

# Heart Disease Prediction System using hybrid model of Multi-layer perception and XGBoost algorithms

Israa Nadheer<sup>1\*</sup>

<sup>1</sup> Alnahrain university, Baghdad ,Iraq

**Abstract.** Multi-layer perceptron (MLP) algorithms play a critical role in improving the accuracy and effectiveness of heart disease diagnosis in the context of the machine learning research. This paper presents an approach of heart disease prediction involves RReliefF-based feature importance assessment then MLP-based classification of features into three groups based on importance scores is proposed. The study employs three feedforward neural networks to classify effectively the clustered groups. Furthermore, an integrated approach utilizes XGBoost ensemble classification, leveraging boosted ensemble learning to enhance overall classification of the outputs of FNN models. By partitioning Cleveland dataset into 70% training and 30% testing sets creates independent datasets, the incorporation of MLP outputs into the XGBoost model yields satisfied testing performance. The confusion matrix showcases accurate classifications, with 96.67% accuracy, 95.92% sensitivity, and 97.92% precision. The F1-Score, at 96.91%, validates the model's balanced performance in precision and recall. This study exemplifies the efficacy of integrating data processing, feature engineering, and ensemble learning techniques for robust cardiovascular disease prediction, providing a reliable and efficient methodology for healthcare applications.

## 1 Introduction

The data related to patients, diseases, and diagnoses is increased in medical field. However, this data is not being leveraged effectively to yield the expected outcomes. Heart disease and stroke are among the leading causes of death [1]. A report by the World Health Organization states that cardiovascular diseases result in over 17.8 million deaths annually. The lack of adequate analysis means that the vast data on patients, diseases, and diagnoses in the healthcare industry does not have the desired impact on patient health [2]. Cardiovascular diseases (CVDs) include conditions such as coronary artery disease, myocarditis, and vascular disease. Stroke and heart disease illness represent 80% of all CVD-related deaths, with 3/4 of these deaths happening in people younger than 70. Risk factors for cardiovascular infection incorporate orientation, smoking, age, family ancestry, horrible eating routine, lipid levels, actual idleness, hypertension, weight gain, and liquor utilization. Hereditary factors, for example, hypertension and diabetes can increase the risk of cardiovascular illness [3]. The state of heart disease diagnosis is in a state of disarray and there is a clear need for better big-data analysis in order to redesign the cardiovascular systems and improve the patient outcomes. The problem is that the data is often noisy, incomplete, and contains ambiguities that make it difficult to extract clear, precise, and well-supported conclusions. In the face of these difficulties, machine learning has emerged as a new era in medical diagnostics especially through the use of machine-learning algorithms. These algorithms have great potential to efficiently manage specific medical centers and to analyze complex datasets. However, predicting heart disease, which is influenced by many factors like age, cholesterol level, lifestyle, etc., is a formidable challenge due to the sheer number of features involved. Classification within machine learning is hindered by the complexity of the data, which affects the performance and reduces the accuracy [4,5]. In this situation, multi-Perceptron algorithms come as a game changer. The use of machine learning systems to diagnose heart disease has been significantly enhanced by their capacity to handle intricate data landscapes. Multi-perceptron algorithms play a critical role in improving the accuracy and effectiveness of heart disease diagnosis in the context of the broader machine learning community [6]. Pre-processing techniques are used to interpret the data collected with various ML models.

\*Corresponding author: [asraa.nadheer88@nahrainuniv.edu.iq](mailto:asraa.nadheer88@nahrainuniv.edu.iq)

A standard set of algorithms, and their variations, are used to detect the emergence of the first stage of heart failure in both hereditary and healthy controls. Algorithms like multi-layer perceptrons (MLP), decision trees (DT), support vector machines (SVM), logistic regression (LR), and random forests (RF) are used to predict heart attack [7,8].

In this paper, we make several important contributions to this field:

- Introducing and applying RRelieFF based feature importance assessment to overcome the limitations of traditional methods
- Proposing a novel classification of features into 3 groups based on importance scores
- Utilizing feedforward neural networks to effectively classify within each group
- Utilizing integrated XGBoost (eXtreme Gradient Boosting) ensemble classification to enhance overall classification performance using boosted ensemble learning

## 2 Literature review

The research in question primarily focused on classifying heart disease using Cleveland dataset of 13 features [9]. The goal was to use these features to predict whether a patient had heart disease. These studies used multiple machine learning techniques, each with different results. Decision trees are a popular method for classifying tasks by their interpretability and achieved an accuracy of 89.1% [10]. Random forests, an ensemble learning technique that works by building multiple decision trees, easily outperformed decision trees with an accuracy of 89.2% [11]. Artificial neural networks (ANNs), inspired by the biological neural networks that make up animal brains, are also being used. The ANN showed an accuracy of 92.7% [12]. This variation may be due to differences in network architecture, training methods, or data preprocessing. The most notable hybrid model is the combination of genetic algorithm and neural network (GA + NN) [13], which achieved an impressive accuracy of 94.2%. This suggests that a hybrid model that combines the strengths of several individual models may improve performance. The best models using principal component analysis (PCA), a technique that emphasizes variation and emphasizes strong patterns within a dataset, were PCA-regression and PCA1-NN [14]. These models achieved 92.0% and 95.2% accuracy, respectively, demonstrating the potential benefits of dimensionality reduction in improving model performance.

Chaddha and Mayank [15] obtained a remarkable 100% accuracy with a neural network model that had just 8 features and three layers of neurons. However, obtaining 100% accuracy may signal overfitting, as it is uncommon for a model to precisely capture all of the intricacies in a dataset without memorizing the information. Shah Devansh et al. [16] attained the maximum accuracy with K-NN (90.789%). However, the performance of K-NN can be sensitive to the choice of the 'K' parameter and the metric used for calculating distance. Kumar Dwivedi [17] used multiple algorithms on the Statlog Heart Disease dataset, with Logistic Regression achieving the highest accuracy of 85%. It would be interesting to know if any parameter tuning was done for these models, as that could potentially improve their performance. Deepika K and Seema S [18] reported high accuracies for multiple models, with SVM achieving the highest at 95.2%. However, SVMs can be computationally expensive for large datasets. Parthiban G and Srivatsa SK [19] achieved a 74% accuracy predicting heart disease in diabetic patients using Naive Bayes. This relatively lower accuracy might indicate that the model could benefit from additional feature engineering or a different algorithm. Vembandasamy et al. [20] achieved an 86.4% accuracy with Naive Bayes. However, Naive Bayes assumes feature independence, which might not hold true in medical datasets. Otoom et al. [21] achieved the same accuracy rate with Naive Bayes and SVM, both at 83.8%. However, they didn't mention if any hyperparameter tuning was done, which could potentially improve the models' performance. In general, while these studies have achieved promising results, it's important to remember that model performance can vary based on the dataset used, the preprocessing steps taken, and the specific parameters used for each algorithm.

One methodology introduced a symptomatic framework that utilized a profound brain organization (DNN) for grouping and a X2 factual model for highlight refinement, accomplishing 93.3% exactness [22]. In any case, the ideal width of one secret layer in the DNN and ANN model was researched utilizing a framework search calculation, recommending the chance of additional complicated calculations like the hereditary calculation. One more methodology assembled two different order models in light of a versatile neuro-fuzzy surmising framework (ANFIS). Exactness was 75% with preparing information and 76.6% with test information [23]. This alludes to the precision of the model and the estimation of the Jacobian network to further develop the union speed of the outcome. Notwithstanding, the precision actually needs improvement. The new technique offers a PSO (Molecule Multitude Streamlining) based highlight determination with a refreshed Fluffy ANN classifier for CVD risk expectation. The technique was 88.82% exact in foreseeing men with the sickness and 88.05% precise in ladies [24]. This proposes that precision should be improved to further develop prediction speed.

## 3 Materials and Methods

### 3.1. Dataset

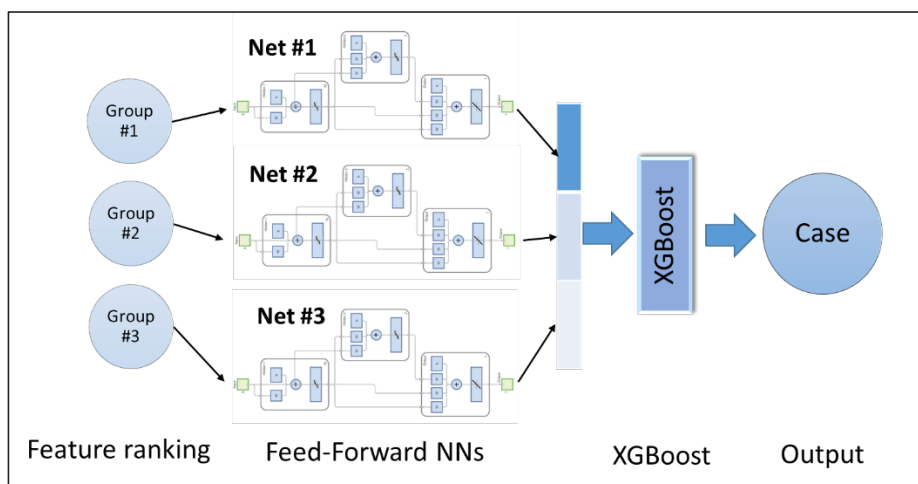
The dataset utilized in this scientific investigation is derived from the Cleveland Heart Disease dataset [9], serving as an imbalanced classification dataset comprising a total of 303 instances. This dataset encompasses 13 distinct features and one target variable, each contributing to the comprehensive characterization of the individuals under scrutiny. The features are elucidated as follows:

- Age 'age': Denoting the age of the individual.
- Sex 'sex': Indicating the gender of the individual, with '0' corresponding to female and '1' to male.
- Chest-pain type 'cp': Representing the type of chest pain experienced, with categories including typical angina (1), atypical angina (2), non-anginal pain (3), and asymptomatic (4).
- Resting Blood Pressure 'restbps': Expressing the individual's resting blood pressure measured in mmHg.
- Serum Cholesterol 'chol': Conveying the serum cholesterol level in mg/dL.
- Fasting Blood Sugar 'fbs': Comparing an individual's fasting blood sugar value with a threshold of 120 mg/dL, denoted as '1' for true (if fasting blood sugar >120 mg/dL) and '0' for false.
- Resting ECG 'restecg': Presenting resting electrocardiographic results, categorized as normal (0), having ST-T wave abnormality (1), or left ventricular hypertrophy (2).
- Max Heart Rate Achieved 'thalach': Reflecting the maximum heart rate attained by the individual.
- Exercise-induced Angina 'exang': Indicating the presence (1) or absence (0) of exercise-induced angina.
- ST Depression Induced by Exercise Relative to Rest 'oldpeak': Depicting the value, which may be an integer or float.
- Peak Exercise ST Segment 'slope': Describing the peak exercise ST segment as upsloping (1), flat (2), or downsloping (3).
- Number of Major Vessels Colored by Fluoroscopy 'ca': Representing the count of major vessels (ranging from 0 to 3) colored by fluoroscopy, expressed as an integer or float.
- Thal 'thal': Indicating thalassemia, with categories including normal (3), fixed defect (6), and reversible defect (7).
- Diagnosis of Heart Disease: Disclosing whether the individual is afflicted by heart disease, categorized as absence (0) or presence (1).

The following steps are necessary for the data preparation process and guarantee the accuracy of our dataset: addressing missing values. Doing so helps to prevent skewed analysis results. MATLAB's `fillmissing` function was utilized to address missing data using the 'linear' interpolation method. This technique involves estimating missing data based on a linear trend in the dataset. With the goal of using linear interpolation to replace null values with interpolated values, we aim to create a more continuous and precise representation of the data. As opposed to traditional methods like row dropping or mean substitution, this technique offers a more flexible solution. The use of linear interpolation ensures that the dataset retains its structural integrity, as it effectively resolves missing values without disrupting the overall trend of the variables. This approach is particularly useful when dealing with a small number of missing values, as it adds little variation to the dataset while effectively addressing data loss issues.

### 3.2. Proposed method

The proposed method (Figure 1) involves a structured data processing approach to extract valuable insights from a cardiovascular dataset. The dataset undergoes thorough preparation, including loading and filling in missing values using linear interpolation. It is then divided into training and testing sets, with 70% allocated for training and 30% for testing. The training data is divided into three subsets—`data1`, `data2`, and `data3`—each containing unique attributes based on their relevance ranking. Three Multi-Layer Perceptron (MLP) neural networks—`net1`, `net2`, and `net3`—are then independently trained on their respective feature subsets using the Levenberg-Marquardt backpropagation technique. The outputs of the trained MLP networks are then merged into a consolidated feature set, showcasing the methodology's algorithmic context. This merged feature set is then used to train an XGBoost model via an ensemble learning strategy. The performance of the XGBoost model is evaluated using testing data, and the results are compared to the real labels. A confusion matrix is generated, revealing the model's performance in terms of true positive, true negative, false positive, and false negative predictions. Lastly, the model's accuracy is determined as the ratio of correctly predicted occurrences to the total instances.



**Figure 1. The proposed methodology.**

### 3.3. Feature ranking

The ReliefF algorithm is a feature selection technique that evaluates and ranks the significance of features in a dataset. It provides insight into each feature's predictive value for classification tasks. In this section, we discuss the concept of Rank Importance of Predictors using ReliefF, as well as the results obtained from applying this method to a dataset with appropriate feature labels. The ReliefF algorithm assesses the value of features by examining their ability to distinguish between instances of the same class and instances of other classes. It calculates feature significance using the difference in feature values between the closest cases with the same and different class labels. Features with larger differences are considered more relevant for distinguishing between examples of various classes. RReliefF (Randomized ReliefF) is an extension of ReliefF that uses randomization to improve efficiency and scalability. It calculates feature significance on a random selection of examples, reducing computational cost compared to evaluating all instances in the dataset exhaustively. By implementing ReliefF on Cleveland features (Figure 2), the features are grouped as follows:

*3.1.1 Top-ranked Features (Group #1): The features at the top of the ranked list (e.g., 'age' and 'sex') are deemed more essential by the ReliefF or RReliefF algorithm. These traits make a major contribution to the classification process and are regarded as critical for distinguishing between various occurrences. The characteristics are: 'age', 'cp', 'chol' and 'fbs'.*

*3.1.2 Middle-ranked Features (Group 2): qualities in the center of the list are also important, but to a lesser extent than the top-ranked qualities. Where the characteristics are 'trestbps', 'restecg', 'thal' and 'slope'.*

*3.1.3 Bottom-ranked Features (Group #3): Features toward the bottom of the list are deemed less essential in the context of the classification job. The characteristics include 'oldpeak', 'thalach', 'sex', 'ca' and 'exang'.*

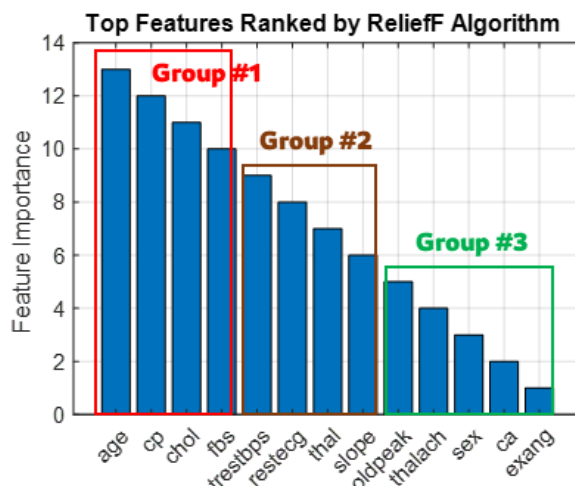


Fig. 2. The result of implementing ReliefF method and grouping the features into 3 sets.

### 3.4. Classifiers

The neural network models use the cascade forward net architecture and the Levenberg-Marquardt (trainlm) backpropagation training technique. Each model has two hidden layers, with ten neurons in the first layer and two neurons in the second. These models are set up for feedforward neural network tasks like regression or classification, and the cascade-forward-net design enables adaptation to capture complicated correlations in the data. The maximum number of epochs, or loops over the dataset during training, is critical. Setting a sufficient number of 1000 epochs ensures that the models have enough iterations to reach convergence while avoiding overfitting. The performance target, which in this example is  $1e-6$ , decides when the training should be interrupted according to the desired level of accuracy. The learning rate, which is the key parameter for weight updates, is recommended to be 0.01 to balance convergence speed and stability. Regulation is applied with a recommended value of 0.01 to avoid over-tuning. These parameters are then applied to each neural network model (net1, net2, and net3), resulting in consistent and successful cascade architecture training using the Levenberg-Marquardt backpropagation training process. Depending on the unique characteristics of the dataset and the intended performance of the model, fine-tuning may be necessary.

XGBoost (eXtreme Gradient Boosting) is a powerful and efficient machine learning technique used for supervised learning tasks including classification and regression. It falls under ensemble learning approaches, especially the gradient descent framework, and has become popular due to its good predictive performance. XGBoost uses ensembles of weak learners (decision trees) to build predictive models. This approach optimizes an objective function by adding trees to correct existing group errors. Each tree is added with the goal of reducing the residual error of the previous model. The algorithm uses a fixed objective function that combines the mystery function of the problem at hand with a set time. Trees are pruned during training to avoid breakage, and the final prediction is a weighted sum of predictions from all trees. The parameters are:

These parameter settings represent a balanced starting setup for training an XGBoost model. The learning rate of 0.1 governs each tree's contribution, and a maximum depth of 3 prevents overfitting. Subsample and colsample\_bytree values of 0.8 introduce stochasticity, which helps with generalization and promotes model variety. A gamma value of 0.1 determines the minimal loss reduction for partitioning, resulting in regularization.

### 3.5. Performance metrics

We performed a comprehensive analysis of each model's performance using a set of standard classification metrics, including precision, accuracy, recall, F1 score, and area under the curve (AUC) for the receiver operating characteristic (ROC) curve. These metrics are based on the confusion matrix values: true positive (TP), false positive (FP), true negative (TN), false negative (FN). Performance measures are as follows:

- True Positives (TP): the count of positive instances correctly predicted by the model to have heart disease.
- True Negatives (TN): the count of negative instances correctly predicted by the model to be free of heart disease.
- False Positives (FP): the count of negative instances incorrectly predicted by the model to have heart disease.
- False Negatives (FN): the count of positive instances incorrectly predicted by the model to be free of heart disease.

Accuracy, defined as the ratio of correctly predicted samples to all samples, is given by:



$$\text{Accuracy} = (\text{TP} + \text{TN}) / (\text{FP} + \text{FN} + \text{TP} + \text{TN}) \quad (1)$$

Precision, the ratio of correctly predicted positive samples to the total number of correctly predicted positive samples, is calculated as follows:

$$\text{Precision} = \text{TP} / (\text{FP} + \text{TP}) \quad (2)$$

Recall, signifying the ratio of correctly predicted positive instances to all instances in the actual positive class, is given by:

$$\text{Recall} = \text{TP} / (\text{FN} + \text{TP}) \quad (3)$$

The F1 score, a balanced measure of precision and recall, is defined as the weighted average:

$$\text{F1 score} = (2 \times \text{Precision} \times \text{Recall}) / (\text{Precision} + \text{Recall}) \quad (4)$$

The AUC-ROC is constructed by plotting the true positive rate (TPR), which is the same as recovery, against the true positive rate (FPR). FPR is defined as:

$$\text{FPR} = \text{FP} / (\text{FP} + \text{TN}) \quad (5)$$

In addition to the aforementioned performance metrics, a 5-fold cross-validation approach is incorporated to further enhance the robustness of the performance analysis. Cross-validation involves partitioning the dataset into five subsets, or "folds," where each fold serves as a testing set while the remaining four folds collectively constitute the training set. This process is repeated five times, with each fold taking turns as the testing set.

## 4. Results and Discussion

The neural network (NN) models exhibit strong training performance across key metrics (Table 2). In NN Model #1, the accuracy is 96.55%, indicating a correct classification rate of instances. Precision and recall both stand at 96.94%, showcasing a well-rounded performance with a balanced F1-score of 96.94%. Model #2 demonstrates high accuracy at 96.43%, with precision notably higher at 98.92%, while recall is slightly lower at 93.88%. The F1-score reflects a good balance between precision and recall at 96.32%. Model #3 had the good accuracy of 89.73%, featuring precision at 88.88% and recall at 88.88% while F1-score was 90.49%, indicating a strong balance between precision and recall. Overall, these models show robust performance, and the choice between them may depend on specific application requirements, considering trade-offs between false positives and false negatives.

**Table 1.** Training performance of each neural network.

Metric	NN Model #1	NN Model #2	NN Model #3
Accuracy	96.55%	96.43%	89.73%
Precision	96.94%	98.92%	92.20%
Recall	96.94%	93.88%	88.88%
F1-Score	96.94%	96.32%	90.49%

In Group #1, which includes the highest-ranked characteristics, 'age,' 'sex,' 'fbs,' and 'cp' are highlighted as critical contributions to the classification task. According to the algorithm, these parameters play a major role in discriminating across occurrences, meaning that age, gender, fasting blood sugar, and chest pain kind are all considered extremely relevant in predicting heart disease. The middle-ranked features in Group #2 are 'trestbps,' 'chol,' 'ca,' and 'thalach.' Although these traits are not at the top of the ranking list, they are nonetheless important, although to a lower level than those in Group #1. 'trestbps' and 'chol' are blood pressure and cholesterol readings, respectively, while 'ca' and 'thalach' may give further insights into coronary artery problems and maximal heart rate. Lastly, in Group #3, the bottom-ranked features, namely 'exang,' 'restecg,' 'oldpeak,' 'thal,' and 'slope,' are considered less crucial for the classification task. These features, according to the algorithm, contribute less significantly to distinguishing between instances with or without heart disease.

The testing performance of incorporating the outputs of NN models as input for XGBoost demonstrates highly favorable results (Figure 3). The confusion matrix reveals that the majority of instances are correctly classified, with 47 true negatives, 40 true positives, 2 false positives, and 1 false negative. This translates to an accuracy of 96.67%, indicating the model's proficiency in making accurate predictions. Sensitivity, or recall, stands at 95.92%, emphasizing the model's capability to effectively identify positive instances. Precision is notably high at 97.92%, highlighting the low rate of false positives. The F1-Score, a balanced measure of precision and recall, is 96.91%, further validating the model's robust performance.

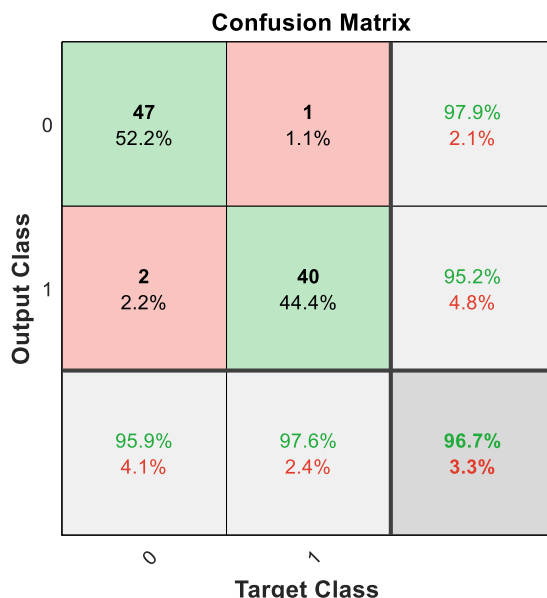


Fig. 3. Confusion matrix of XGBoost testing performance.

Comparing these results with related work (Table 2), Chaddha and Mayank [15] achieved a perfect 100% accuracy with a Neural Network model but cautioned that such high accuracy might indicate overfitting. Shah Devansh et al. [16] achieved a notable accuracy of 90.789% using K-NN; however, the sensitivity of K-NN to parameter choices and distance metrics raises concerns about its reliability. Kumar Dwivedi [17] reported an 85% accuracy with Logistic Regression, and Deepika K and Seema S [18] achieved 95.2% accuracy with SVM. In contrast, the proposed approach outperforms these models, highlighting the effectiveness of incorporating NN outputs into XGBoost. Parthiban G and Srivatsa SK [19] achieved a 74% accuracy with Naive Bayes, suggesting a potential need for additional feature engineering or a different algorithm. Vembandasamy et al. [20] achieved an 86.4% accuracy with Naive Bayes, and Ootom et al. [21] obtained an 83.8% accuracy with Naive Bayes and SVM. The suggested technique outperforms these results, implying better predictive performance. In addition, innovative techniques were introduced: a diagnostic system with a DNN obtaining 93.3% accuracy [22], two ANFIS-based classification models achieving 76.6% accuracy [23], and a PSO-based attribute selection with an upgraded Fuzzy ANN classifier achieving up to 88.82% accuracy [24]. These techniques provide exciting opportunities for additional investigation and eventual development. In conclusion, while several research have yielded promising results, the suggested technique of using NN model outputs for XGBoost integration displays competitive and better performance, implying that it is useful in improving predictive capacities for heart disease categorization. The comparison illustrates the significance of adopting many techniques and continually refining models to achieve peak performance.

Table 2. Comparison with related work.

Study	Model/Approach	Accuracy
Chaddha and Mayank [15]	Neural Network	100%
Shah Devansh et al. [16]	K-NN	90.789%
Kumar Dwivedi [17]	Logistic Regression	85%
Deepika K and Seema S [18]	SVM	95.2%
Parthiban G and Srivatsa SK [19]	Naive Bayes	74%
Vembandasamy et al. [20]	Naive Bayes	86.4%
Ootom et al. [21]	Naive Bayes and SVM	83.8%
Novel Approaches	DNN [22], ANFIS [23], PSO-based Fuzzy ANN [24]	93.3%, 76.6%, up to 88.82%
Current Study	XGBoost with Multi-perceptron outputs	96.67%

## 4 Conclusions and Future Directions

In conclusion, we show that incorporating MLP-NN models into XGBoost after grouping features is promising approach for improving predictive capabilities for heart disease classification and achieving accuracy, sensitivity, precision, and F1 scores. Although this approach is more promising than existing models, care should be taken when interpreting the results as they may be too large and require further validation. A limitation of this study is the lack of technology exploration or algorithm selection that affects the model and its robustness. Future work will include a deeper study of technical methods and the selection of algorithms for performance optimization. In addition, it explores the potential of new approaches, such as system analysis and deep neural networks, classification models based on adaptive neuro-fuzzy inference systems, and feature selection based on particle optimization, paving the way for detection and improved prognosis of heart diseases. Overall, this study provides important insights into the effectiveness of integrating outputs from NN and XGBoost models and also suggests areas for improvement and future exploration.

## References

1. Tsao, C. W., Aday, A. W., Almarzooq, Z. I., Alonso, A., Beaton, A. Z., Bittencourt, M. S., ... & American Heart Association Council on Epidemiology and Prevention Statistics Committee and Stroke Statistics Subcommittee. (2022). Heart disease and stroke statistics—2022 update: a report from the American Heart Association. *Circulation*, 145(8), e153-e639.
2. Chatzinikolaou, A., Tzikas, S., & Lavdaniti, M. (2021). Assessment of Quality of Life in Patients With Cardiovascular Disease Using the SF-36, MacNew, and EQ-5D-5L Questionnaires. *Cureus*, 13(9).
3. Srinivasan, S., Gunasekaran, S., Mathivanan, S. K., M. B, B. A. M., Jayagopal, P., & Dalu, G. T. (2023). An active learning machine technique based prediction of cardiovascular heart disease from UCI-repository database. *Scientific Reports*, 13(1), 13588.
4. Dai, H., Younis, A., Kong, J. D., Puce, L., Jabbour, G., Yuan, H., & Bragazzi, N. L. (2022). Big data in cardiology: state-of-art and future prospects. *Frontiers in Cardiovascular Medicine*, 9, 844296.
5. Huang, J. D., Wang, J., Ramsey, E., Leavey, G., Chico, T. J., & Condell, J. (2022). Applying artificial intelligence to wearable sensor data to diagnose and predict cardiovascular disease: a review. *Sensors*, 22(20), 8002.
6. Ogunpola, A., Saeed, F., Basurra, S., Albarrak, A. M., & Qasem, S. N. (2024). Machine Learning-Based Predictive Models for Detection of Cardiovascular Diseases. *Diagnostics*, 14(2), 144.
7. Sanyal, S., Das, D., Biswas, S. K., Chakraborty, M., & Purkayastha, B. (2022, May). Heart Disease Prediction Using Classification Models. In *2022 3rd International Conference for Emerging Technology (INCET)* (pp. 1-6).
8. Hossain, M. I., Maruf, M. H., Khan, M. A. R., Prity, F. S., Fatema, S., Ejaz, M. S., & Khan, M. A. S. (2023). Heart disease prediction using distinct artificial intelligence techniques: performance analysis and comparison. *Iran Journal of Computer Science*, 1-21.
9. UCI heart disease data set Retrieved from <http://archive.ics.uci.edu/ml/datasets/heart+disease> (2018, September 26), (2024, January 15).
10. Sen, S. K. (2017). Predicting and diagnosing of heart disease using machine learning algorithms. *International Journal of Engineering and Computer Science*, 6(6), 21623-21631.
11. Khan, S. (2017). Prediction of Angiographic Disease Status using Rule Based Data Mining Techniques Prediction of Angiographic Disease Status using Rule Based Data Mining Techniques. 8.
12. Das, R., Turkoglu, I., & Sengur, A. (2009). Effective diagnosis of heart disease through neural networks ensembles. *Expert systems with applications*, 36(4), 7675-7680.
13. Amma, N. B. (2012, February). Cardiovascular disease prediction system using genetic algorithm and neural network. In *2012 international conference on computing, communication and applications* (pp. 1-5).
14. Santhanam, T., & Ephzibah, E. P. (2013). Heart disease classification using PCA and feed forward neural networks. In *Mining Intelligence and Knowledge Exploration: First International Conference, MIKE 2013, Tamil Nadu, India, December 18-20, 2013*. Proceedings (pp. 90-99).
15. Chadha, R., & Mayank, S. (2016). Prediction of heart disease using data mining techniques. *CSI transactions on ICT*, 4, 193-198.



16. Shah, D., Patel, S., & Bharti, S. K. (2020). Heart disease prediction using machine learning techniques. *SN Computer Science*, 1, 1-6.
17. Dwivedi, A. K. (2018). Performance evaluation of different machine learning techniques for prediction of heart disease. *Neural Computing and Applications*, 29, 685-693.
18. Deepika, K., & Seema, S. (2016, July). Predictive analytics to prevent and control chronic diseases. In 2016 2nd international conference on applied and theoretical computing and communication technology (iCATecT) (pp. 381-386).
19. Parthiban, G., & Srivatsa, S. K. (2012). Applying machine learning methods in diagnosing heart disease for diabetic patients. *International Journal of Applied Information Systems*, 3(7), 25-30.
20. Vembandasamy, K., Sasipriya, R., & Deepa, E. (2015). Heart diseases detection using Naive Bayes algorithm. *International Journal of Innovative Science, Engineering & Technology*, 2(9), 441-444.
21. Otoom, A. F., Abdallah, E. E., Kilani, Y., Kefaye, A., & Ashour, M. (2015). Effective diagnosis and monitoring of heart disease. *International Journal of Software Engineering and Its Applications*, 9(1), 143-156.
22. Ali, L., Rahman, A., Khan, A., Zhou, M., Javeed, A., & Khan, J. A. (2019). An automated diagnostic system for heart disease prediction based on  $\chi^2$  statistical model and optimally configured deep neural network. *Ieee Access*, 7, 34938-34945.
23. Sagir, A. M., & Sathasivam, S. (2017). A Novel Adaptive Neuro Fuzzy Inference System Based Classification Model for Heart Disease Prediction. *Pertanika Journal of Science & Technology*, 25(1).
24. Narasimhan, B., & Malathi, A. (2019). Altered particle swarm optimization based attribute selection strategy with improved fuzzy Artificial Neural Network classifier for coronary artery heart disease risk prediction. *Int. J. Adv. Res. Ideas Innov. Technol*, 5, 1196-1203.