THE JOURNAL OF COGNITIVE SYSTEMS

Journal Homepage: https://dergipark.org.tr/en/pub/jcs

Vol.6, No.1, 2021

Performance Evaluation of Multilayer Perceptron Artificial Neural Network Model in the Classification of Heart Failure

Mehmet Onur Kaya

Firat University, Faculty of Medicine, Department of Biostatistics and Medical Informatics, Elazig, Tturkey. (e-mail: mokaya@firat.edu.tr).

ARTICLE INFO

Received: Apr.,19.2021
Revised: May.,02.2021
Accepted: May.,22.2021

Keywords:
Classification
Multilayer perceptron
Artificial neural network
Heart failure

Corresponding author: Mehmet Onur Kaya
✉ mokaya@firat.edu.tr
☎ +90 424 2370000 / 4672

ISSN: 2548-0650

DOI: https://doi.org/10.52876/jcs.913671

ABSTRACT

Objective: The aim of this study was to compare the classification performance of heart failure using the MLP ANN model on an open-access "heart failure clinical records" data set, as well as to identify risk factors that may be linked to heart failure.

Material and Methods: The open-access "heart failure" data collection MLP ANN model was used to classify nephritis of the renal pelvis, and risk factors that may be involved were discovered. Different output metrics are used to demonstrate MLP ANN's progress.

Results: It has been shown that the classification of renal pelvic nephritis is quite high with MLP ANN model (AUC = 0.925, Accuracy = 93.9%, Balanced Accuracy = 89.2%, Sensitivity = 98.4%, Specificity = 80.0%). Furthermore, the MLP ANN model showed that "time" is the most significant variable among the risk factors linked to heart failure.

Conclusion: Consequently, in the analysis with the heart failure data collection, the MLP ANN model generated very positive results. Moreover, this model has gained important information in identifying risk factors that may be associated with heart failure. Thus, it has been understood that the relevant model will provide reliable information about any disease to be used in preventive medicine practices.

1. INTRODUCTION

HEART failure is a common clinical syndrome characterized by dyspnea, fatigue, and signs of volume overload, which may include peripheral edema and pulmonary rales resulting from impaired systolic and/or diastolic function of the heart. There is no single diagnostic test for heart failure. Therefore, the diagnosis is made with the patient's anamnesis, physical examination, and laboratory test. Symptoms of heart failure can vary depending on systolic or diastolic dysfunction. The prognosis of heart failure is closely related to the stage of the disease and the duration of the disease. Survival is 89.6 percent at one month after diagnosis, 78 percent at one year, and only 57.7 percent at five years. Given the poor prognosis, proper diagnosis and treatment of heart failure are important [1].

Although many diagnostic criteria have been specified in the diagnosis of heart failure, Framingham criteria are widely used in diagnosis. Framingham primary criteria are defined as Acute pulmonary edema, Cardiomegaly, Hepatojugular reflex, Neck vein distension, Paroxysmal nocturnal dyspnea or orthopnea, Rales, Third heart sound as gallop, and minor criteria are defined as Ankle edema, Dyspnea on exertion, Hepatomegaly, Nocturnal cough, Pleural effusion, Tachycardia (> 120 beats per minute). Diagnosis is made by considering these criteria [2]. Heart failure is diagnosed when there are two major criteria or one major and two minor criteria. Although the Framingham criteria are widely used in diagnosis, two different studies reported that heart failure occurs when Framingham criteria are not met [3,4]. It is known that early diagnosis and treatment in heart failure are also essential to take into account the poor prognosis. In this context, new developments are needed in diagnosis to start early diagnosis and treatment.

Artificial Neural Networks (ANNs) are computer systems developed to automatically realize skills such as generating new information, creating and discovering new information through learning, which are characteristics of the human brain, without any help [5]. ANNs can model nonlinearly without the need for any assumptions or prior knowledge of input and output variables. In addition, the Multilayer Perceptron (MLP) is one of the most commonly used ANN models for the solution of nonlinear problems. There is at least one layer between the input and output layers in this feed-forward backpropagation network. The relation weight values between the layers are updated to decrease the calculated error value in the reverse propagation stage when the network's output and error value are measured in the forward propagation stage [6,7].

The aim of this study was to compare the classification performance of heart failure using the MLP ANN model on an open-access "heart failure" data set, as well as to identify risk factors that may be linked to heart failure.

2. MATERIAL AND METHODS

2.1. Dataset

In this study, the open-access data set named "Heart failure clinical records" was used to evaluate the performance of the MLP ANN model and to determine the possible risk factors [8] with heart failure.

The number of cases that make up the study sample is 299 patients. The patients ranged in age from 40 to 95 years old, with 105 women and 194 men among them. After a certain follow-up period, 96 (32.1%) of these patients were reported to die. Table 1 shows the variables in the data set and the properties of the variables.

TABLE 1
PROPERTIES OF THE VARIABLES IN THE STUDY

Variable	Variable Description	Variable Type	Variable Role
Age	Decrease of red blood cells or hemoglobin	Quantitative	Input
Anemia	Level of the CPK enzyme in the blood (mcg/L) (boolean)	Qualitative	Input
Creatinine phosphokinase	Level of the CPK enzyme in the blood (mcg/L)	Quantitative	Input
Diabetes	If the patient has diabetes (boolean)	Qualitative	Input
Ejection fraction	Percentage of blood leaving the heart at each contraction (percentage)	Quantitative	Input
High blood pressure	If the patient has hypertension (boolean)	Qualitative	Input
platelets	Platelets in the blood (kiloplatelets/mL)	Quantitative	Input
Serum creatinine	Level of serum creatinine in the blood (mg/dL)	Quantitative	Input
Serum sodium	Level of serum sodium in the blood (mEq/L)	Quantitative	Input
Gender	Woman or man (binary)	Qualitative	Input
Smoking	If the patient smokes or not (boolean)	Qualitative	Input
Time	Follow-up period (days)	Quantitative	Input
Death event	If the patient deceased during the follow-up period (boolean)	Qualitative	Output

2.2. Multilayer Perceptron Artificial Neural Network Model

The output of the MLP ANN method on the heart failure data set was investigated in this analysis, and risk factors that may be linked to heart failure were identified. Predictive ANNs are particularly useful in applications with a complicated mechanism. ANNs are currently gaining popularity as a solution to problems that cannot be solved with traditional methods, and they have been used successfully in a variety of medical applications. ANNs, unlike conventional spectral analysis, approaches model signals as well as generate signal classification solutions. Another advantage of ANNs over other methods for analyzing biomedical signals is that they are speedy after being educated. The MLP ANN model is a nonparametric artificial neural network technique that can perform a wide range of detection and prediction tasks [9].

Approximately 73% and 27% of the entire dataset is used for the training and testing process, respectively, to create the MLP ANN model. The number of units in the input layer was 12, the number of units in the hidden layer was 2, the hidden layer activation function was a hyperbolic tangent, the number of units in the output layer was 2, the output layer activation function was Softmax, and the error function was Cross-entropy. The scaled conjugate gradient approach was used to optimize the model's hyperparameters.

2.3. Performance Evaluation of the Models

The MLP ANN model was evaluated using performance metrics including Accuracy, Balanced accuracy, Sensitivity, Specificity, Positive predictive value (PPV), Negative predictive value (NPV), F-score, and Area under the curve (AUC) with an appropriate 95% confidence interval (CI).

2.4. Data Analysis

IBM SPSS Statistics 26.0 program was used for all analyzes in the study. The compliance of continuous variables to the normal distribution was evaluated using the Shapiro Wilk test. The comparison of continuous variables that do not provide the assumption of normal distribution between two independent groups (Death Event: Survived and Died) was performed with the Mann Whitney U test, and descriptive statistics were presented as median (min-max). In the analysis of categorical variables, the Pearson chi-square test was used, and descriptive statistics were presented as frequency (%). The significance level was accepted as 0.05.

3. RESULTS

The statistical analysis findings of the variables included in the data set are presented in Table 2. In the table, statistically significant p values are presented in bold.

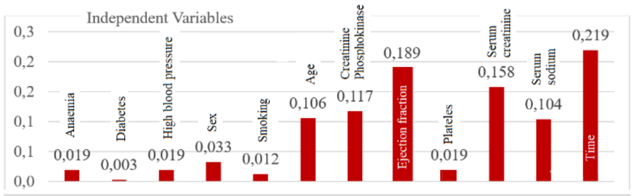

Figure 1. The importance values of possible risk factors associated with heart failure.

When the table is examined, it is seen that Age, Ejection Fraction, Serum Creatinine, Serum Sodium, Time, and Smoking variables are statistically significantly different between the groups. Table 3 shows the classification matrices for the MLP ANN model's training and testing measures.Table 4 shows the performance metrics results obtained from the MLP ANN model for classifying hearth failure in the test step.

TABLE 2
STATISTICAL COMPARISONS OF VARIABLES

Variables		**Death Event**		**p-value**
		Survived (n=203)	**Dead (n=96)**	
Age [Median (Min-Max)]		60(40-90)	65(42-95)	**<0.001**
Creatinine Phosphokinase [Median (Min-Max)]		245(30-5209)	259(237-861)	0.684
Ejection Fraction [Median (Min-Max)]		38(17-80)	30(14-70)	**<0.001**
Platelets [Median (Min-Max)]		263000(25100-850000)	258500(47000-621000)	0.425
Serum Creatinine [Median (Min-Max)]		1(0.5-6.10)	1.3(0.6-9.40)	**<0.001**
Serum Sodium [Median (Min-Max)]		137(113-148)	136(116-146)	**<0.001**
Time [Median (Min-Max)]		172(12-285)	45(4-241)	**<0.001**
Gender n(%)	Female	71(35.00%)	34(35.40%)	0.941
	Male	132(65.00%)	62(64.60%)	
Anemia n(%)	Absence	120(59.10%)	50(52.10%)	0.252
	Presence	83(40.90%)	46(47.90%)	
Diabetes n(%)	Absence	118(58.10%)	56(58.30%)	0.973
	Presence	85(41.90%)	40(41.70%)	
High blood pressure n(%)	Absence	137(67.50%)	57(59.40%)	0.170
	Presence	66(32.50%)	39(40.60%)	
Smoking n(%)	Absence	137(67.50%)	66(68.75%)	0.827
	Presence	66(32.50%)	30(31.25%)	

TABLE 3
CLASSIFICATION MATRICES OF TRAINING AND TESTING STEPS

Training Step

		Reference		**Total**
		Survived	**Dead**	
Predicted	**Survived**	127	11	138
	Dead	23	56	79
	Total	150	67	217 (72.6%)

Testing Step

		Reference		**Total**
		Survived	**Dead**	
Predicted	**Survived**	61	4	65
	Dead	1	16	17
	Total	62	20	82 (27.4%)

TABLE 4
CLASSIFICATION PERFORMANCE METRICS OF THE MLP ANN MODEL

Performance Metrics	**Value (95% CI)**
Accuracy (%)	93.9 (88.7 – 99.1)
Balanced Accuracy (%)	89.2 (82.5 – 95.9)
Sensitivity (%)	98.4 (91.3 – 1.00)
Specificity (%)	80.0 (56.3 – 94.3)
PPV (%)	93.8 (85.0 – 98.3)
NPV (%)	94.1 (71.3 – 99.9)
F1-score (%)	96.1 (91.9 – 1.00)
AUC	0.925

Table 5 and Figure 1 show the significance values of the input variables in the data set obtained from the MLP ANN model, respectively.

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

TABLE 5
THE IMPORTANCE VALUES OF THE INPUT VARIABLES

Independent Variables	Importance
Anemia	0,019
Diabetes	0,003
High blood pressure	0,019
Sex	0,033
Smoking	0,012
Age	0,106
Creatinine phosphokinase	0,117
Ejection fraction	0,189
Platelets	0,019
Serum creatinine	0,158
Serum sodium	0,104
Time	0,219

4. DISCUSSION

Heart failure is one of the most common cardiovascular diseases in the world, the last stage of all heart diseases, a health problem with increasing prevalence and incidence. The incidence of heart failure continues to increase in our country and worldwide, and mortality rates are still at very high levels. Individuals' survival and lifespans are increasing as a result of advances in cardiovascular disease care. As a result, monitoring and treating patients with heart failure is becoming increasingly relevant, and it continues to be a fertile ground for new research and innovations[10], [11].

Researchers successfully use various artificial neural networks to conduct difficult diagnostic tasks in medical applications, such as classifying diseases such as heart disease and diabetes. The ability of ANNs to process a large amount of data during the training phase and reduce the necessary diagnostic time is the reason for their success [12], [13]. The most abstracted type of ANNs is the single-layer perceptron (SLP), which has only two input and output layers [14]. It has been shown that SLPs are incapable of handling nonlinearly separable patterns effectively [15]. In this regard, multilayer perceptron (MLP) NNs have been proposed, which employ one or more hidden layers in the ANNs to avoid the drawbacks of SLPs. As a result, the MLP edition of ANNs is the most widely used [16]. MLP's learning speed, nonlinearity, parallelism, fault tolerance, robustness to noise, and excellent generalization capacity are just a few of its main features [13].

The MLP model, which is one of the ANN models, was applied to the open-source data set "Heart Failure" in this analysis. The MLP ANN model was used to estimate various factors (explanatory variables) that may be correlated with heart failure (dependent variable), and performance metrics were obtained. The model's accuracy, balanced accuracy, sensitivity, specificity, PPV, NPV, F1-score, and AUC values were 93.9%, 89.2%, 98.4%, 80.0%, 93.8%, 94.1%, 96.1%, and 0.925, respectively, according to the experimental results. The results of a study using the same data set were obtained using a variety of machine learning models, with the "Random Forest" model reporting the highest accuracy of 0.74 [8]. Another study using the same data set stated that the accuracy was 0.86 [17]. When this study is compared with the other two studies using the same data set, it is seen that the accuracy value obtained with the MLP ANN model used in this study is higher than the others with 93.9%.

Consequently, in the analysis with the heart failure data collection, the MLP ANN model generated very positive results. Moreover, this model has gained important information in identifying risk factors that may be associated with heart failure. Thus, it has been understood that the relevant model will provide reliable information about any disease to be used in preventive medicine practices.

REFERENCES

[1] Loehr, L. R., Rosamond, W. D., Chang, P. P., Folsom, A. R., & Chambless, L. E. (2008). Heart failure incidence and survival (from the Atherosclerosis Risk in Communities study). *The American journal of cardiology*, *101*(7), 1016-1022.

[2] Lindenfeld, J. O. A. N. N., Krause-Steinrauf, H., & Salerno, J. U. D. I. T. H. (1997). Where are all the women with heart failure?. *Journal of the American College of Cardiology*, *30*(6), 1417-1419.

[3] Maestre, A., Gil, V., Gallego, J., Aznar, J., Mora, A., & Martín-Hidalgo, A. (2009). Diagnostic accuracy of clinical criteria for identifying systolic and diastolic heart failure: cross-sectional study. *Journal of evaluation in clinical practice*, *15*(1), 55-61.

[4] Gil, V., Merino, J., García, M., Jordán, A., & Guerrero, L. (2006). Validity of Framingham criteria as a clinical test for systolic heart failure. *Revista clinica espanola*, *206*(10), 495-498.

[5] Öztemel, E. (2003). Yapay sinir ağlari. *PapatyaYayincilik, Istanbul*.

[6] Haykin, S. (1999). A comprehensive foundation on neural networks.

[7] Ayşe, A. R. I., & Berberler, M. E. (2017). Yapay sinir ağları ile tahmin ve sınıflandırma problemlerinin çözümü için arayüz tasarımı. *Acta Infologica*, *1*(2), 55-73.

[8] Chicco, D., & Jurman, G. (2020). Machine learning can predict survival of patients with heart failure from serum creatinine and ejection fraction alone. *BMC medical informatics and decision making*, *20*(1), 1-16.

[9] Orhan, U., Hekim, M., & Özer, M. (2010, April). Discretization approach to EEG signal classification using Multilayer Perceptron Neural Network model. In *2010 15th National Biomedical Engineering Meeting* (pp. 1-4). IEEE.

[10] Bleumink, G. S., Knetsch, A. M., Sturkenboom, M. C., Straus, S. M., Hofman, A., Deckers, J. W., ... & Stricker, B. H. C. (2004). Quantifying the heart failure epidemic: prevalence, incidence rate, lifetime risk and prognosis of heart failure: the Rotterdam Study. *European heart journal*, *25*(18), 1614-1619.

[11] Mosterd, A., & Azadas, A. W. (2007). Epidemiología clínica de la insuficiencia cardíaca. *Heart*, *93*, 1137-1146.

[12] Heidari, A. A., Faris, H., Aljarah, I., & Mirjalili, S. (2019). An efficient hybrid multilayer perceptron neural network with grasshopper optimization. *Soft Computing*, *23*(17), 7941-7958.

[13] Faris, H., Aljarah, I., & Mirjalili, S. (2016). Training feedforward neural networks using multi-verse optimizer for binary classification problems. *Applied Intelligence*, *45*(2), 322-332.

[14] Hu, Y. C. (2014). Nonadditive similarity-based single-layer perceptron for multi-criteria collaborative filtering. *Neurocomputing*, *129*, 306-314.

[15] Ojha, V. K., Abraham, A., & Snášel, V. (2017). Metaheuristic design of feedforward neural networks: A review of two decades of research. *Engineering Applications of Artificial Intelligence*, *60*, 97-116.

[16] Chen, J. F., Do, Q. H., & Hsieh, H. N. (2015). Training artificial neural networks by a hybrid PSO-CS algorithm. *Algorithms*, *8*(2), 292-308.

[17] KUCUKAKCALİ, Z. T., ÇİÇEK, İ. B., GÜLDOĞAN, E., & ÇOLAK, C. (2020). Assessment of Associative Classification Approach for Predicting Mortality by Heart Failure. *The Journal of Cognitive Systems*, *5*(2), 41-45.

BIOGRAPHIES

Mehmet Onur Kaya received his BSc from Fırat University in 2009. He received his MSc degree from Fırat University Applied Statistics in 2011 and his PhD degree from Bursa Uludağ University Biostatistics Department in 2016. His research interests are clinical biostatistics, meta-analysis, data mining, expert systems, artificial intelligence, bioinformatics and statistical shape analysis. He joined the Fırat University Biostatistics and Medical Informatics Department as an Assistant Professor in 2017 and is still a faculty member. He is engaged in teaching and research activities in artificial intelligence, data mining, statistical shape analysis and meta-analysis.