INTERNATIONAL JOURNAL OF Telecommunications
AIR DEFENSE COLLEGE

Heart Diseases Diagnosis Based on Artificial Neural Network

Mohamed E. EL-Bouridy [1], Amira Salah EL-Batouty [2] , Mohamed A. Massoud [3], and Wael Abouelwafa Ahmed [4]

[1]Alexandria Higher Institute of Engineering and Technology (AIET), mohamed.elbouridy@aiet.edu.eg, [2] Alexandria Higher Institute of Engineering and Technology (AIET) amira.elbatouty@aiet.edu.eg, [3] Department of Biomedical, Minia University, Minia, 61519, Egypt. dr.massoud@mu.edu.eg; massoud300@yahoo.com, [4] Department of Biomedical, Minia University, Minia, 61519, Egypt. wael.wafa@minia.edu.eg.
.

Citation: Elbouridy , Mohamed E .; EL-Batouty, Amira Salah.; Ahmed , Massoud, Mohamed; Ahmed , Wael Abouelwafa. . Heart Diseases Diagnosis Based on Artificial Neural Network.

Inter. Jour. of Telecommunications, ***IJT'2024****, Vol. 04, Issue 01, pp. 01-7, 2024.*

Editor-in-Chief: *Youssef Fayed.*

Received: 04/03/2024.

Accepted: 27/04/2024.

Published: 27/04/2024.

Publisher's Note: The International Journal of Telecommunications, IJT, stays neutral regard-ing jurisdictional claims in published maps and institutional affiliations.

Abstract: Heart disease is the term used to describe any ailment that is unavoidable and damages the heart. Machine learning (ML) techniques have been used by computer systems to support diagnosis in the field because of the vast amount of electronic health data that is now available. The data in this study is classified using a multi-layered perceptron (MLP) that was trained using a back-propagation (BP) artificial neural network (ANN). Heart diseases with or without heart attacks are classified using the MLP. By repeatedly going through layers of functions, ANN may identify patterns in the data and subsequently build a model. This study spans more than a thousand epochs and uses sigmoid activation functions. Experiments with different learning rates and neuron number values produced the greatest results. The results showed 25 neurons and a 0.25 learning rate with a high accuracy of 80.66%. It is discovered that ANN can be applied to categorize cases of heart disease.

Keywords: ML1; MLP2; BP3; ANN4; KNN5.

1. Introduction

Over 17.9 million people lost their lives from cardiovascular disease, and 80% of the total deaths are caused by coronary heart disease [1]. The coronary arteries in the blood veins that provide blood to the heart might narrow or get blocked in one of the many different types of heart disease. This steadily develops over time and is known as coronary artery disease. This is the main cause of heart attacks in people. Heart disease is any unpreventable circumstance that is affecting a human's heart [2]. Heart illness comes in a wide variety of forms, and each has a unique effect on the heart. Coronary heart disease, congenital heart problems, arrhythmia, myocardial infarction, and other prevalent kinds of heart illness are only a few. Damage to the heart or blood arteries can result in heart disease [3]. It may also occur if there is a deficiency in nutrition and oxygen. By performing physical examinations and blood testing, heart disease can be identified. The following are examples of noninvasive tests: echocardiograms, stress tests, carotid ultrasounds, Holter monitors, and others [4]. Due to the abundance of electronic health data available today, computer systems have entered the area to support diagnosis using machine learning techniques.

Machine learning (ML) is a powerful tool that helps computers learn how to solve problems. Its flexibility and power make it a great fit for many different fields, including medical research [5]. There are many studies focused on applying machine learning to medical diagnosis, and this is just one field where its potential is high. Frank Rosenblatt put up the ANN method in 1958. While we have enough computational power to run the algorithm, this technology is flourishing now. Heart illness was frequently classified using ANN. There are many ways to categorize cardiac illness, including statistical classification, machine learning techniques, and

symbolic classification, often known as rule-based classification. K-Nearest Neighbor (KNN) is a straightforward algorithm that is frequently and widely used in classification problems. Recent research has shown that neural network-based classification produces promising results, in which the neural network represents the problem layer by layer using an N-dimensional tensor. It essentially uses distance functions to allocate and group the nearest data. Euclidean, and Manhattan, distance are the available functions [6]. The functions will determine the separation between the points, and if the points are close together, they will be grouped. This allows any data to be categorized into a group based on its location. The data are then plotted in n-dimensional space by Support Vector Machine (SVM) to begin working. The "N" stands for the number of features a dataset contains. After that, a classifier will be used to divide the data into various groups by drawing lines. In three or more dimensions, the line is referred to as a hyperplane [7]. It strengthens the data points by locating the planes with the greatest margin, making the classification outcomes more reliable. By processing labeled training datasets, ANN can learn. An ANN can determine the trends in the data after performing layers of functions, and after that, it can create a model. With input testing data, the model can be utilized to do categorization [8]. Additional training data will boost the quality and accuracy of the model.

Thus, some earlier studies have established automatic and semi-automatic diagnoses utilizing ANN to increase the accuracy of diagnosis and decrease reliance on human specialists. Performing the test to investigate whether they had any heart attacks that they snapped out of, could help the doctor's assistance in further identifying them.

In this presented research paper, it will be dealt with through a review of related work, methodology, data gathering, results and discussion and finally the conclusion.

2. RELATED WORK

ANN has long been employed in the diagnosis of medical conditions. With just a few characteristics, the neuron model can help clinicians make diagnoses [9]. It is difficult to grasp how categorization accuracy and data properties relate to one another. Certain classification methods are only effective with a particular set or variety of features. That will significantly alter the outcomes [10]. In addition to characteristics, input data for algorithms also affects results. Several algorithms determine data memorization in different ways. A set of labeled data is required for supervised learning to complete the training [11].

Only a few academics have contrasted the accuracy of various ML algorithms. The techniques include Logistic Regression, Hybrid Random Forest with a linear model, Decision Tree (C5.0), Neural Network, Support Vector Machine (SVM), K-Nearest Neighborhood (KNN), and Neural Network (HRFLM). They discovered that C5.0 had the highest accuracy (93.0%) based on the outcome. With an accuracy of 88.7%, 88.37%, 86.05%, and 80.23%, the remaining models are HRFLM, KNN, SVM, and neural networks. They have discovered that input and dataset can affect a model's performance [12].

One intriguing SVM classification result for heart disease demonstrated greater accuracy with fewer dataset features. With only 6 of the 13 features, SVM achieved an accuracy of 72.55%. It triumphs over a test that employs all 13 features, which only achieves accuracy of 61.93% [12]. Second, the research noted that to get results that are more accurate, extensive tests should be conducted utilizing a larger dataset [13]. In, [14] asserted that hybrid random forest with a linear model (HRFLM) had an accuracy level of 88.7% in the prediction of heart disease. Moreover, [15] provides an overview of various Machine Learning (ML) and Deep Learning (DL) approaches for the disease diagnosis, that provided a comparative analysis of ML classifiers and DL models where the accuracy of classification using Convolutional Neural Network (CNN), (97.6%) and Random Forest (RF) (96.93%) have outperformed other algorithms.

The purpose of this study is to evaluate how well MLP taught by BP can categorize heart illness given that ANN can detect heart disease. Next, this study will use unprocessed photos and deep learning neural networks to support automated heart disease diagnosis systems.

2. Methodology

The three stages of this study are data collection, experiment planning, and evaluation. As shown in Figure 1, the three steps were carried out in order. The procedure of creating an appropriate dataset for this study and thoroughly describing each dataset attribute are the first phases. The data will be compiled in the second stage. The next step for researchers is to train and test the dataset using various input, such as learning rate and number of neurons. The analysis of the outcome is the last stage.

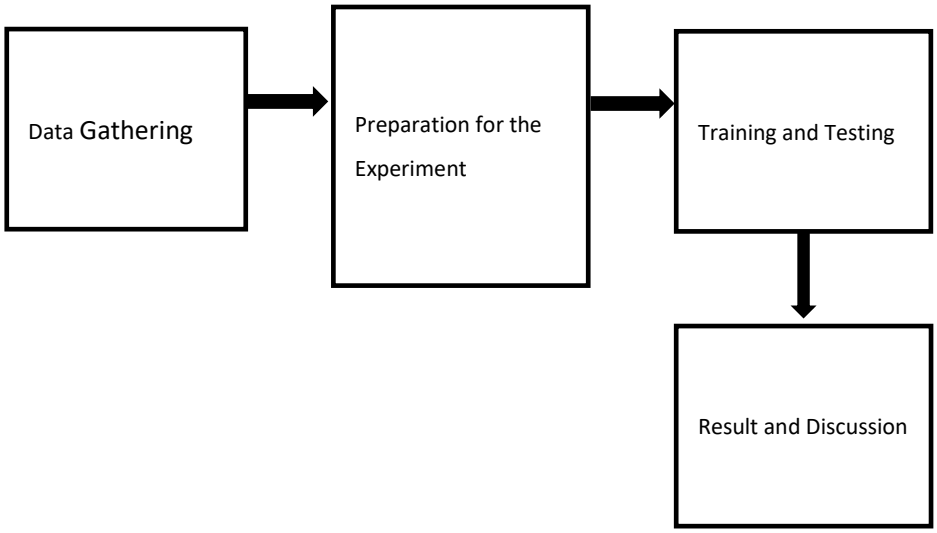

Fig. 1. Methodology Steps.

2.1 Data Gathering

During the data collection phase, a dataset was produced from (http://archive.ics.uci.edu/ml/datasets/heart+disease). The dataset consists of 303 cases in total, and 6 entries include some missing values. The following phase uses the remaining 297 instances after 6 entries were eliminated from the dataset. The characteristics included in Table I [10] include age, sex, the type of chest pain, and more. The effectiveness of the model created using the 13 characteristics and ML approach is tracked.

The data has all been converted to floating-point normal form. A 7:3 data allocation suggests that 70% (208 instances) of the sample will be utilized as a training set and 30% (89 instances) as a testing set.

Preparation for the Experiment

The input, output, and hidden layers of ANN nodes are divided into three groups. The ANN's structure can be succinctly explained as follows. The net inputs are as follows for each neuron j in the hidden layer and neuron k in the output layer:

$$\boldsymbol{net_j = \sum_i w_{ji} * o_i} \quad \boldsymbol{and} \quad \boldsymbol{net_k = \sum_j w_{kj} * o_j} \qquad (1)$$

where *i(j)* is a neuron from the previous layer, *oi(oj)* is a node's output, and *wji (wki)* is the weight of the connection between *i(j)* and *j(k)*. These are examples of the neuron outputs:

$$o_i = net_i \quad (2)$$

$$o_i = \frac{1}{1 + e^{-(net_i + \theta_i)}} = f_i(net_i, \theta_i) \quad (3)$$

If *j (k)* is a bias and net_j (net_K) is the input signal from the external source to node *j(k)* in the input layer. The most often used transformation function to date is the sigmoid function depicted in Equations (2) and (3). The effectiveness of MLP who had BP training will be examined in the study. In this investigation, an epoch of 1000 is employed with the sigmoid activation function [16]. Table II displays the performance model neurons with various learning rates and hidden layer neuron counts.

3. Results and Discussion

Researchers put the ANN to the test using various inputs to assess the performance results. The learning rates range from 0.25 to 0.75 to 1.0, accordingly. For each learning rate, different numbers of neurons (5, 10, 15, 20, 25) are examined. Figure 2 compares how well different parameters for learning rate and number of neurons number settings for learning rate and number of neurons.

adjustments to performance. The maximum accuracy was measured at an epoch of 1000 and a learning rate of 0.25 with 25 neurons. The findings show that the MLP neural network trained by the BP algorithm is capable of classifying heart illness that includes heart attack symptoms or not.

4. CONCLUSIONS

One of the biggest problems facing the modern world and a leading cause of death worldwide is heart disease. Recent advancements in the application of machine learning (ML) demonstrate that early detection of cardiac illness is achievable with the use of patient data and electrocardiograms. An important duty that can help doctors and technologists in their screening procedure is the automated classification of cardiac disease. To perform the task, many ML models have recently been adopted or modified. The use of BP learning algorithms in facilitating MLP models has been demonstrated by this work. Instead of focusing only on theoretical techniques and simulations, future work on this research will involve studies for real-world datasets. The results point to promising future research that will help the medical field. It is also recommended to further study the impact of using deep learning theories due to the accurate results they achieve in diagnosis, as well as study the effect of increasing noise on the model presented in the research.

TABLE I. The Characteristics Description.

characteristics	The Description
Age	years of age
Sex	0= female, 1= male
Cp	If there is no anginal pain, the condition is asymptomatic.
Ca	Major vessel count (0–3) colored by fluorescence
Fbs	Blood sugar level at dawn was 120 mg/dl. 0 = false, 1 = true.
Chol	Mg/dl of serum cholesterol
Trestbps	mm of resting blood pressure
Thal	Normal is three, fixed defects are six, and reversible defects are seven.
Oldpeak	Exercise includes ST depression compared to reset
Restecg	Electrocardiographic results when at rest
Exang	Angina brought on by exercise 1=yes, 0=No
Slope	Fluorescence-colored main vessels (0:3) in number
Thalach	Reached maximum heart rate

TABLE II. The train and test datasets with various learning raters based on the number of neurons.

Percentage Accuracy	Learning rates	Number of neurons
70.09 %	1	5
64.46 %		10
56.64%		15
45.8 %		20
48.51 %		25
71.78 %	0.75	5
68.66 %		10
73.65 %		15
45.67 %		20
55.56 %		25
75.89 %	0.5	5
79.27 %		10
78.04 %		15
79.47 %		20
55.63 %		25
79.58 %	0.25	5
80.26 %		10
78.42 %		15
78.79 %		20
80.66 %		25

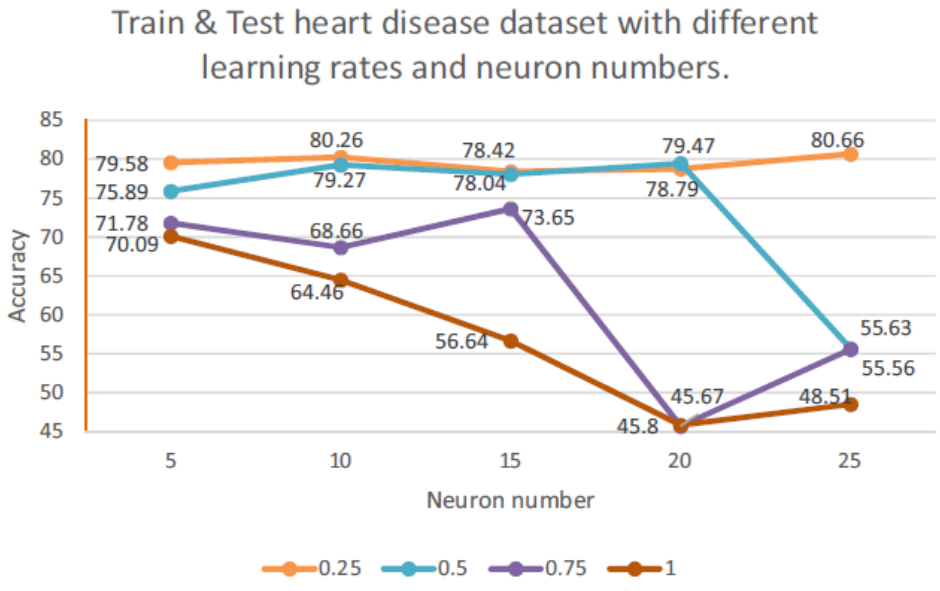

Fig. 2. The train and test datasets with various learning raters based on the number of neurons.

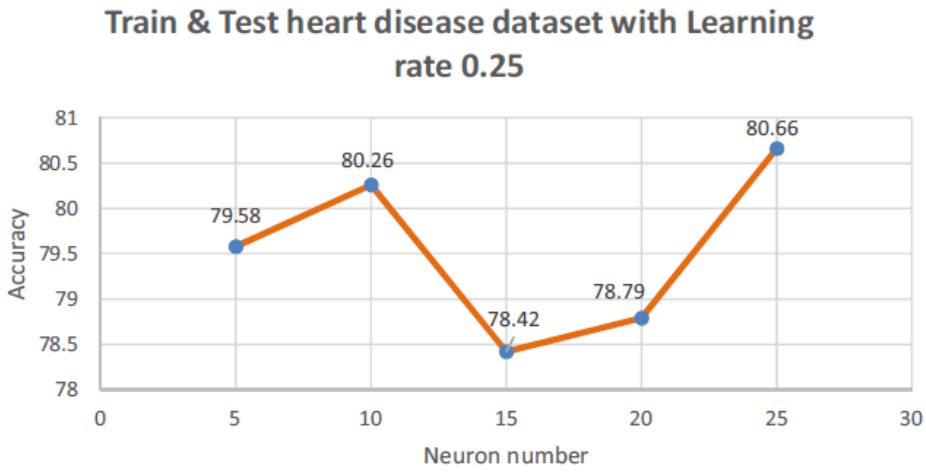

Fig. 3. The train and test datasets with learning rater 0.25.

REFERENCES

[1] World Heart Federation What Is CVD 2020 [Internet] Retrieved from https://world-heart-federation.org/world-heart-day/about-whd/world-heart-day-2020 on July 15, 2021.

[2] Medical News Today Heart Disease: Types Causes and Treatments 2020 [Internet] Retrieved from https://www.medicalnewstoday.com/articles/237191 on July 15th, 2021.

[3] New Straits Time Heart disease top killer of Malaysians in 2019 2020 [Internet] Retrieved from https://www.nst.com.my/news/nation/2020/11/644515/heart-disease-top-killer-malaysians-2019 on July 15th, 2021.

[4] Beunza JJ Puertas E García O Villalba G Condes E et al 2019 Comparison of machine learning algorithms for clinical event prediction (risk of coronary heart disease) J Biomed 97 103527.

[5] Carbonell J Michalski R and Mitchell T 1983 An Overview Of Machine Learning. Machine Learning 3-23.

[6] Abdar M Kalhori SRN Sutikno T Subroto IMI Arji G 2015 Comparing Performance of Data Mining Algorithms in Prediction Heart Diseases International Journal of Electrical and Computer Engineering (UECE) 5 1569.

[7] Towards Data Science Support Vector Machines (SVM) - An Overview Towards data science 2019 [Internet] Retrieved from https://towardsdatascience.com/https-medium-com-pupalerushikesh-svm-f4b42800e989 on July 15th, 2021.

[8] Mastery What is Deep Learning Machine Learning 2020 [Internet] Retrieved from https://machinelearningmastery.com/what-is-deep-learning on July 15th, 2021.

[9] Akhil M Deekhatulu BL Chandr P 2003 Classification of heart disease using K- nearest neighbor and genetic algorithm Int Conf Comput Intell Model Tech Appl Proc Technol 10 85-94.

[10] Sellappan P Awang R 2008 Intelligent Heart Disease Prediction System Using Data MiningTechniques LJCSNS International Journal of Computer Science and Network Security 8 (8).

[11] Maind SB Wankar 2014 P Research Paper on Basic of Artificial Neural Network International Journal on Recent and Innovation Trends in Computing and Communication 2 96-100.

[12] Sumit B Praveen P Pillai GN 2008 SVM Based Decision Support System for Heart Disease Classification with Integer-Coded Genetic Algorithm to Select Critical Features WCECS Proceedings of the World Congress on Engineering and Computer Science 22-24.

[13] Foody G McCulloch M Yates W 1995 The effect of training set size and composition on artificial neural network classification International Journal of Remote Sensing 16 1707-1723.

[14] Mohan S Thirumalai C Srivastava G 2019 Effective heart disease prediction using hybrid machine learning techniques IEEE Access 7 81542-81554

[15] Meghavi Rana & Megha Bhushan, 2022, Machine learning and deep learning approach for medical image analysis: diagnosis to detection, Multimedia Tools and Applications (2023) 82:26731–26769, https://doi.org/10.1007/s11042-022-14305-w

[16] Jeremy Jordan Neural networks: activation functions 2019 [Internet] Jeremy Jordan Retrieved from https://www.jeremyjordan.me/neural-networks-activation-functions on July 15th, 2021.