EMERGING ADVANCES IN INTEGRATED TECHNOLOGY
e-ISSN: 2773-5540

EmAIT

Vol. 5 No. 1 (2024) 52-58
https://publisher.uthm.edu.my/ojs/index.php/emait

Enhancing Heart Disease Classification: Performance Evaluation of the Machine Learning Model

Laveniya[1], Nor Hazlyna Harun[1,2]*,

[1] *School of Computing*
Universiti Utara Malaysia (UUM), Sintok, Kedah, MALAYSIA

[2] *Data Science Research Lab (DSRL), School of Computing*
Universiti Utara Malaysia (UUM), Sintok, Kedah, MALAYSIA

*Corresponding Author: hazlyna@uum.edu.my
DOI: https://doi.org/10.30880/emait.2024.05.01.007

Article Info

Received: 23 January 2024
Accepted: 13 June 2024
Available online: 30 June 2024

Keywords

Heart disease, machine learning, MLP, NN, filtering heart disease, classification

Abstract

Heart disease, mainly caused by coronary artery disease, poses a significant health challenge by impeding oxygen-rich blood flow to the heart muscle. Addressing this, our study utilizes a dataset of 1,026 cases from Kaggle, focusing on the application of MultiLayer Perceptron (MLP) in diagnosing heart conditions. This dataset, rich in attributes indicative of heart disease, comprises 14 variables, with the class variable determined by 13 attributes acting as the focal point for disease classification. Our methodology involved applying various machine learning algorithms, with a particular emphasis on the MLP model, to categorize cases as diseased or non-diseased. The MLP model demonstrated superior performance, achieving an accuracy of 92.39% (AUC=0.923), and outperforming several comparative techniques in terms of precision, recall, and mean squared error (MSE). The analysis of the dataset under this model revealed insightful patterns and trends, underscoring the effectiveness of MLP in heart disease diagnosis and potentially guiding future research in medical diagnostics.

1. Introduction

One month after diagnosis, the survival rate stands at 89.6 percent. After one year, the survival rate decreases to 78 percent, and at the five-year mark, it further diminishes to 57.7 percent. Considering the poor prognosis of heart failure, the disease must be diagnosed and treated properly. In 2018, the Australian Institute of Health and Welfare (AIHW) stated that cardiovascular disease (CVD) was the top cause of death in Australia, responsible for 42% of all deaths [1]. Consequently, complications such as pulmonary failure and peripheral edema may manifest [2]. The diagnosis of heart failure does not hinge on a solitary diagnostic test; rather, it necessitates a comprehensive approach involving the patient's medical history, a thorough physical examination, and various laboratory tests. Diastolic dysfunction and systolic dysfunction can produce different heart failure symptoms [3].

The field of clinical data has advanced in the organization, design, storage, and transmission of medical information for various purposes. One of these objectives is to enhance decision-making processes in healthcare services, thereby enhancing human capacity for pathological condition analysis, treatment, and evaluation. Heart disease is increasing in prevalence in modern society, leading to millions of global fatalities annually [2].

Manual techniques for tasks related to heart disease and classification pose inherent challenges. They heavily rely on human judgment, introducing subjectivity that may result in varying interpretations of the same set of symptoms or diagnostic results among healthcare professionals. This inconsistency can lead to diverse diagnoses, impacting the overall reliability of manual techniques. Moreover, these methods are time-consuming and often lack scalability, making them inefficient for handling a large number of cases.

To address these issues, previous works have proposed solutions by exploring evaluation techniques and methodologies in heart disease classification research. In this context, machine learning models have emerged as promising tools to assist in the diagnostic process. These models can analyze extensive patient data, identify patterns, and offer valuable insights to healthcare professionals, contributing to more accurate and timely diagnoses.

This study presents an approach to heart disease classification using a Multilayer Perceptron (MLP) model. Our research focuses on leveraging machine learning techniques and methodologies to enhance diagnostic accuracy. The core of our investigation involves the MLP in the evaluation and training processes. The primary objective of this paper is to propose a robust solution for improving the classification of heart diseases.

2. Related Works

The MLP model's classification of heart failure using an open-access dataset titled "heart failure clinical records," aims to identify potential risk factors for heart failure. The MLP model exhibits notably reliable outcomes. Additionally, this model has acquired significant data for recognizing risk factors that might be related to cardiovascular breakdown [3]. Consequently, it has been perceived that the important model will provide solid data about any sickness to be utilized in preventive medication rehearsals [3]. However, the research problem of preventing the model from overfitting the training data and ensuring that it generalizes well to unseen data, such as new patient cases, is a major concern. The current research considers using ensemble methods, such as bagging or boosting, to combine multiple MLP models. This can help improve generalization by reducing the variance.

A heart disease diagnosis system classifies two cases as normal and abnormal and five specific cases with high probability. The system uses two types of databases, one from the UCI learning repository and another gathered from the Ibn Al-Bitar Emergency Clinic and Baghdad Clinical City. These information bases incorporate thirteen clinical elements pivotal for heart disease determination. Two classifiers, the MLP and SVM, are proposed for heart disease classification. Results show that the MLP classifier accomplishes 98% exactness for two heart disease cases, while the SVM classifier reaches 96% accuracy. In classifying four types of heart diseases and normal cases, the MLP outperforms the SVM with an accuracy of 81%.

Traditional classifiers like MLP, PCA, Jordan, GFF, Modular, RBF, SOFM, SVM NNs, DA, and CART for heart disease classification. The MLP neural network, with optimized parameters, achieves an average classification of 98% and a best classification of 100% when trained and tested on cross-validation datasets. Compared to other neural networks and statistical models, MLP consistently demonstrates higher overall accuracy (96.67%), sensitivity (96%), and specificity (100%). The results highlight MLP's potential for heart disease classification, and its structural simplicity, requiring fewer free parameters, suggests feasibility for online and hardware implementation. Refer to Table 1 for a summary of the related works.

Table 1 *Literature review matrix* [4–8]

No	Aim	Datasets	Model/ Techniques	Results
1.	The MLP model in the classification of heart failure and aims to identify potential risk factors	Open-access "heart failure clinical records"	MLP ANN model	MLP model (AUC=0.925, Accuracy = 93.9%)
2.	A diagnostic system for heart disease has been created, aiming to categorize cases into two groups: Normal and Abnormal	UCI learning data set repository - Ibn Al Bitar Hospital Cardiac Surgery	MLP and SVM	MLP classifies has 98% accuracy then the SVM classifies has 96% accuracy
3.	To investigate and diagnose heart disease using biomedical test values, a feature selection algorithm, and an MLP with Backpropagation	UCI learning data set repository- heart disease	MLP with Back-Propagation	The accuracy dropped by 1.1% in the training dataset when reducing from 13 features to 8 features
4.	To develop a multiplayer perceptron-based decision support system for heart disease diagnosis	352 medical records of heart disease	MLP-based decision support system	MLP-based decision support systems are accurate, exceeding 90%, and have small intervals, less than 5%, showing they're great for helping with decisions about heart diseases

5.	The research uses a three-layer multilayer perceptron (MLP) neural network to create a decision support system for diagnosing five major heart diseases	38 input variables, extracted from a large number of patient cases	MLP with three layers	MLP can achieve high accuracy levels ranging from 63.6% to 82.9% in classifying heart diseases
6.	To determine MLP with a deep learning approach for data pre-processing helps enhance the quality of data	Framingham Heart Study Dataset	MLP with deep learning approach for data preprocessin g	MLP achieves a high accuracy of 96.50%
7.	To implement multilayer perceptron neural networks and support vector machines are compared on heart disease datasets	Dataset healthcare industry daily	MLP and SVM	SVM can classify more accurately
8.	The multilayer perceptron (MLP) and radial basis function (RBF) neural networks were applied to distinguish among patients (n = 266) experiencing one of these diseases	Use d 42 clinic data	MLP and RBF	The MLP achieved a sensitivity of 83.9% and, a specificity of 86%, and an area under the receiver operating characteristic curve (AUC) and RBF network yielded a sensitivity of 81.8%, a specificity of 88.4%
9.	the study aims to pinpoint potential risk factors associated with heart failure through the use of the MLP ANN model	open-access – heart failure clinical records	MLP ANN model	The MLP ANN model shows strong performance in classifying renal pelvic nephritis with an AUC of 0.925, Accuracy of 93.9%, Balanced Accuracy of 89.2%, Sensitivity of 98.4%, and Specificity of 80.0%
10.	To propose are committee machines (CM) using ensembles of MLP to distinguish between diagnoses of five major heart diseases	352 heart disease database	MLP Model	The committee machines, featuring ensembles of MLP, perform better than the single MLP on both counts
11.	To propose a method for identifying the most important features of cardiac disease characteristics through the application of a feature selection technique	Framingham heart disease dataset (FHS). KAGGLE Machine Learning repository.	MLP Model	The highest accuracy for this dataset was achieved CBFS with MLP
12.	The processed data undergoes testing with four classifiers to gauge the accuracy of the classification	Cleveland Heart Disease data	Deep Neural Net MLP	Involves tasks like Classification in Data Mining and Machine Learning
13.	A committee of classifiers, combining feed MLP and RBF is employed to diagnose heart disease	Cleveland heart disease dataset	MLP model	Achieves an impressive classification accuracy of 95.45%
14.	This paper aims to create a model for classifying medical data, specifically focusing on using Artificial Neural Networks (ANN) for diagnosing heart disease	Relevant dataset	MLP ANN model	The use of Artificial Neural Networks (ANN) is highly accurate

3. Proposed MLP Approach for Heart Disease Classification

In this segment, introduce the method for classifying heart disease using MLP [9]. We detail the steps involved in data pre-processing, describe the architecture of the MLP, explain the testing and training procedures, discuss the evaluation metrics employed, and compare our approach with other models. Emphasizing the robustness of the proposed method, we underscore its true capacity for reasonable applications in heart disease classification. Summarily, this section provides an overview of the MLP model, its exhibition, and its importance in accomplishing exact heart disease arrangements [10].

3.1 Data Preprocessing

The dataset utilized in this study comprises 1,026 cases with 14 attributes and 1 target attribute, obtained from Kaggle. The objective is to classify the target variable into "disease" or "non-disease" categories employing various machine learning algorithms, with a specific focus on the MLP model. In data preprocessing, the Impute widget was utilized to eliminate instances with unknown values. Subsequently, the data preprocessing was completed, and the results can be viewed on the Data Table widget. The outcome reveals 653 instances with 14 features and 1 target attribute.

The objective of the feature selection process was to identify the most effective characteristics for classifying heart conditions. Through assessment, it was seen that utilizing each of the 14 elements reliably accomplished elite execution across assessment measurements, showing the worth of all highlights for the order task. Thus, holding each of the 14 highlights for the ensuing analysis was chosen.

3.2 Model Building

Fig. 1 and 2 (a) show a sample of the model's precision across over 1000 iterations. It gives a visual depiction of how the precision metric is created through the readiness cycle. The figure highlights whether the model's precision improves, levels, or sways for a long time. Fig. 1 and 2 (b) display an illustration of the model's Mean Squared Error (MSE) across more than 1000 iterations. It illustrates variations in the model's MSE metric throughout the training cycle. The figure aids in evaluating the alignment between the model's predictions and real-world observations., with lower MSE values indicating better performance. There are three hidden layers with decreasing numbers of nodes (64, 32, 16). Each node in the hidden layers applies the Rectified Linear Unit (ReLU).

3.3 Training and Testing Procedure

The dataset has been partitioned into training (70%) and testing (30%) sets, comprising It contains 715 training instances and 307 testing instances. In the four graphs, the results of the training and testing phases are shown. Fig. 1 (a) and (b) show the training accuracy over epochs and the training mean square error (MSE) over epochs. Fig. 2 (a) and (b) show the testing accuracy over epochs and the testing mean square error (MSE) over epochs.

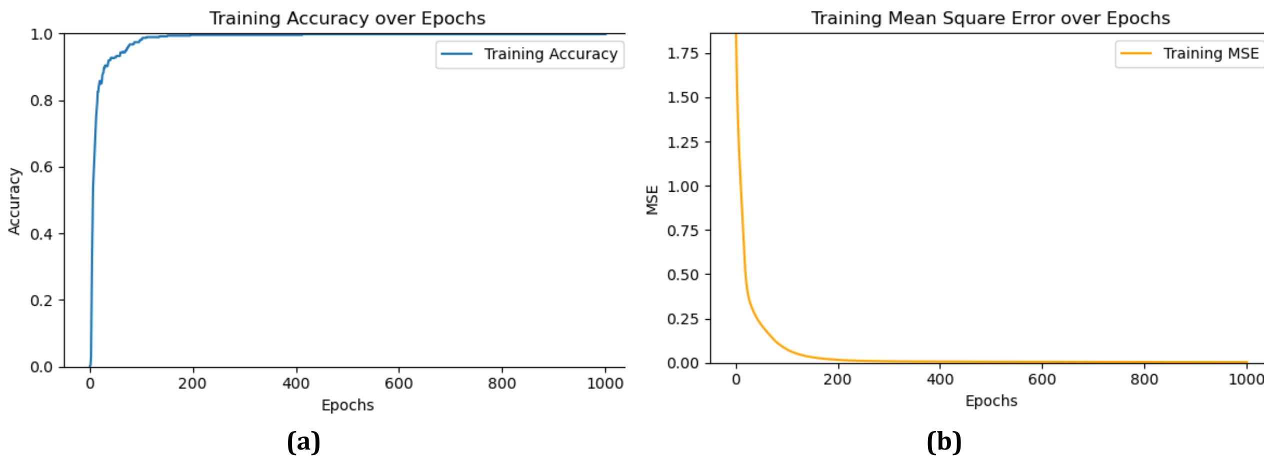

Fig. 1 *(a)Training accuracy over Epochs; (b) Training Mean Square Error over Epochs*

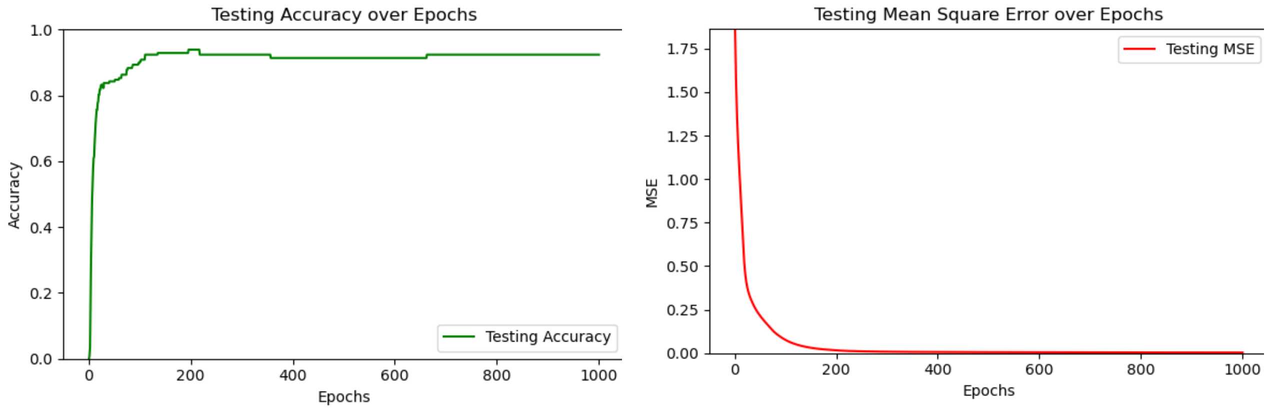

Fig. 2 *(a) Testing accuracy over Epochs; (b) Testing Mean Square Error over Epochs*

3.4 Evaluation Metrics

The table presented delineates the differences in performance metrics evaluation among various machine learning algorithm models applied to a given dataset. An evaluation of a classification model's performance requires a confusion matrix. In it, the predictions of the model are detailed by comparing them to the actual outcomes in tabular form. Table 2 and 3 show the training and testing set confusion matrix:

Table 2 *Confusion matrix for training set*

	Predicted	
	Negative (hearth disease)	Positive (legitimate)
Actual negative (hearth disease)	210	0
Actual positive (legitimate)	0	246

Table 3 *Confusion matrix for testing set*

	Predicted	
	Negative (hearth disease)	Positive (legitimate)
Actual negative (hearth disease)	83	13
Actual positive (legitimate)	2	99

As evident from the provided tables, both the training and testing sets demonstrate high accuracy in classifying instances of legitimate links. Moreover, the accuracy in predicting heart disease links, while still considerable, is slightly lower compared to the accuracy achieved in classifying legitimate links.

3.5 Comparison With Other Models

The evaluation of the MLP model's ability to characterize heart disease. This more extensive structure for assessment improves our cognizance of its assets and shortcomings. By providing insights into prediction confidence, the standard metric MSE complements conventional measurements such as precision, accuracy, recall, and the F1-Score. Table 4 below shows the performance metrics of the MLP and other models.

Table 4 outlines the evaluation metrics of different machine learning models in the dataset. The Multilayer Perceptron (MLP) model demonstrates impressive training performance with high accuracy (99.56%), F1 score (99.35%), precision (99.13%), and recall (99.56%). However, on the testing dataset, its accuracy drops to 91.37%, suggesting a potential overfitting issue.

The k-Nearest Neighbors model exhibits strong overall performance with an accuracy of 98.3%, while the decision tree and SVM models also exhibit high accuracies of 97.8% and 99.0%, respectively. The Random Forest model achieves perfect scores across all metrics, indicating potential overfitting as well.

On the other hand, the Naïve Bayes model performs moderately with an accuracy of 92.3%, an F1 score of 85.0%, precision of 85.0%, and recall of 85.0%. While not achieving the same level of accuracy as some other models, Naïve Bayes provides a balance between precision and recall.

Table 4 *Evaluating performance metrics for the MLP and other models*

Dataset	Performance			
	Accuracy	F1	Precision	Recall
MLP (Training)	0.99	0.99	0.99	0.99
MLP (Testing)	0.91	0.91	0.92	0.91
kNN	0.98	0.92	0.92	0.92
SVM	0.98	0.95	0.95	0.95
Tree	0.99	0.98	0.98	0.98
Random Forest	1.00	1.00	1.00	1.00
Naïve bayes	0.92	0.85	0.85	0.85

In summary, these performance metrics provide valuable insights into the strengths and weaknesses of each model. It's crucial to consider factors such as overfitting and the specific requirements of the problem at hand when selecting the most suitable model for deployment in real-world scenarios. Additionally, further analysis, such as cross-validation, can enhance the robustness of these findings.

4. Conclusion

Utilizing a database of 1026 instances and 14 attributes, the MLP method for heart disease characterization exhibits great execution, really recognizing people with and without heart disease. The MLP model outperforms different calculations, achieving a noteworthy 92.39% precision on the testing set. Using this method, people can lower their risk of heart disease and their likelihood of dying from cardiovascular problems.

In any case, the limits of this work rotate around time and cost imperatives. The undertaking included different emphases of the 'Attempt and Blunder' technique. Albeit the MLP-based technique requests significant preparation information and computational power, it stays a vigorous methodology for heart disease characterization. Future exploration ought to focus on extending the preparation information and upgrading assets to improve execution and versatility, in this way laying out viable symptomatic frameworks and protecting people from heart-related issues.

Acknowledgement

We express our sincere appreciation to all individuals who played a crucial role in the development of this research project on heart disease. Our mentors, coworkers, friends, and the academic community have played a crucial role in shaping our work with their support, guidance, and valuable insights. Additionally, we extend heartfelt thanks to our families and loved ones for their unwavering support throughout this project. Their encouragement has been a source of strength, motivating us to contribute significantly to the understanding and prevention of heart disease.

Conflict of Interest

Authors declare that there is no conflict of interests regarding the publication of the paper.

Author Contribution

The authors confirm contribution to the paper as follows: ***study conception and design:*** *Laveniya, Nor Hazlyna Harun;* ***data collection:*** *Laveniya;* ***analysis and interpretation of results:*** *Laveniya, Nor Hazlyna Harun;* ***draft manuscript preparation:*** *Laveniya, Nor Hazlyna Harun. All authors reviewed the results and approved the final version of the manuscript.*

References

[1] Kuruvilla, A. M., & Balaji, N. . (2021). Heart disease prediction system using Correlation Based Feature Selection with Multilayer Perceptron approach. *IOP Conference Series: Materials Science and Engineering*, *1085*(1), 012028. https://doi.org/10.1088/1757-899x/1085/1/012028

[2] Kaya, M. O. (2021). Performance Evaluation of Multilayer Perceptron Artificial Neural Network Model in the Classification of Heart Failure. *The Journal of Cognitive Systems*, *6*(1), 35–38. https://doi.org/10.52876/jcs.913671

[3] Zheng, J., Jiang, Y., & Yan, H. (2006). Committee machines with ensembles of multilayer perceptron for the support of diagnosis of heart diseases. *2006 International Conference on Communications, Circuits and*

Systems, ICCCAS, Proceedings, *3*, 2046–2050. https://doi.org/10.1109/ICCCAS.2006.285080

[4] Krishna, C. L., & Reddy, P. V. S. (2019). An Efficient Deep Neural Network Multilayer Perceptron Based Classifier in Healthcare System. *2019 Proceedings of the 3rd International Conference on Computing and Communications Technologies, ICCCT 2019*, 1–6. https://doi.org/10.1109/ICCCT2.2019.8824913

[5] Vadicherla, D., & Sonawane, S. (2013). Decision support system for heart disease based on sequential minimal optimization in support vector machine. *International Journal of Engineering Sciences & Emerging Technologies*, *4*(2), 19–26.

[6] Yan, H., Zheng, J., Jiang, Y., Peng, C., & Li, Q. (2003). Development of a decision support system for heart disease diagnosis using multilayer perceptron. *Proceedings - IEEE International Symposium on Circuits and Systems*, *5*. https://doi.org/10.1109/iscas.2003.1206411

[7] Karaduzovic-Hadziabdic, K., & Köker, R. (2015). Diagnosis of heart disease using a committee machine neural network. *Proceedings of the 9th International Conference on Applied Informatics*, *May*, 351–360. https://doi.org/10.14794/icai.9.2014.1.351

[8] Masih, N., Naz, H., & Ahuja, S. (2021). Multilayer perceptron based deep neural network for early detection of coronary heart disease. *Health and Technology*, *11*(1), 127–138. https://doi.org/10.1007/s12553-020-00509-3

[9] Naraei, P., Abhari, A., & Sadeghian, A. (2017). Application of multilayer perceptron neural networks and support vector machines in classification of healthcare data. *FTC 2016 - Proceedings of Future Technologies Conference*, 848–852. https://doi.org/10.1109/FTC.2016.7821702

[10] Tarle, B., & Jena, S. (2017, July 2). An Artificial Neural Network Based Pattern Classification Algorithm for Diagnosis of Heart Disease. *2017 International Conference on Computing, Communication, Control and Automation, ICCUBEA 2017*. https://doi.org/10.1109/ICCUBEA.2017.8463729