
Data Mining for Heart Disease Prediction Based on Echocardiogram and Electrocardiogram Data

Tb Ai Munandar

Department of Informatics, Universitas Bhayangkara Jakarta Raya, Indonesia

Article Info

Article history:

Received Feb 12, 2023

Revised April 05, 2023

Accepted May 19, 2023

Keywords:

Data mining

Echocardiogram

Electrocardiogram

Heart disease

Multi-layer perceptron

ABSTRACT

Traditional methods of detecting cardiac illness are often problematic in the medical field. The doctor must next study and interpret the findings of the patient's medical record received from the electrocardiogram and echocardiogram. These tasks often take a long time and require patience. The use of computational technology in medicine, especially the study of cardiac disease, is not new. Scientists are continuously striving for the most reliable method of diagnosing a patient's cardiac illness, particularly when an integrated system is constructed. The study attempted to propose an alternative for identifying cardiac illness using a supervised learning technique, namely the multi-layer perceptron (MLP). The study started with the collection of patient medical record data, which yielded up to 534 data points, followed by pre-processing and transformation to provide up to 324 data points suitable to be employed by learning algorithms. The last step is to create a heart disease classification model with distinct activation functions using MLP. The degree of classification accuracy, k-fold cross-validation, and bootstrap are all used to test the model. According to the findings of the study, MLP with the Tanh activation function is a more accurate prediction model than logistics and Relu. The classification accuracy level (CA) for MLP with Tanh and k-fold cross-validation is 0.788 in a data-sharing situation, while it is 0.672 with Bootstrap. MLP using the Tanh activation function is the best model based on the CA level and the AUC value, with values of 0.832 (k-fold cross-validation) and 0.857 (bootstrap).

Corresponding Author:

Tb Ai Munandar,

Informatics Department, Faculty of Computer Science, Universitas Bhayangkara Jakarta Raya,

Jl. Raya Perjuangan Bekasi Utara, Kota Bekasi, Jawa Barat 17121, Indonesia

Email: tbaimunandar@gmail.com

1. INTRODUCTION

Heart disease is one of the main causes of death in many parts of the world [1], [2], [3], and it may affect anybody, regardless of age, gender, or socioeconomic status [4]. If this sickness infects humans, it is extremely deadly [5] since it is difficult to anticipate adequately and efficiently [6]. Traditional analytical approaches are used by clinicians, particularly in hospitals, to diagnose heart illness. The study was based on the patient's observation findings, which were then combined with the results of the ECG and echocardiography data examinations. However, normal medical analysis is frequently ineffectual and has little consequence on diagnosis [7]. Even doctors make mistakes while diagnosing heart issues [8]. Traditional medical analysis frequently struggles to match diagnostic findings with other diagnostic data due to the enormous amount of medical record data.

The health business now generates a large amount of medical record data [9], but they lack a very powerful analysis tool that can be used to diagnosis disease kinds based on hidden relationship information in the data [10]. An artificial intelligence approach may be used in computer science to

extract hidden information from data for decision-making purposes. Artificial intelligence is not a new concept, particularly in the detection of cardiovascular disease. A decade ago, several research on the application of artificial intelligence to the difficulty of diagnosing heart illness were conducted [11], [12], [13], [14], and [15]. Data mining is a branch of artificial intelligence that is frequently used to detect heart illness.

Data mining is the process of detecting patterns and correlations in data for data analytic purposes such as trend forecasting, classification, and grouping [9]. Furthermore, it has the potential to predict heart disease and identify unique illnesses such as cardiac amyloidosis (CA) [16]. Several other studies have used data mining to diagnose cardiac illness, including [17], [18], [19], and [20]. The bulk of researchers execute data mining tasks using machine learning [21], [22], and [23].

Several studies have been conducted to diagnose cardiac disease using simply echocardiography or electrocardiogram (ECG) data. Echocardiogram parameters have been used in studies [24], [25], and [26] to identify heart disease and other conditions. Meanwhile, [27], [28], [29], and [30] employ ECG characteristics to detect cardiac disease and other conditions. The present study employs a data mining method to integrate these two factors in order to generate more accurate identification results.

Despite significant research into using data mining and machine learning techniques to predict heart disease, it remains difficult for health practitioners and academics to use data mining to aid in research and diagnose illnesses based on patient medical record data. This work contributes to a comprehensive understanding of the stages of data mining in the health industry, particularly in the prediction of heart disease. The discussion chapter uses ECG data to show how a multilayer perceptron technique with three activation functions may be utilized to forecast heart disease. This work employs training and testing data sharing approaches such as k-fold cross-validation and bootstrapping to assess the accuracy of the three MLP activation functions utilized to test the MLP model. The metrics area under the curve (AUC), classification accuracy (CA), F1, precision, and recall are used to determine the optimum activation function of MLP.

2. METHOD

This study starts with identifying the issue and ends with the collection of patient medical data from one of Banten Province's government general hospitals. The information was gathered manually by collecting patient medical records, which were then combined into a single file to aid in the analytic process. There were 354 data occurrences in the initial data collection results. Following data cleaning and preprocessing, 324 valid datasets were retrieved. There are fifteen predictor variables and one response variable in the dataset. The fifteen prediction criteria were established based on echocardiography and ECG data gathered from the medical records of heart disease patients. The data is ready for use after the preparation step. The *.csv file extension is associated with data analysis. The data is then separated into training data and testing data. For training data, there are two distribution possibilities. The first example uses k-fold cross-validation, whereas the second uses bootstrapping. After that, the training data is sent into the MLP algorithm, which generates a model. Three activation functions are used for training the algorithm using training data. The three roles of activation are logistics, Tanh, and Relu. Validation of the model is performed using scenario-specific test data. The last step is model selection. The chosen model is then used to predict new data, which is then examined. This study also used Orange Data Mining (ODM), a visual programming tool. Several widgets are used to facilitate analysis. The *.xls extension is changed to *.csv for simpler display on ODM work screens.

3. RESULT AND DISCUSSION

3.1. Result

3.1.1. Business and data understanding and preparation

Data mining begins with business knowledge and progresses to data comprehension. In the first phase, we define research objectives, which are subsequently translated into issue descriptions and handled by data mining. We also created a research goal at this time, which was to create an MLP-based model for predicting cardiac illness.

Following the identification of objectives and impediments, the next stage of data understanding is data gathering. The data was gathered from the medical records of heart disease patients at one of Banten Province's regional public facilities. There are fifteen predictor factors and one response variable. The fifteen characteristics were generated from the medical records of individuals who had undergone echocardiogram (ECHO) and electrocardiogram (ECG). Echocardiography data comprises cardiac functions, EDD (end-diastolic diameter), ESD (end-systolic diameter), IVS Diastole, IVSSystole, PW (posterior wall) Diastole, and PW (posterior wall) Systole. HR (heart rate in pulses per minute), PR (PQ), QRS, QT, QTC, and P waves are ECG variables.

Each characteristic has its own heart disease distribution. In the LEFT-ATRIUM variable, for example, the greatest distribution of disease types occurs between 40 and 42:48, with hypertensive heart disease (41.67%), coronary artery disease (39.58%), left ventricular HSHHD (12.50%), rheumatic heart disease (4.17%), and normal resting echocardiography (2.07%) dominating the distribution. Numerous kinds of cardiovascular illness have varying distributions as well. Figures 1 and 2 show the distribution and relative percentages of the LEFT-ATRIUM and end-diastolic diameter (EDD) variables for all forms of heart disease.

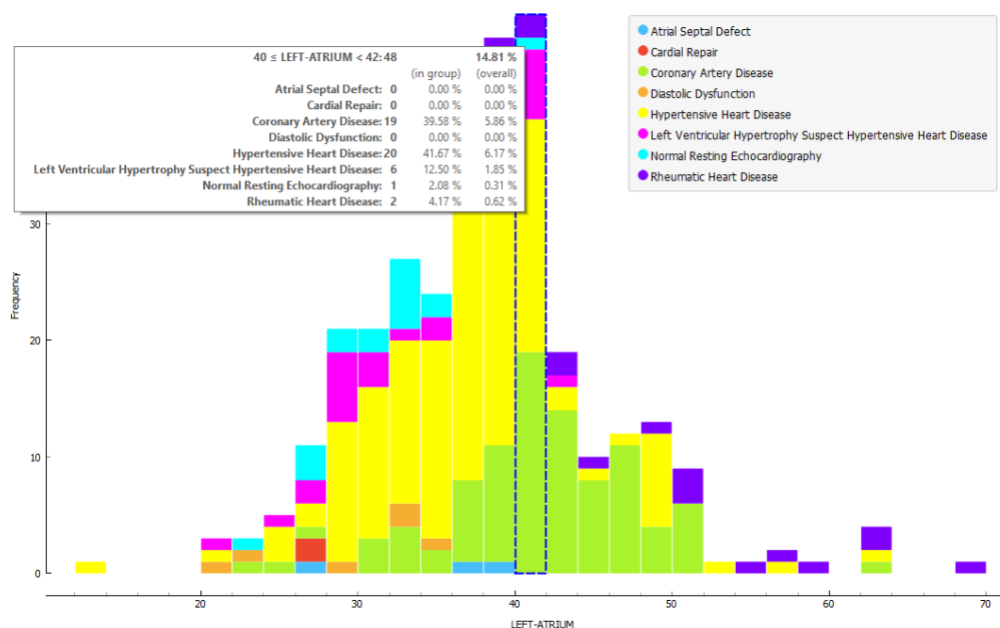


Figure 1. Left-Atrium Variable Value Range Distribution for Disease Types

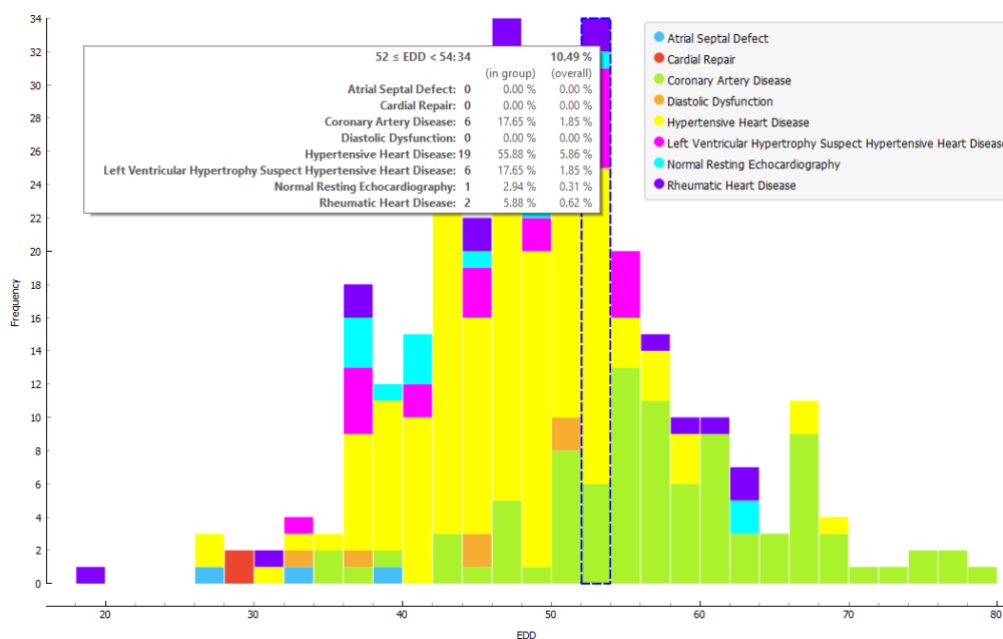


Figure 2. The range of End-Diastolic Diameter (EDD) values for different kinds of disease

Following a physician's diagnosis, 354 patient medical record files comprising eight types of heart disease were first obtained. The process of selecting attributes and purifying the data occurs during the data comprehension step. The data purification technique yielded 324 useable data instances with eight class labels corresponding to disorder types. The eight categories are as follows: atrial septal defect, cardiac repair, coronary artery disease, diastolic dysfunction, hypertensive heart disease, left ventricular hypertrophy indicative of hypertensive heart disease, normal resting echocardiography, and rheumatic heart disease. The distribution of heart disease categories throughout the sample is seen in Figure 3. The distribution of heart disease categories throughout the sample is seen in Figure 3.

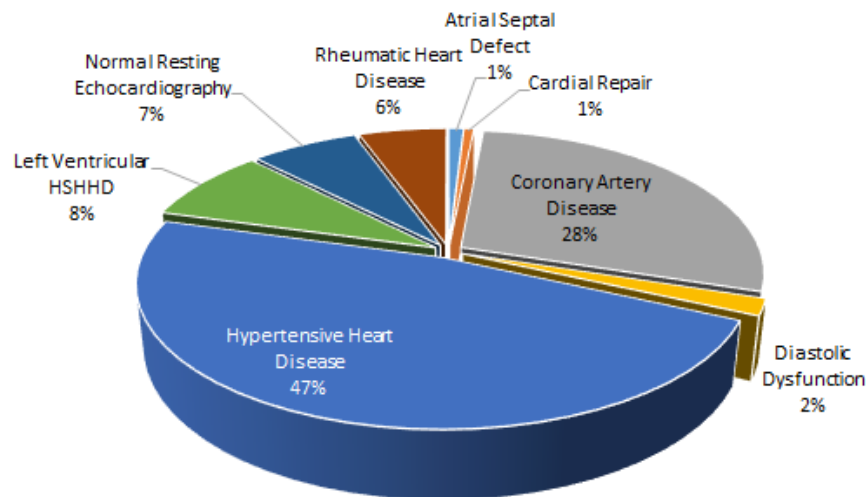


Figure 3. Distribution of types of heart disease in the dataset

The goal of the data preparation step is to create datasets that are useable and analyzable. Data tabularization with columns representing attributes and rows representing data instances. To simplify later analysis, the data is saved in *.csv files.

3.1.2. Building a classification model

The difficulty of constructing a classification model is part of data mining. The multilayer perceptron (MLP) technique is employed as a learning algorithm in this stage. To determine the best accurate MLP with activation, three activation functions are utilized. There are two main possibilities for model creation. The current scenario is primarily concerned with how the dataset is divided into training and testing data before being incorporated in the algorithm. The first case employed k-fold cross-validation (k-FCV), whereas the second used Bootstrap.

3.1.2.1 Classification model with k-fold cross-validation division

Because the k-FCV approach has a value of 10, the dataset is divided into ten parts, and ten iterations of training and accuracy tests are done until the model's final average accuracy is obtained. Training data is often separated into 291 training data sets divided into nine divisions and 33 testing data sets (with a separate section for model performance testing). The model was created using the multilayer perceptron (MLP) method. Training data are used to train the MLP algorithm. Logistic, Relu, and Tanh are examples of activation functions that have been employed. The purpose of comparing the three activation functions is to determine which one has the highest degree of accuracy, so that it may be utilized as the most accurate model for predicting heart disease.

The model is then evaluated using test data. According to the model test results, the logistic activation function and Relu have the same level of classification accuracy (CA), which is 0.758, while the Tanh activation function has a higher level of CA, which is 0.788. Despite having the same accuracy as Logistic, Relu has a superior activation function in this scenario. The fact that Relu has a greater accuracy rating than logistics confirms this. This precision grade shows the correctness of the predicted findings in relation to the actual data. Figure 4 displays an in-depth examination of the CA, F1, accuracy, and recall component qualities.

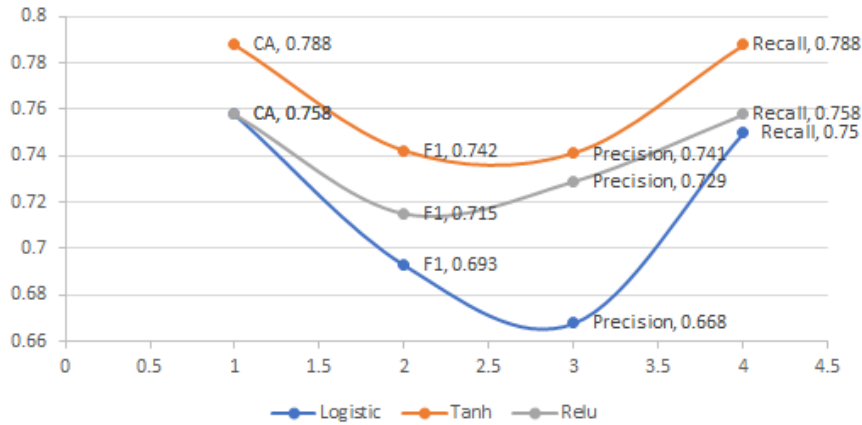


Figure 4. Comparison of CA, F1, Precision, and Recall by Dividing the k-fold Cross Validation Dataset

3.1.2.2 Classification model with Bootstrap

When constructing a training set, Bootstrap splits the data into random samples. Random sampling is employed on the dataset. Orange Data Mining separates the dataset into training and testing data using the Bootstrap option. For each occurrence of data in the dataset, the proportion estimated from training data is 32:81. The great majority of the data is utilized for training, with the remainder being used for testing. The training data is then utilized to train the MLP algorithm with three distinct activation functions, as in the k-fold cross-validation scenario. Data testing is used to validate the model that has been created. The Tanh activation function has a greater classification accuracy (CA) than the other two activation functions, according to the accuracy test results. The Logistic, Tanh, and Relu activation functions had classification accuracy scores of 0.664, 0.672, and 0.688, respectively. Tanh has a higher accuracy value than logistcs and relu by default. Figure 5 shows a comparison of the three activation functions' accuracy.

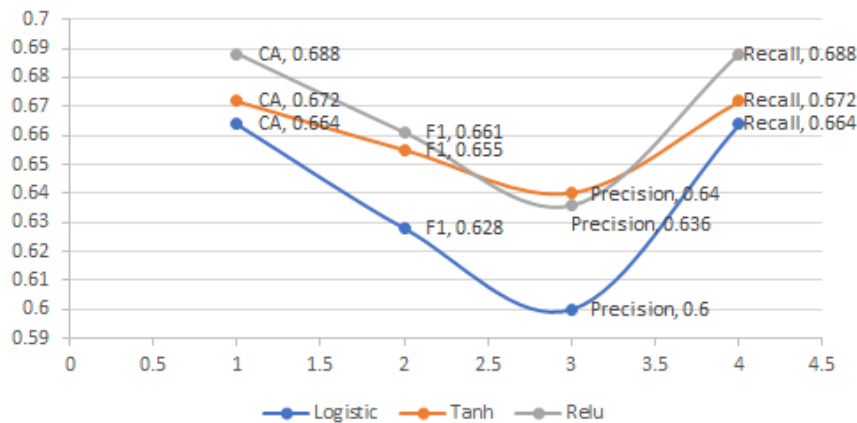


Figure 5. Comparison of CA, F1, precision, and recall by dividing the Bootstrap dataset

3.1.3 Model evaluation

During this stage, it is determined which activation function-based model will be used to forecast the kind of heart disease. Classification accuracy (CA), F1, precision, recall, and AUC data are used to assess models. comparing the evaluation component values for the two instances According to the results of the first scenario model, the MLP with the Tanh activation function achieves the highest degree of accuracy when compared to the other two activation functions. Tanh's activation function provides superior classification accuracy than Logistic and Relu in the second scenario outcome model.

Tanh's MLP model is often not better than Relu's MLP model in the first scenario, despite having a greater degree of classification accuracy than other activation functions based on the area under the receiver operating characteristic (AUC) value. AUC may have values ranging from 0 to 1. The closer the model is to 1, the better its performance. In this case, Relu's accuracy and AUC results are inconsistent.

Relu has the greatest AUC value among the three activation functions, although the degree of accuracy between predicted and actual data is the same as Tanh's.

In the second situation, the MLP using the Tanh activation function has the maximum classification accuracy (CA). Tanh outperforms Logistics and Relu in AUC, accuracy, and recall. Based on the results of model testing, the MLP with the Tanh activation function is the best classification model for both the first and second instances. Table 1 compares the evaluation of models based on AUC, CA, F1, accuracy, and recall values for the first and second situations.

Table 1. Comparison of MLP for the First and Second Scenarios

	K-fold Cross Validation			Bootstrap		
	Logistic	Tanh	Relu	Logistic	Tanh	Relu
AUC	0.832	0.864	0.868	0.845	0.857	0.843
CA	0.758	0.788	0.758	0.664	0.672	0.688
F1	0.693	0.742	0.715	0.628	0.655	0.661
Precision	0.668	0.741	0.729	0.6	0.64	0.636
Recall	0.75	0.788	0.758	0.664	0.672	0.688

3.2 Discussion

In this research, the activation function of the best model was used to predict the type of heart disease. Twenty patient data points are supplied in order to assess the patient's heart disease classification. We use physician-identified data from current medical records to confirm the classification findings. Prediction using MLP with the Tanh activation function According to the prediction findings for the extra patient data, up to fifteen patients accurately recognized the disease type that matches the actual data and predicted outcomes, while up to five patients were categorized incorrectly. This shows that the classification accuracy of the prediction model on extra data is 75%. Table 2 shows the results of the heart disease category for new data on twenty individuals.

Table 2. Prediction results for types of heart disease for additional data

Patient	Actual (Doctor's Diagnosis)	Predicted
PSJ-484	Coronary Artery Disease	Coronary Artery Disease
PSJ-489	Coronary Artery Disease	Coronary Artery Disease
PSJ-491	Coronary Artery Disease	Coronary Artery Disease
PSJ-492	Coronary Artery Disease	Coronary Artery Disease
PSJ-495	Coronary Artery Disease	Coronary Artery Disease
PSJ-496	Coronary Artery Disease	Hypertensive Heart Disease
PSJ-480	Hypertensive Heart Disease	Hypertensive Heart Disease
PSJ-482	Hypertensive Heart Disease	Hypertensive Heart Disease
PSJ-483	Hypertensive Heart Disease	Hypertensive Heart Disease
PSJ-486	Hypertensive Heart Disease	Hypertensive Heart Disease
PSJ-497	Hypertensive Heart Disease	Hypertensive Heart Disease
PSJ-485	Left Ventricular HSHHD	Hypertensive Heart Disease
PSJ-487	Left Ventricular HSHHD	Hypertensive Heart Disease
PSJ-490	Left Ventricular HSHHD	Hypertensive Heart Disease
PSJ-493	Left Ventricular HSHHD	Hypertensive Heart Disease
PSJ-478	Normal Resting Echocardiography	Normal Resting Echocardiography
PSJ-479	Normal Resting Echocardiography	Normal Resting Echocardiography
PSJ-488	Normal Resting Echocardiography	Normal Resting Echocardiography
PSJ-494	Normal Resting Echocardiography	Normal Resting Echocardiography
PSJ-481	Rheumatic Heart Disease	Rheumatic Heart Disease

In Table 3, the probability column displays the likelihood of the expected result based on MLP using the Tanh activation function. Patients with the identities PSJ-496, PSJ-485, PSJ-487, PSJ-490, and PSJ-493 had a discrepancy between actual data and predicted results, with probability values of 0.46, 0.65, 0.64, 0.46, and 0.85. The patient with the code PSJ-493 had a higher projected probability value than the other four patients whose predictions did not match.

Table 3. Probability of MLP Prediction Results with the Tanh Activation Function

Patient	CAD	HHD	Left Ventricular HSHHD	NRE	RHD
PSJ-484	0.97	0.00	0.00	0.00	0.03
PSJ-489	0.83	0.09	0.06	0.00	0.01
PSJ-491	0.82	0.10	0.02	0.02	0.02
PSJ-492	0.94	0.03	0.00	0.00	0.02
PSJ-495	0.97	0.01	0.00	0.00	0.01
PSJ-496	0.00	0.46	0.27	0.16	0.06

Patient	CAD	HHD	Left Ventricular HSHHD	NRE	RHD
PSJ-480	0.00	0.94	0.05	0.00	0.00
PSJ-482	0.35	0.40	0.02	0.00	0.22
PSJ-483	0.05	0.84	0.08	0.00	0.02
PSJ-486	0.12	0.71	0.07	0.05	0.03
PSJ-497	0.00	0.95	0.02	0.01	0.00
PSJ-485	0.01	0.65	0.27	0.02	0.01
PSJ-487	0.08	0.64	0.26	0.00	0.01
PSJ-490	0.02	0.46	0.41	0.02	0.00
PSJ-493	0.01	0.85	0.09	0.01	0.03
PSJ-478	0.02	0.45	0.04	0.47	0.01
PSJ-479	0.00	0.10	0.06	0.78	0.00
PSJ-488	0.03	0.22	0.03	0.66	0.01
PSJ-494	0.01	0.25	0.11	0.45	0.00
PSJ-481	0.01	0.14	0.00	0.00	0.84

Patients with the PSJ-496 code had a 0.46 chance of having hypertensive heart disease, a 0.27 chance of having left ventricular HSHHD, and a 0.16 chance of having a normal resting echocardiogram (NRE). It can also be seen that the likelihood of the real data being reselected as a candidate for the prediction result is 0.00, suggesting that the projected result is very unlikely to match the actual outcome. Patients with PSJ-485, PSJ-487, and PSJ-490 codes show a similar pattern. Table 3 displays the MLP's anticipated probability using the Tanh activation function.

The classification accuracy (CA) of supplementary data prediction results is 0.75, and the area under the curve (AUC) is 0.928%, both of which are near-excellent for a model. Clearly (precision), the model's accuracy in predicting data impacts the value of CA in the prediction outcomes of fresh data. With a precision value of 0.675%, a classification accuracy (CA) of 0.75 is quite satisfactory. Regardless of the kind of heart illness, the MLP model with the Tanh activation function reliably predicts it.

4. CONCLUSION

The classification accuracy of a multilayer perceptron (MLP) with several activation functions changes depending on the test data, according to the study. First, the classification accuracy (CA) for the three activation functions used differs. The logistic and Relu activation functions have CA values that are similar. Tanh has more precise activation functions than Logistics and Relu. Relu, on the other hand, maintains the model's quality level in the first scenario based on the AUC value for MLP. The Tanh activation function and the high quality of the developed classification model produce the most exact classification in the second case. In general, the MLP with Tanh activation offers the most accuracy. According to test findings on fresh data, the model can predict cardiovascular disease with 75.0% accuracy. In this work, a comparison of the MLP activation function to the Relu and Logistic activation functions reveals that Tanh is an activation function that may be selected for the prediction of heart disease. Thus, utilizing the MLP with the Tanh activation function, an alternate approach for predicting heart disease is possible.

REFERENCES

- [1] C.A. Hassan, J. Iqbal, R. Irfan, S. Hussain, A. D. Algarni, S.S.H. Bukhari, N. Alturki, and S.S. Ullah, "Effectively Predicting the Presence of Coronary Heart Disease Using Machine Learning Classifiers", *Sensors*, Vol. 22, 7227, 2022, <https://doi.org/10.3390/s22197227>
- [2] Z. He, H. Zhang, X. Chen, J. Shi, L. Bai, Z. Fang, and R. Wang, "Hemorrhagic risk prediction in coronary artery disease patients based on photoplethysmography and machine learning," *Scientific Reports*, 12:19190, 2022, <https://doi.org/10.1038/s41598-022-22719-7>
- [3] H.D. Masethe, and M.A. Masethe, "Prediction of Heart Disease using Classification Algorithms", *Proceedings of the World Congress on Engineering and Computer Science 2014*, Vol II, 22-24 October, 2014
- [4] Anderies, J.A.R.W. Tchinn, P.H. Putro, Y.P. Darmawan, A.A.S. Gunawan, "Prediction of Heart Disease UCI Dataset Using Machine Learning Algorithms", *Engineering, Mathematics and Computer Science*, Vol.4, No.3, 2022, pp. 87-93, DOI: 10.21512/emacsjournal.v4i3.8683
- [5] Y. Muhammad, M. Tahir, M. Hayat, and K.T. Chong, "Early and accurate detection and diagnosis of heart disease using intelligent computational model", *Scientific Reports*, 10:19747, 2020, <https://doi.org/10.1038/s41598-020-76635-9>
- [6] H. Jindal, S. Agrawal, R. Khera, R. Jain and P. Nagrath, "Heart disease prediction using machine learning algorithms", *IOP Conf. Series: Materials Science and Engineering 1022*, pp. 1 - 10, 2021, doi:10.1088/1757-899X/1022/1/012072
- [7] B. Umadevi, and M. Snehapriya, "A Survey on Prediction of Heart Disease Using Data Mining Techniques", *International Journal of Science and Research (IJSR)*, Volume 6, Issue 4, pp. 2228 - 2232
- [8] S. Kodati, and R. Vivekanandam, "Analysis of Heart Disease using in Data Mining Tools Orange and Weka", *Global Journal of*

- Computer Science and Technology*, Volume XVIII, Issue I, Version I, pp. 17 - 21, 2018
- [9]. S.H. Shadid, A.Shafkat, F. Yasmeen, S.A. Sibli, and M.R.Rafi, "Prediction of Heart Disease Using Data Mining Techniques: A Case Study", *International Journal of Research Publications*, Volume: 41, Issue: 1, 2019, available at <https://ijrp.org/paper-detail/826>
- [10]. B.Bahrami, and M.H. Shirvani, "Prediction and Diagnosis of Heart Disease by Data Mining Techniques", *Journal of Multidisciplinary Engineering Science and Technology (JMEST)*, Vol. 2, Issue 2, pp. 164 - 168, 2015
- [11]. O.Yu. Atkov, S.G. Gorokhova, A.G. Sboev, E.V. Generozov, E.V. Muraseyeva, S.Y. Moroshkina, N.N. Cherniy, "Coronary heart disease diagnosis by artificial neural networks including genetic polymorphisms and clinical parameters", *Journal of Cardiology*, 59, pp. 190—194, 2012, doi:10.1016/j.jcc.2011.11.005
- [12]. R. Chitra, and V. Seenivasagam, "Review of Heart Disease Prediction System Using Data Mining And Hybrid Intelligent Techniques", *ICTACT Journal on Soft Computing*, Vol. 3, No. 4, pp. 605 - 609, 2013, DOI : 10.21917/ijsc.2013.0086
- [13]. Taneja, "Heart disease Prediction System Using data Mining Techniques", *Oriental Journal of Computer Science & Technology*, Vol. 6, No. (4): Pp. 457-466, 2013
- [14]. C.S. Dangare, and S.S. Apte, "Improved Study of Heart Disease Prediction System using Data Mining Classification Techniques", *International Journal of Computer Applications*, Vol. 47, No.10, 2012
- [15]. A.K. Pandey, P. Pandey, K.L. Jaiswal, A.K. Sen, "A Heart Disease Prediction Model using Decision Tree", *IOSR Journal of Computer Engineering (IOSR-JCE)*, Vol. 12, No. 6, pp. 83-86, 2013
- [16]. S. Goto, K. Mahara, L. Beussink-Nelson, H. Ikura, Y. Katsumata, J. Endo, H. K. Gaggin, S. J. Shah, Y. Itabashi, C. A. MacRae, and R. C. Deo, "Artificial intelligence-enabled fully automated detection of cardiac amyloidosis using electrocardiograms and echocardiograms", *NATURE COMMUNICATIONS* | (2021) 12:2726, pp. 1-12, 2021, <https://doi.org/10.1038/s41467-021-22877-8>
- [17]. Akansha Jain, Manish Ahirwar, Rajeev Pandey, "A Review on Intutive Prediction of Heart Disease Using Data Mining Techniques", *International Journal of Computer Sciences and Engineering*, 2019, ff10.26438/ijcse/v7i7.109113
- [18]. Thamizhiniyal, and A. Vidhya, "Intelligent Risk Analysis and Prediction of Heart Disease Using Data Mining Techniques", *Journal of Information and Computational Science*, Vol. 10, No. 3, pp. 175 - 180, 2020
- [19]. R. Fadnavis, K. Dhore, D. Gupta, J. Waghmare, and D. Kosankar, "Heart disease prediction using data mining", *Journal of Physics: Conference Series* 1913, 012099, 2021, doi:10.1088/1742-6596/1913/1/012099
- [20]. P. Singh, S. Singh, and G.S. Pandi-Jain, "Effective heart disease prediction system using data mining techniques", *International Journal of Nanomedicine*, Vol. 13, pp. 121 - 124, 2018
- [21]. S.Y. Cho, S.H. Kim, S.H. Kang, K.J. Lee, D. Choi, S. Kang, S.J. Park, T.Kim, C.H. Yoon, T.J. Youn, and I.C. Chae, "Pre-existing and machine learning-based models for cardiovascular risk prediction", *Scientific Reports*, 11:8886, 2021, <https://doi.org/10.1038/s41598-021-88257-w>
- [22]. C. Krittanawong, H.U.H. Virk, S. Bangalore, Z. Wang, K.W. Johnson, R. Pinotti, H. Zhang, S. Kaplin, B. Narasimhan, T. Kitai, U. Baber, J.L.Halperin and W.H.W. Tang, "Machine learning prediction in cardiovascular diseases: a meta-analysis", *Scientific ReportS*, 10:16057, 2020, <https://doi.org/10.1038/s41598-020-72685-1>
- [23]. R. Alizadehsani, M. Roshanzamir, M. Abdar, A. Beykikhoshk, A. Khosravi, M. Panahiazar, A. Koohestani, F. Khozeimeh, S. Nahavandi, and N. Sarrafzadegan, "A Database For Using Machine Learning and Data Mining Techniques for Coronary Artery Disease Diagnosis", *Scientific Data* | (2019) 6:227, pp. 1 -13, 2019, <https://doi.org/10.1038/s41597-019-0206-3>
- [24]. C. Balakrishnan, and V.D.A. Kumar, IoT-Enabled Classification of Echocardiogram Images for Cardiovascular Disease Risk Prediction with Pre-Trained Recurrent Convolutional Neural Networks, *Diagnostics* 2023, 13, 775. <https://doi.org/10.3390/diagnostics13040775>
- [25]. R. Nedadur, B. Wang, and W. Tsang, Artificial intelligence for the echocardiographic assessment of valvular heart disease *Heart* 2022;108:1592-1599.
- [26]. N. Kumar, and D. Kumar, Machine Learning based Heart Disease Diagnosis using Non- Invasive Methods: A Review, *Journal of Physics: Conference Series* 1950, ICMIAI 2021, pp. 1 - 14, doi:10.1088/1742-6596/1950/1/012081
- [27]. M. B. Abubaker and B. Babayiğit, "Detection of Cardiovascular Diseases in ECG Images Using Machine Learning and Deep Learning Methods," in *IEEE Transactions on Artificial Intelligence*, vol. 4, no. 2, pp. 373-382, April 2023, doi: 10.1109/TAI.2022.3159505.
- [28]. T. Dhamdhare, S. Bhutambere, A. Bansode, and S. Umare, Heart Disease Detection Using ECG Waves: A Survey, *Journal of Emerging Technologies and Innovative Research (JETIR)*, Vol. 7, Issue 4, pp. 1761 - 1764
- [29]. A. S. . Sharma and D. H. . Hota, "ECG Analysis-Based Cardiac Disease Prediction Using Signal Feature Selection with Extraction Based on AI Techniques", *Int. j. commun. netw. inf. secur.*, vol. 14, no. 3, pp. 73–85, Dec. 2022.
- [30]. M.H. Vafaie, M. Ataei, and H.R. Koofgar, Heart diseases prediction based on ECG signals' classification using agenetic-fuzzy system and dynamical model of ECG signals, *Biomedical Signal Processing and Control* 14, 2014, pp. 291 -296