Table 6, from high to low. The most important variable is thal (0,086966) followed by age (0,082903) and ca (0,081892) respectively.

TABLE VII
IMPORTANCE VALUES OF EXPLANATORY VARIABLES ACCORDING TO DEEP LEARNING

Variables	Importance
thal	0,086966
age	0,082903
ca	0,081892
oldpeak	0,080578
exang	0,079877
sex	0,079097
cp	0,077245
trestbps	0,076191
slope	0,07378
fbs	0,072506
thalach	0,06997
restecg	0,069702
chol	0,069293

The values for these importance percentages are shown in Figure 2.

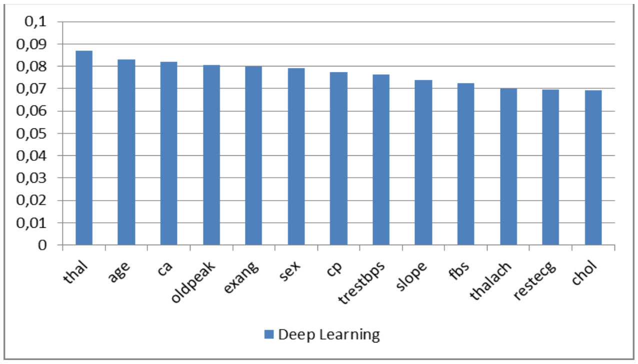


Fig. 2. The importance values for possible risk factors

6. DISCUSSION

Ischemic heart disease (IHD) is one of the diseases that cause mortality and morbidity worldwide [10]. One of these diseases is MI, defined as myocardial cell damage caused by prolonged ischemia. Heart attack is a physiological condition that occurs with severe chest pain and is likely to result in death, as a result of failure after a disorder in the coronary arteries of the heart. Heart attack occurs as a result of oxygen interruption due to a sudden decrease or interruption of blood flow in the vessels that feed the heart. It can cause various degrees of damage or death of the heart muscle fed by the occluded vessel [11]. Heart attack is the most important health problem in developed countries and a serious health problem of increasing importance in developing countries. MI is an important public health problem, which is frequently seen in the productive age group of the population, causes serious problems due to post-acute period complications and can result in death. According to the data of the World Health Organization (WHO), 16.7 million people die each year due to a heart attack. This number is one-third of deaths in the world [12].

Machine learning is a science that deals with the design and development processes of algorithms that allow computers to learn based on data types. Machine learning is not only a database problem but also a field of artificial intelligence that allows experiences gained from existing data to predict and model future events [13]. Deep learning methods are an increasingly popular machine learning method. With the development of graphics processing units (GPU), the Deep learning approach has gained popularity [14]. The deep learning approach is designed inspired by the working principle of the brain. It is a method developed by taking human learning systems as an example and consists of many hidden layers and neurons. With the contribution of hardware features that develop over time (especially the participation of graphics processors in the calculation), the deep learning method is used for processing large-sized data sets. Deep learning works in areas such as motion detection, face recognition, health technologies, object recognition, and object detection [15].

In the study, deep learning method was applied to the data set named “Health care: Data set on Heart attack possibility” which is an open-source data set. Accuracy, sensitivity, specificity, positive predictive value and negative predictive value obtained from the model were 0.814, 0.804, 0.823, 0.809 and 0.834 respectively. Deep learning method gave successful predictive results in the classification of having a heart attack according to the results of the performance criteria calculated in this study. In addition, the importance of risk factors related to having a heart attack was obtained with the experimental findings. The most important 3 factors that may be associated with having a heart attack were obtained as thal, age, ca.

In conclusion, the findings obtained from this study showed that the classification of heart attack status gives successful predictions. In addition, with the classification model used, the importance values of the factors associated with having a heart attack were estimated and the most important factors were determined.

REFERENCES

[1] G. Abanonu, "Koroner arter hastalığı majör risk faktörleri ve C-Reaktif proteinin değerlendirilmesi," Yayınlanmış Uzmanlık Tezi İstanbul, 2005.
[2] Z. Halıcı, H. S. Yasin Bayır, E. Çadırcı, M. S. Keleş, and E. Bayram, "Amiodaron’un Sıçanlarda İsoproterenol ile Oluşturulan Akut ve Kronik Miyokard İnfarktüsü Modelinde Eritropoetin Seviyeleri Üzerine Etkilerinin İncelenmesi."
[3] A. B. Storrow and W. B. Gibler, "Chest pain centers: diagnosis of acute coronary syndromes," Annals of emergency medicine, vol. 35, pp. 449-461, 2000.
[4] S. Şentürk, "İsoproterenol ile miyokart infarktüsü oluşturulmuş ratlarda l-lizin'in total sialik asit düzeylerine etkisinin incelenmesi," 2008.
[5] Y. Bengio, Learning deep architectures for AI: Now Publishers Inc, 2009.
[6] L. Deng and D. Yu, "Deep learning: methods and applications," Foundations and trends in signal processing, vol. 7, pp. 197-387, 2014.
[7] K. Kayaalp and A. Süzen, "Derin Öğrenme ve Türkiye’deki Uygulamaları," Yayın Yeri: IKSAD International Publishing House, Basım sayısı, vol. 1, 2018.
[8] Y. LeCun, Y. Bengio, and G. Hinton, "Deep learning," nature, vol. 521, pp. 436-444, 2015.
[9] I. Mierswa and R. Klinkenberg, "Rapidminer studio," ed, 2019.

[10] B. Thippeswamy, S. Thakker, S. Tubachi, G. Kalyani, M. Netra, U. Patil, et al., "Cardioprotective effect of Cucumis trigonus Roxb on isoproterenol-induced myocardial infarction in rat," American journal of pharmacology and toxicology, vol. 4, pp. 29-37, 2009.
[11] S. Ateş, "Coğrafi Bilgi Sistemleri İle Kalp Krizi Vakalarına Yönelik En Uygun Ambulans Yerlerinin Belirlenmesi," Fen Bilimleri Enstitüsü, 2010.
[12] A. Upaganlawar, H. Gandhi, and R. Balaraman, "Isoproterenol induced myocardial infarction: protective role of natural products," J Pharmacol Toxicol, vol. 6, pp. 1-17, 2011.
[13] E. Alpaydin, Introduction to machine learning: MIT press, 2020.
[14] A. Şeker, B. Diri, and H. H. Balık, "Derin öğrenme yöntemleri ve uygulamaları hakkında bir inceleme," Gazi Mühendislik Bilimleri Dergisi, vol. 3, pp. 47-64, 2017.
[15] J. Schmidhuber, "Deep learning in neural networks: An overview," Neural networks, vol. 61, pp. 85-117, 2015.

BIOGRAPHIES

Zeynep KUCUKAKCALI obtained her BSc. degree in mathematics from Çukurova University in 2010. She received MSc. degree in biostatistics and medical informatics from the Inonu University in 2018. She currently continues Ph.D. degrees in biostatistics and medical informatics from the Inonu University. In 2014, she joined the Department of Biostatistics and Medical Informatics at Inonu University as a researcher assistant. Her research interests are cognitive systems, data mining, machine learning, deep learning.

İpek BALIKCI CİCEK obtained her BSc. degree in mathematics from Çukurova University in 2010. She received MSc. degree in biostatistics and medical informatics from the Inonu University in 2018. She currently continues Ph.D. degrees in biostatistics and medical informatics from the Inonu University. In 2014, she joined the Department of Biostatistics and Medical Informatics at Inonu University as a researcher assistant. Her research interests are cognitive systems, data mining, machine learning, deep learning.

Emek GÜLDOĞAN obtained his BSc. degree in Computer Engineering from Middle East Technical University in 2001. He received MSc. degree in biostatistics and medical informatics from the Inonu University in 2005, and Ph.D. degrees in biostatistics and medical informatics from the Inonu University in 2017. He is currently working as an assistant professor of the Department of Biostatistics and Medical Informatics at Inonu University and as the information processing manager at Turgut Özal Medical Center. His research interests are cognitive systems, data mining, machine learning, deep learning.

THE JOURNAL OF
COGNITIVE SYSTEMS

an international, peer-reviewed
indexed, and open-access periodical

PUBLICATION ETHICS

All who participate in producing The Journal of Cognitive Systems conduct themselves as authors, reviewers, editors, and publishers in accord with the highest level of professional ethics and standards. Plagiarism or self-plagiarism constitutes unethical scientific behaviour, and is never acceptable. By submitting a manuscript to this journal, each author explicitly confirms that the manuscript meets the highest ethical standards for authors and co-authors. The undersigned hereby assign(s) to The Journal of Cognitive Systems (JCS) copyright ownership in the above paper, effective if and when the paper is accepted for publication by JCS, and to the extent transferable under applicable national law. This assignment gives JCS the right to register copyright to the paper in its name as claimant, and to publish the paper via any print or electronic medium.

Authors, or their employers, in the case of works made for hire, retain the following rights.

+ all proprietary rights other than copyright, including patent rights
+ the right to make and distribute copies of the Paper for internal purposes
+ the right to use the material for lecture or classroom purposes
+ the right to prepare derivative publications based on the Paper, including books or book chapters, journal papers, and magazine articles, provided that publication of a derivative work occurs subsequent to the official date of publication by JCS.
+ the right to post an author-prepared version or an official version (preferred version) of the published paper on an internal or external server controlled exclusively by the author/employer, provided that (a) such posting is non-commercial in nature, and the paper is made available to users without charge; (b) a copyright notice and full citation appear with the paper, and (c) a link to JCS's official online version of the abstract is provided using the Document Object Identifier (DOI) link

THE JOURNAL OF COGNITIVE SYSTEMS

an international, peer-reviewed, indexed, and
open-access periodical

VOLUME 05, NUMBER 02

DECEMBER 2020

CONTENTS

www.dergipark.gov.tr/jcs

ISSN 2548-0650