PERFORMANCE EVALUATION OF DIFFERENT ARTIFICIAL NEURAL NETWORK MODELS IN THE CLASSIFICATION OF TYPE 2 DIABETES MELLITUS

E. Guldogan, Z. Tunc, A. Acet and C. Colak

Abstract**—** Objective: In this study, it is aimed to classify type 2 Diabetes Mellitus (DM), compare the estimates of the Artificial Neural Network models and determine the factors related to the disease by applying Multilayer Perceptron (MLP) and Radial Based Function (RBF) methods on the open-access dataset.

Material and Methods: In this study, the data set named "Pima Indians Diabetes Database" was obtained from https://www.kaggle.com/uciml/pima-indians-diabetes-database. The dataset contains 768 records with 268 (34.9%) type 2 diabetes patients and 500 (65.1%) people without diabetes, which have 9 variables (8 inputs and 1 outcome). MLP and RBF methods, which are artificial neural network models, were used to classify type 2 DM. Factors associated with type 2 DM were estimated by using artificial neural network models.

Results: The performance values obtained with MLP from the applied models were accuracy 78.1%, specificity 81.2%, AUC 0.848, sensitivity 71%, positive predictive value 61.7%, negative predictive value 86.8% and F-score 66%. In relation to RBF model, the performance metrics were accuracy obtained 76.8%, specificity 82.1%, AUC 0.813, sensitivity 66.0%, positive predictive value 64.6%, negative predictive value 83% and F-score 65.3%, respectively. When the effects of the variables in the data set examined in this study on Type 2 DM are analyzed; The three most important variables for the MLP model were obtained as Glucose, BMI, Pregnancies respectively. For RBF, it was obtained as Glucose, Skin Thickness, and Insulin.

Conclusion: The findings obtained from this study showed that the models used gave successful predictions for Type 2 DM classification. Besides, unlike similar studies examining the same dataset, the significance values of the factors associated with the models created were estimated.

Keywords**—***Classification, Multilayer perceptron neural network, Radial-based function neural network, Type 2 Diabetes Mellitus.*

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

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

Ayça ACET, Inonu University, Department of Computer Engineering, Faculty of Engineering, , Malatya, Turkey, (a.ayceacet@gmail.com)

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

Manuscript received May 08, 2020; accepted May 29, 2020.
Digital Object Identifier:

1. INTRODUCTION

DIABETS mellitus (DM) is a chronic disease that seriously affects both daily life and quality of life. This disease cannot be cured completely, but when it is managed well and precautions are taken, its negative effects in the short and long term can be minimized [1, 2]. DM is a chronic and metabolic disease characterized by abnormalities in protein, carbohydrate, and fat metabolism caused by absolute or relative insulin deficiency and accompanying clinical, biochemical findings [3]. Type 1 DM is an autoimmune disease and is caused by the destruction of pancreatic beta cells. Type 2 DM is defined as the combination of insulin resistance and impairment of pancreatic beta cells in insulin secretion [4].

Type 2 DM is a heterogeneous disorder caused by a large number of genetic and environmental factors. Although the pathogenesis of type 2 diabetes is quite complex, it is characterized by two main pathophysiological causes.
A decrease in insulin sensitivity or insulin resistance.

Dysfunction of pancreatic beta cells in addition to relative insulin deficiency (insulin secretion defect) [5].

Type 2 DM accounts for 80-90% of all diabetes cases. The frequency of the disease is increasing gradually all over the world. The prevalence of type 2 DM increases with age. Factors such as the transition from traditionally accepted lifestyle to western lifestyle, the increase in the number of overweight and obese individuals, decrease in activities such as exercise and sports, and unhealthy diet contributed to the prevalence of the disease [6].

Artificial Neural Networks (ANNs) are computer systems developed to directly realize the features of learning, which is one of the features of the human brain, such as the ability to derive, create and discover new information without any help [7]. ANN can provide nonlinear modeling without needing any prior knowledge between input and output variables, without any assumptions [8]. Artificial neural networks are a successful method in solving many daily life problems such as classification, modeling, and prediction. Multilayer Perceptron (MLP) is a frequently used ANN model for the solution of nonlinear problems. It is a feed-forward, backpropagation network using at least one layer between the input and output layers consisting of at least three layers [9]. In the forward propagation stage, while calculating the output and error value of the network, the link weight values between the layers are updated to minimize the calculated error value in the reverse propagation stage [10].

Radial-based function (RBF) neural networks are feed-forward networks consisting of a 3-layer structure: an input layer, an output layer, and a single hidden layer. This hidden layer is the layer using radial functions that give the network its name as a transfer function. While the inputs of this network are not linear, the output is linear [11]. The input layer consists of source nodes and provides the connection of the network with the environment. The only hidden layer in the network makes a nonlinear transformation from the input area to the hidden area. The conversion from the input layer to the hidden layer is a nonlinear constant transformation with radial-based transfer functions. The output layer is linear and is the layer that responds to the network, which is the transfer signal applied to the input layer. An adaptive and linear transformation is performed from the hidden layer to the output layer [12].

In this study, it is aimed to compare the classification performance of Type 2 DM and to determine the risk factors that may be associated with Type 2 DM by applying MLP and RBF models on the open-access Type 2 DM data set.

2. Material and methods

2.1. Dataset

In this study, the data set named "Pima Indians Diabetes Database" was obtained from *https://www.kaggle.com/uciml/pima-indians-diabetes*-database to examine the working principle of MLP and RBF ANN models and to determine risk factors. In the data set used, there were 768 individuals with 268 (34.9%) type 2 diabetes patients and 500 (65.1%) people without diabetes. DM was described as a concentration of plasma glucose greater than 200 mg/dl two hours after ingestion of a carbohydrate solution with 75 gm. All the subjects were females and ≥ 21 years old at the time of the index examination [13, 14]. Explanations about the variables in the data set and their properties are given in Table 1.

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

Variable	Variable Description	Variable Type	Variable Role
Pregnancies	Number of pregnancies	Quantitative	Input
Glucose	2-hour plasma glucose concentration in the oral glucose tolerance test	Quantitative	Input
Blood Pressure (BP)	Diastolic blood pressure (mmHg)	Quantitative	Input
Skin Thickness (ST)	Triceps skinfold thickness (mm)	Quantitative	Input
Insulin	2-hour serum insulin (mu U / ml)	Quantitative	Input
BMI	Body mass index [weight in kg / (height in m)^2]	Quantitative	Input
Diabetes Pedigree Function (DPF)	Diabetes family tree function	Quantitative	Input
Age	Age (years)	Quantitative	Input
Outcome	Class variable (type 2 DM; 0: absent, 1: present)	Qualitative	Output

3. Artificial Neural Network Models

In this study, classification performance was compared by applying MLP and RBF methods on artificial neural network models on Type 2 DM data set and risk factors that may be associated with Type 2 DM were determined. Because of its power, flexibility, and ease of use, artificial neural networks are the preferred tool for many predictive data mining applications. Predictive neural networks are particularly useful in applications where the mechanism underlying them is complex. In recent years, interest in the application of neural networks has increased for problems that cannot be solved with classical techniques, and ANN has been used successfully in many medical applications. Unlike traditional spectral analysis methods, artificial neural networks not only model signals but also produce solutions for the classification of signals. Another advantage of artificial neural networks compared to the methods available for the analysis of biomedical signals is that after their training, they are very fast. MLP is a nonparametric artificial neural network technique that performs many detection and prediction operations [15].

Radial Based Function Neural Network (RBFNN) was developed in 1988 inspired by the effect response behaviors seen in biological nerve cells and entered the history of ANN by applying it to the filtering problem. It is possible to see the training of RBFNN models as a curve-fitting approach in multi-dimensional space. For this reason, the training performance of the RBFNN model turns into an interpolation problem, finding the most suitable surface for the data in the output vector space. RBFNN models are defined in three layers as the input layer, hidden layer, and output layer, similar to general ANN architecture. However, unlike conventional ANN structures, RBFNNs use radial-based activation functions and nonlinear cluster analysis in the transition from the input layer to the hidden layer. The structure between the hidden layer and the output layer continues to function as in other ANN types [16].

In the construction of MLP and RBF models, nearly 60% and 40% of the whole dataset were used for training and testing stages, respectively. Rescaling method for the variables was standardized for both models, the number of units in the hidden layer was 6 for MLP and 5 for RBF, hidden layer activation function was hyperbolic tangent for MLP and softmax for RBF, the number of units in output layer was 2 for both models, and output layer activation function was softmax for MLP and identity for RBF. Hyperparameters of the models were optimized by the scaled conjugate gradient method.

3.1. Performance Evaluation of the Models

In the performance evaluation of MLP and RBF artificial neural network models for predicting the factors that may be associated with type 2 DM, different metrics that can be calculated from the values in the confusion matrix (Table 2) given below have been obtained.

TABLE II
CONFUSION 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

The metrics considered in the performance evaluation of the models in this study are given below.

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)

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

4. DATA ANALYSİS

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. Whether there is a statistically significant difference between categories of the dependent variable in terms of input variables, the Mann-Whitney U test was used for the analyses. P<0.05 values were considered statistically significant. In all analyses, 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. There is a statistically significant difference between the dependent variable classes in terms of other variables other than the insulin variable.

In this study, descriptive statistics of the factors examined according to the type 2 DM variable are summarized in Table 3. According to these findings, while there was a difference between the presence and absence of type 2 DM in terms of Pregnancies, Glucose, BP, ST, BMI, DPF and Age (p<0.05), no statistically significant difference was found for the Insulin factor (p>0.05).

TABLE III
DESCRIPTIVE STATISTICS ABOUT INPUT AND OUTPUT VARIABLES

	Outcome (Diabetes Mellitus)		
Variable	**Absent (n=500)**	**Present (n=268)**	**p value**
Statistics	*Median (Min-Max)*	*Median (Min-Max)*	
Pregnancies	2 (0-13)	4 (0-17)	<0.001
Glucose	107 (0-197)	140 (0-199)	<0.001
BP	70 (0-122)	74 (0-114)	<0.001
ST	21 (0-60)	27 (0-99)	0.013
Insulin	39 (0-744)	0 (0-846)	0.066
BMI	30.1 (0-57,3)	34.3 (0-67.1)	<0.001
DPF	0.336 (0.078-2.329)	0.449 (0.088-2.42)	<0.001
Age	27 (21-81)	36 (21-70)	<0.001

BMI: Body mass index; ST: Skin Thickness; DPF: Diabetes Pedigree Function; BP: Blood pressure;

Classification matrices of the testing stages for MLP and RBF models are shown in Tables 4 and 5, respectively.

TABLE IV
CLASSIFICATION MATRIX OF THE TESTING STAGE FOR THE MLP MODEL

Real / Predicted	**Present**	**Absent**	**Total**
Present	66	41	107
Absent	27	177	204
Total	93	218	311

TABLE V
CLASSIFICATION MATRIX OF THE TESTING STAGE FOR THE RBF MODEL

Real / Predicted	**Present**	**Absent**	**Total**
Present	64	35	99
Absent	33	161	194
Total	97	196	293

Table 6 presents the performance criteria values calculated from the models to classify type 2 DM in the testing stage.

 ISSN: 2548-0650 http://dergipark.gov.tr/jcs

TABLE VI
PERFORMANCE METRIC VALUES CALCULATED FROM THE GENERATED MODELS IN THE TESTING STAGE

Metric \ Model	ANN type	
	MLP	**RBF**
Accuracy (%)	78.1	76.8
Specificity (%)	81.2	82.1
AUC	0.848	0.813
Sensitivity (%)	71.0	66.0
Positive predictive value (%)	61.7	64.6
Negative predictive value (%)	86.8	83
F-score (%)	66.0	65.3

AUC: Area under the ROC curve; MLP: Multilayer perceptron neural network; RBF: Radial-based function neural network

The values related to performance criteria obtained from MLP and RBF models are demonstrated in Figure 1.

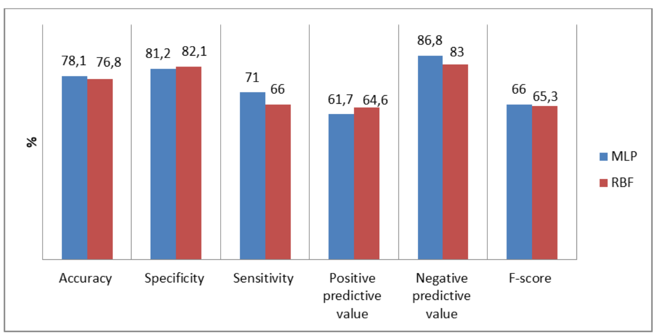

Fig.1. Performance criteria values obtained from MLP and RBF models in the testing stage (MLP: Multilayer perceptron; RBF: Radial-based function)

TABLE VII
IMPORTANCE VALUES OF EXPLANATORY VARIABLES ACCORDING TO MLP AND RBF MODELS

Explanatory Variables	MLP	RBF
Glucose	0.287	0.175
BMI	0.219	0.144
Pregnancies	0.134	0.074
DPF	0.125	0.135
Age	0.077	0.068
Insulin	0.072	0.159
BP	0.057	0.078
ST	0.03	0.167
Total	100.00	100.00

BMI: Body mass index; ST: Skin Thickness; DPF: Diabetes Pedigree Function;
BP: Blood pressure; MLP: Multilayer perceptron; RBF: Radial-based function

In this study, the importance values of the factors related to diabetes mellitus are given in Table 7, while the values for these importance percentages are shown in Figure 2.
Table 7: Importance values of explanatory variables according to MLP and RBF models

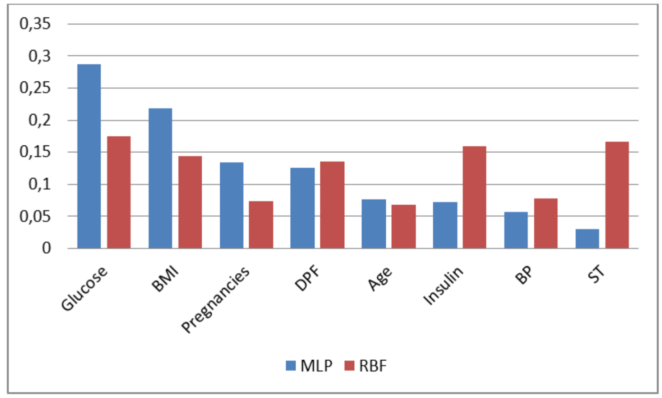

Fig.2. The importance values for possible risk factors (BMI: Body mass index; ST: Skin Thickness; DPF: Diabetes Pedigree Function; BP: Blood pressure; MLP: Multilayer perceptron; RBF: Radial-based function)

6. DISCUSSION

An artificial neural network is a very successful technique that solves the classification and prediction problems and is a mathematical model promoted by the regulation and functional feature of artificial neural networks. Neural networks include input and output layers and (in most cases) hidden layers that convert the input to output. When artificial neural network architecture is used to predict any disease, the ANN model can be generally built in two stages: training and testing. First, the ANN model is trained on the specified dataset and the weights of the connections between the neurons are fixed. Second, the model examined is validated to determine the classification of a new data set. The performance of the models constructed is evaluated using different criteria. [17].

In this study, it was aimed to apply multilayer perceptron and radial-based function from artificial neural network models on an open-source type 2 DM dataset and to compare classification estimates. In this framework, various factors (explanatory variables) that may be associated with type 2 DM (dependent variable) are estimated by multilayer perceptron and radial-based function artificial neural network models, and the use of artificial intelligence methods in the classification problem of interest is revealed. Also, the importance levels of factors that may be associated with type 2 DM for use in preventive medicine applications were obtained from these models.

According to the results of the performance criteria (accuracy, AUC, sensitivity, negative predictive value, and F-score) calculated in this study, the MLP model gave better predictive results than the RBF model in the classification of type 2 DM. However, when the criteria of selectivity and positive predictive value are taken into consideration compared to the MLP model, the higher classification rates were obtained. The three most important risk factors that can be associated with type 2 DM were Glucose, BMI, and Pregnancies according to the MLP model, while the RBF model defined Glucose, Skin Thickness, and Insulin, respectively.

In a study using the same data set, the accuracy, sensitivity, and specificity performance criteria used in the classification made with support vector machines were obtained as 78%, 80%, and 76.5, respectively [18]. In another study using the same data set, classification was made using six different machine learning models. The best classification performances were achieved from the Hoeffding Tree algorithm based on these models, and precision, Recall, F-criterion, and area under the ROC curve values calculated from this algorithm were obtained as 0.757, 0.762, 0.759 and 0.816, respectively. [19]. In this study, the classification of DM was performed with MLP and RBF models on the same data set, and higher classification performances were obtained from the experimental results of the current study.

As a result, the findings obtained from this study showed that the classification of Type 2 DM performed successful predictions. Also, unlike similar studies examining the same dataset, the importance values of the factors associated with the type 2 DM were estimated with the classification models.

REFERENCES

[1] J. A. Cramer, "A systematic review of adherence with medications for diabetes," Diabetes care, vol. 27, no. 5, pp. 1218-1224, 2004.

[2] R. Shobhana, R. Begum, C. Snehalatha, V. Vijay, and A. Ramachandran, "Patients' adherence to diabetes treatment," The Journal of the Association of Physicians of India, vol. 47, no. 12, pp. 1173-1175, 1999.

[3] N. Başkal, "Diabetes Mellitus Tanım, Klasifikasyon, Tanı, Klinik, Laboratuar ve Patogenez," Erdoğan G. Klinik Endokrinoloji. Anıtıp AŞ yayınları, Ankara, pp. 207-233, 2003.

[4] A. Cameron, "The metabolic syndrome: validity and utility of clinical definitions for cardiovascular disease and diabetes risk prediction," Maturitas, vol. 65, no. 2, pp. 117-121, 2010.

[5] M. Arslan, "Diabetes mellitusta tanı ve sınıflandırma," İliçin G, Biberoğlu K, Süleymanlar G, Ünal S (editörler). İç Hastalıkları, vol. 2, pp. 2279-2295, 2003.

[6] T. Yılmaz, "Diabetes mellitusun tanı kriterleri ve sınıflaması," Diabetes Mellitus' un Modern Tedavisi, birinci baskı, İstanbul, Türkiye Diyabet Vakfı, 2003.

[7] E. Öztemel, "Yapay Sinir Ağları, Papatya Yayıncılık, 2," Baskı, İstanbul, pp. 29-57, 2006.

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

[9] H. Batar, "EEG işaretlerinin dalgacık analiz yöntemleri kullanılarak yapay sinir ağları ile sınıflandırılması," Fen Bilimleri Enstitüsü, Yüksek Lisans Tezi, Kahramanmaraş, 89s, 2005.

[10] A. Arı and M. E. Berberler, "Yapay Sinir Ağları ile Tahmin ve Sınıflandırma Problemlerinin Çözümü İçin Arayüz Tasarımı," Acta Infologica, vol. 1, no. 2, pp. 55-73, 2017.

[11] E. Kilic, U. Ozbalci, and H. Ozcalik, "Lineer Olmayan Dinamik Sistemlerin Yapay Sinir Ağları ile Modellenmesinde MLP ve RBF Yapılarının Karşılaştırılması," ELECO2012 Elektrik-Elektronik ve Bilgisayar Mühendisliği Sempozyomu,(29.11. 2012-01.12. 2012), 2012.

[12] S. S. Haykin, "Neural networks and learning machines/Simon Haykin," ed: New York: Prentice Hall, 2009.

[13] M. Barale and D. J. I. J. o. C. A. Shirke, "Cascaded modeling for PIMA Indian diabetes data," vol. 139, no. 11, pp. 1-4, 2016.

[14] J. W. Smith, J. Everhart, W. Dickson, W. Knowler, and R. Johannes, "Using the ADAP learning algorithm to forecast the onset of diabetes mellitus," in Proceedings of the Annual Symposium on Computer Application in Medical Care, 1988, p. 261: American Medical Informatics Association.

[15] U. Orhan, M. Hekim, and M. Özer, "Discretization approach to EEG signal classification using Multilayer Perceptron Neural Network model," in 2010 15th National Biomedical Engineering Meeting, 2010, pp. 1-4: IEEE.

[16] O. Kaynar, Y. Görmez, Y. E. Işık, and F. Demirkoparan, "Değişik Kümeleme Algoritmalarıyla Eğitilmiş Radyal Tabanlı Yapay Sinir Ağlarıyla Saldırı Tespiti," in International Artificial Intelligence and Data Processing Symposium (IDAP'16), 2016.

[17] I. M. Nasser and S. S. Abu-Naser, "Lung Cancer Detection Using Artificial Neural Network," International Journal of Engineering and Information Systems (IJEAIS), vol. 3, no. 3, pp. 17-23, 2019.

[18] V. A. Kumari and R. Chitra, "Classification of diabetes disease using support vector machine," International Journal of Engineering Research and Applications, vol. 3, no. 2, pp. 1797-1801, 2013.

[19] F. Mercaldo, V. Nardone, and A. Santone, "Diabetes mellitus affected patients classification and diagnosis through machine learning techniques," Procedia computer science, vol. 112, pp. 2519-2528, 2017.

BIOGRAPHIES

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 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.

Zeynep Tunç 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.

Ayca Acet received the BSc degree in Computer Engineering from TOBB University of Economics and Technology, Ankara, Turkey, in 2017. She currently continues M.S. degree in Computer Engineering, at Inonu University, Malatya, Turkey. Her general research interests include machine learning methods, deep learning and data mining.

Cemil Çolak obtained his BSc. degree in statisticsfrom Ondokuz Mayıs University in 1999. He received MSc. degree in statistics from the InonuUniversity in 2001, and Ph.D. 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, analysis.