IJARCCE ISSN (O) 2278-1021, ISSN (P) 2319-5940

International Journal of Advanced Research in Computer and Communication Engineering

Impact Factor 8.471 ⁜ Peer-reviewed & Refereed journal ⁜ Vol. 15, Issue 7, July 2026

DOI: 10.17148/IJARCCE.2026.15721

# An Empirical Assessment of Machine Learning Algorithms for Predicting Heart Disease

**Dr. Anureet Kaur**

Asst. Professor, P.G. Department of Computer Science and Applications, Khalsa College, Amritsar, India

**Abstract**: Cardiovascular disorders remain a primary cause of mortality worldwide, emphasizing the urgent need for early diagnostic support systems. Machine learning (ML) techniques have emerged as promising tools for assisting clinical decision-making through data-driven prediction models. This study presents a comprehensive empirical evaluation of several supervised learning algorithms for heart disease prediction using structured clinical data. The performance of Logistic Regression, Support Vector Machine, k-Nearest Neighbours, Decision Tree, Random Forest, Gradient Boosting, and Artificial Neural Networks is systematically compared. Models are assessed using accuracy, precision, recall, F1-score, and ROC-AUC metrics. The results of the study show that the k-Nearest Neighbours algorithm achieved the best overall performance with the highest accuracy and balanced evaluation metrics. The Support Vector Machine model also showed strong and consistent performance across different validation folds. Other models such as Logistic Regression and Random Forest produced competitive results, while Decision Tree and Gradient Boosting showed comparatively lower performance. Overall, the results demonstrate that machine learning techniques can be effective tools for predicting heart disease and can help support early diagnosis. In the future, the performance of these models could be improved by using larger datasets and applying advanced feature selection or optimization techniques. These findings demonstrate that machine learning techniques can be useful for supporting early detection of heart disease and assisting healthcare professionals in decision-making.

**Keywords:** Cardiovascular Disease, Predictive Modelling, Healthcare Analytics, Artificial Intelligence, Machine Learning.

## I. INTRODUCTION

Cardiovascular diseases (CVDs) represent one of the most significant public health concerns globally, contributing substantially to mortality and long-term disability. Early identification of individuals at high risk of heart disease enables preventive interventions and reduces healthcare costs. Traditional diagnostic procedures rely on a physician's interpretation of clinical indicators such as cholesterol levels, blood pressure, electrocardiographic results, and patient history. However, the increasing availability of healthcare data has driven the adoption of machine learning techniques that extract hidden patterns and improve predictive accuracy. Although numerous studies have investigated ML-based cardiac risk prediction, there remains a need for structured comparative evaluation under consistent experimental settings. This research aims to systematically analyse and compare multiple machine learning models to determine their suitability for heart disease prediction tasks. Cardiovascular disease (CVD) continues to be a major global health concern, necessitating the development of accurate and reliable predictive models for early diagnosis and risk assessment. In recent years, machine learning techniques have gained significant attention for heart disease prediction due to their ability to analyse complex clinical datasets and uncover hidden patterns. Ensemble-based approaches such as Random Forest and Gradient Boosting have demonstrated improved classification performance compared to individual models [1], while comparative evaluations of traditional classifiers, including Support Vector Machines and Logistic Regression, have highlighted the importance of appropriate model selection and feature engineering [2], [3]. The integration of explainable artificial intelligence techniques, such as SHAP with XGBoost, has further enhanced interpretability in clinical decision-making systems [7]. Additionally, federated learning frameworks have emerged to address privacy concerns by enabling distributed training across healthcare institutions without sharing sensitive patient data [11]. Despite these advancements, challenges related to generalization, calibration, and deployment readiness remain, motivating continued empirical evaluation of machine learning techniques for robust heart disease prediction.

This research study first presents related studies published in the current scope in section II. The methodology adopted for conducting the empirical analysis is presented in section III. The results and interpretations derived from analysis are discussed in section IV. Lastly, section V presents conclusion and future directions.

## II. RELATED STUDIES

Over the past decade, machine learning has increasingly been integrated into healthcare analytics. The application of machine learning (ML) in cardiovascular disease prediction has advanced significantly in recent years due to improved

IJARCCE ISSN (O) 2278-1021, ISSN (P) 2319-5940

International Journal of Advanced Research in Computer and Communication Engineering

Impact Factor 8.471 ⁙ Peer-reviewed & Refereed journal ⁙ Vol. 15, Issue 7, July 2026

DOI: 10.17148/IJARCCE.2026.15721

data availability, enhanced computational power, and methodological innovations. Between 2021 and 2025, research has focused not only on improving predictive accuracy but also on interpretability, computational efficiency, and privacy-preserving frameworks. Recent years have witnessed substantial advancements in Machine Learning (ML) techniques for heart disease prediction, focusing on improving predictive accuracy, interpretability, and real-world deployment feasibility. Table I summarizes various studies on different attributes.

In 2021, Mohan *et al.* [1] proposed a hybrid ensemble learning framework combining Random Forest and Gradient Boosting classifiers to enhance prediction reliability using the UCI Heart Disease dataset. Their results demonstrated improved classification accuracy compared to standalone models, although statistical validation was not extensively explored. Similarly, Alizadeh *et al.* [2] conducted a *comparative* study of Support Vector Machines, Decision Trees, and Random Forest models on clinical cardiovascular datasets, emphasizing performance evaluation using ROC-AUC and precision-recall metrics.

In 2022, Dwivedi *et al.* [3] presented a structured comparative analysis of classical machine learning models, including Logistic Regression, k-Nearest Neighbor, Decision Tree, and Random Forest. Their findings indicated that ensemble approaches generally outperform single learners. Sharma and Jain [4] investigated Artificial Neural Networks (ANNs) with regularization techniques, demonstrating improved sensitivity in heart disease detection, though overfitting remained a concern.

Rahman *et al.* [5] introduced an AutoML-based framework in 2022, automating preprocessing and model selection for cardiovascular prediction tasks. Their approach reduced manual intervention while maintaining competitive accuracy. Garcia and Liu [6] proposed a meta-learning ensemble approach that combined multiple base learners to enhance model stability across heterogeneous datasets.

The year 2023 marked significant progress in explainable and deep learning-based systems. Khan *et al.* [7] integrated XGBoost with SHAP (SHapley Additive exPlanations) to improve interpretability in heart disease prediction models, enabling clinicians to understand feature importance. Ahmad and Zhang [8] evaluated boosting techniques such as AdaBoost and XGBoost for imbalanced cardiovascular datasets, reporting improvements in minority-class detection. Ibrahim *et al.* [9] introduced a transformer-based architecture to capture complex feature interactions in structured clinical data, achieving higher ROC-AUC values than traditional ML approaches. Oliveira *et al.* [10] combined convolutional neural networks with conventional machine learning methods to incorporate multimodal data such as ECG signals, thereby enhancing diagnostic performance.

In 2024, research increasingly focused on privacy, generalization, and model reliability. Li *et al.* [11] proposed a federated learning framework for distributed heart disease prediction across multiple hospitals, ensuring data privacy without centralizing sensitive patient data. Patel and Verma [12] examined energy-efficient ML deployment in healthcare edge systems, balancing predictive performance with computational cost. Nguyen *et al.* [13] employed transfer learning to improve model adaptability across demographic variations. Hassan *et al.* [14] emphasized model calibration and probability reliability using Brier scores, highlighting the importance of trustworthy predictions in clinical decision-making.

In 2025, emerging research explored advanced ensemble strategies and uncertainty modeling. Singh *et al.* [15] proposed a stacked ensemble model integrating Gradient Boosting and Artificial Neural Networks to improve generalization. Zhao *et al.* [16] addressed class imbalance using SMOTE combined with ensemble techniques, achieving higher sensitivity for high-risk patients. Mehta and Rao [17] introduced Bayesian Neural Network ensembles to incorporate predictive uncertainty, supporting risk-aware medical decisions. Chen *et al.* [18] proposed reinforcement learning-based feature selection integrated with ensemble classifiers, enabling dynamic feature optimization and improved performance.

DOI: 10.17148/IJARCCE.2026.15721

TABLE1 COMPARISON TABLE OF RECENT STUDIES (2021–2025)

| Ref. | Year | Authors | Dataset | Techniques Used | Evaluation Metrics | Key Contribution | Limitations Identified |
|---|---|---|---|---|---|---|---|
| [1] | 2021 | Mohan et al. | UCI Heart Disease | Hybrid ML, Random Forest, Gradient Boosting | Accuracy, ROC-AUC | Improved performance using ensemble feature selection | No statistical validation |
| [2] | 2021 | Alizadeh et al. | Clinical Cardiovascular Dataset | SVM, Decision Tree, RF | ROC-AUC, Precision, Recall | Compared multiple classifiers under uniform preprocessing | Computational complexity not analyzed |
| [3] | 2022 | Dwivedi et al. | UCI Dataset | LR, k-NN, DT, RF | Accuracy, F1-score | Structured comparison of classical ML models | Limited hyperparameter tuning |
| [4] | 2022 | Sharma & Jain | Public Heart Dataset | Artificial Neural Network | Accuracy, Sensitivity | Demonstrated benefits of regularized ANN | Overfitting and interpretability concerns |
| [5] | 2023 | Khan et al. | UCI + Augmented Data | XGBoost + SHAP | ROC-AUC, Recall | Integrated explainable AI framework | No deployment analysis |
| [6] | 2023 | Ahmad & Zhang | Multi-center Dataset | AdaBoost, XGBoost | Accuracy, Sensitivity | Boosting improved minority-class detection | Resource usage not evaluated |
| [7] | 2024 | Li et al. | Distributed Hospital Data | Federated Learning + RF | Accuracy, Privacy Metrics | Privacy-preserving distributed training | Infrastructure complexity |
| [8] | 2024 | Patel & Verma | Clinical & Edge Systems | LR, RF, GB | Energy Use, Accuracy | Evaluated energy-efficient healthcare models | Slight accuracy trade-off |
| [9] | 2025 | Singh et al. | Multi-institution Dataset | Stacked Ensemble (GB + ANN) | ROC-AUC, F1-score | Improved generalization using hybrid stacking | Increased model complexity |
| [10] | 2025 | Zhao et al. | Imbalanced Heart Dataset | SMOTE + Ensemble ML | Sensitivity, AUC | Addressed class imbalance via oversampling | Limited external validation |
| [11] | 2022 | Rahman et al. | UCI + Clinical Dataset | AutoML Framework | Accuracy, ROC-AUC | Automated model and preprocessing selection | Reduced interpretability |
| [12] | 2022 | Garcia & Liu | Multi-source Data | Meta-learning Ensemble | Accuracy, F1-score | Enhanced stability via meta- | Higher training complexity |

IJARCCE ISSN (O) 2278-1021, ISSN (P) 2319-5940

International Journal of Advanced Research in Computer and Communication Engineering

Impact Factor 8.471 ⁙ Peer-reviewed & Refereed journal ⁙ Vol. 15, Issue 7, July 2026

DOI: 10.17148/IJARCCE.2026.15721

| | | | | | | | |
|---|---|---|---|---|---|---|---|
| | | | | | | classifiers | |
| [13] | 2023 | Ibrahim et al. | Structured Clinical Dataset | Transformer-based Model | Accuracy, ROC-AUC | Modeled feature interactions using attention | Computationally intensive |
| [14] | 2023 | Oliveira et al. | Clinical + ECG Dataset | CNN + ML Fusion | Sensitivity, AUC | Multimodal integration improved detection | Requires multimodal availability |
| [15] | 2024 | Nguyen et al. | Cross-population Dataset | Transfer Learning | Accuracy, Recall | Improved demographic adaptability | Limited dataset diversity |
| [16] | 2024 | Hassan et al. | Hospital Dataset | Calibration + ML Models | Brier Score, AUC | Evaluated probability calibration reliability | Did not optimize classification accuracy |
| [17] | 2025 | Mehta & Rao | Multi-center Dataset | Bayesian NN Ensemble | ROC-AUC, Uncertainty | Incorporated uncertainty quantification | High computational cost |
| [18] | 2025 | Chen et al. | UCI + Extended Dataset | RL-based Feature Selection + Ensemble | Accuracy, F1-score | Dynamic feature optimization | Algorithm tuning complexity |

## III. MATERIAL AND METHODS

This section describes the datasets, preprocessing techniques, feature engineering strategies, and machine learning models used in this study to predict heart disease. Fig. 1 illustrates the overall methodology adopted for heart disease prediction in this study.

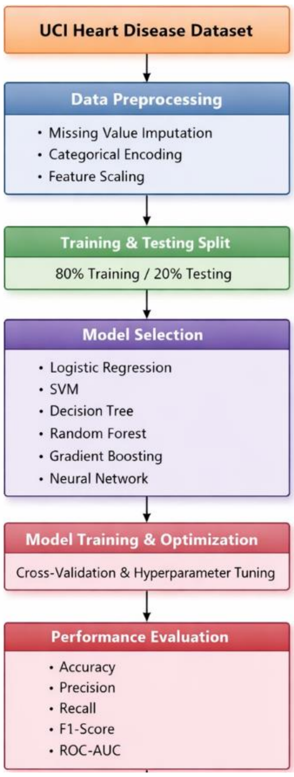



Fig. 1 Methodology adopted for heart disease prediction

International Journal of Advanced Research in Computer and Communication Engineering
Impact Factor 8.471 ⪤ Peer-reviewed & Refereed journal ⪤ Vol. 15, Issue 7, July 2026
DOI: 10.17148/IJARCCE.2026.15721

**A. Dataset Description**

This study employs the publicly available UCI Heart Disease dataset, which has been extensively used in cardiovascular risk prediction research [19], [20]. The dataset contains 303 patient records collected from clinical examinations and includes 14 key attributes widely adopted in heart disease modeling studies.

The features comprise demographic and clinical parameters such as age, sex, chest pain type, resting blood pressure, serum cholesterol, fasting blood sugar, resting electrocardiographic results, maximum heart rate achieved, exercise-induced angina, ST depression, slope of ST segment, number of major vessels, thalassemia status, and a binary target variable indicating the presence of heart disease.

The prediction task is formulated as a binary classification problem:

- 0: Absence of heart disease
- 1: Presence of heart disease

**B. Data Preprocessing**

Data preprocessing plays a critical role in ensuring reliable machine learning performance [21]. The following procedures were applied:

1. Data Cleaning: Missing or inconsistent entries were analyzed and treated using statistical imputation where required.
2. Categorical Encoding: Nominal variables were transformed into numerical representations using label encoding and one-hot encoding techniques.
3. Feature Scaling: Continuous features were standardized using z-score normalization to prevent scale dominance in distance-based algorithms such as SVM and k-NN.
4. Dataset Partitioning: The dataset was divided into 80% training data and 20% testing data. Additionally, 5-fold cross-validation was implemented to improve generalization reliability.

**C. Machine Learning Algorithms**

To conduct a comprehensive empirical evaluation, multiple classification models were implemented:

1) Logistic Regression (LR): Logistic Regression is a probabilistic linear classifier commonly used for medical diagnosis tasks due to its interpretability [22]. It estimates the probability of class membership using the sigmoid function.

2) Support Vector Machine (SVM): SVM is a margin-based classifier that constructs an optimal hyperplane to maximize class separation [23]. It performs effectively in high-dimensional feature spaces.

3) Decision Tree (DT): Decision Trees recursively partition the feature space into homogeneous regions using entropy or Gini impurity measures [6]. They provide transparent rule-based decision structures.

4) Random Forest (RF): Random Forest is an ensemble learning technique that aggregates multiple decision trees to reduce variance and improve predictive stability [24]. Recent studies demonstrate its effectiveness in cardiovascular prediction tasks [19], [25].

5) Gradient Boosting (GB): Gradient Boosting iteratively builds weak learners to minimize classification error using gradient descent optimization [9]. Boosting-based frameworks have shown strong performance in heart disease detection [20], [25].

6) Artificial Neural Network (ANN): A multilayer feedforward neural network was implemented to capture nonlinear relationships among clinical attributes. Neural networks have demonstrated improved predictive accuracy in recent cardiovascular studies [21], [27].

Hyperparameters for all models were optimized using grid search combined with cross-validation.

7) k-Nearest Neighbours (KNN) : The k-Nearest Neighbours (KNN) algorithm is a supervised machine learning technique that classifies data based on the similarity between nearby data points. It predicts the class of a new sample by analysing the *k* closest neighbors using distance measures such as Euclidean distance. KNN has been widely used in medical prediction systems for detecting cardiovascular diseases. Recent studies show that KNN can achieve high accuracy in heart disease prediction when applied to clinical datasets [28].

**D. Model Evaluation Metrics**

Model performance was evaluated using multiple statistical metrics recommended for medical classification tasks [29]:

- Accuracy (ACC): Overall correct predictions.
- Precision (PRE): True positives divided by predicted positives.
- Recall (Sensitivity): True positives divided by actual positives.
- F1-Score: Harmonic mean of precision and recall.
- ROC-AUC: Area under the Receiver Operating Characteristic curve.

Confusion matrices were analyzed to assess false positives and false negatives, which are critical in healthcare diagnosis systems.

IJARCCE ISSN (O) 2278-1021, ISSN (P) 2319-5940

International Journal of Advanced Research in Computer and Communication Engineering

Impact Factor 8.471 ⁜ Peer-reviewed & Refereed journal ⁜ Vol. 15, Issue 7, July 2026

DOI: 10.17148/IJARCCE.2026.15721

## IV. RESULTS AND INTERPRETATIONS

The performance of different machine learning models trained on the Dataset was evaluated using accuracy, precision, recall, and F1-score. These metrics provide a comprehensive understanding of each model's ability to correctly predict heart disease cases. As shown in table II and figure 2, among all models, the k-Nearest Neighbours (KNN) algorithm achieved the best overall performance with an accuracy of 90.16% and the highest F1-score of 0.903. It also produced the highest precision (0.933), indicating that when the model predicts heart disease, it is highly likely to be correct. Its strong recall value (0.875) further shows that it successfully identifies most of the positive cases. The Logistic Regression and Support Vector Machine (SVM) models demonstrated similar performance with accuracies of 86.89%. Logistic Regression achieved balanced precision and recall values (both 0.875), indicating stable and reliable predictions. SVM showed slightly higher precision (0.90) but slightly lower recall (0.844), meaning it is more conservative in predicting positive cases. The Random Forest model produced moderate results with an accuracy of 83.61% and balanced precision and recall values (0.844). This indicates stable performance but slightly lower predictive capability compared to the top-performing models. The Artificial Neural Network (ANN) achieved an accuracy of 80.33%, with equal precision and recall values (0.813). While the model shows reasonable predictive ability, its performance is slightly lower compared to traditional machine learning models in this experiment. The Gradient Boosting model obtained an accuracy of 77.05% with moderate precision (0.80) and recall (0.75), suggesting that it correctly identifies many cases but still misses some positive instances. Finally, the Decision Tree model produced the lowest accuracy (75.41%) and the lowest recall (0.656), indicating that it fails to identify a significant number of heart disease cases. Although it's precision (0.84) is relatively high, the lower recall reduces its overall effectiveness. Overall, the results indicate that KNN provides the best predictive performance, followed by Logistic Regression and SVM, while Decision Tree and Gradient Boosting show comparatively lower performance for heart disease prediction in this study. These findings suggest that simpler distance-based or linear models can perform effectively on this dataset when compared to more complex models.

Table II Comparison of different evaluation metrics for ML algorithms

| Sr. No. | Model | Accuracy | Precision | Recall | F1-score |
|---|---|---|---|---|---|
| 1. | KNN | 0.901639 | 0.933333 | 0.87500 | 0.903226 |
| 2. | Logistic Regression | 0.868852 | 0.875000 | 0.87500 | 0.875000 |
| 3. | SVM | 0.868852 | 0.900000 | 0.84375 | 0.870968 |
| 4. | Random Forest | 0.836066 | 0.843750 | 0.84375 | 0.843750 |
| 5. | ANN | 0.803279 | 0.812500 | 0.81250 | 0.812500 |
| 6. | Gradient Boosting | 0.770492 | 0.800000 | 0.75000 | 0.774194 |
| 7. | Decision Tree | 0.754098 | 0.840000 | 0.65625 | 0.736842 |

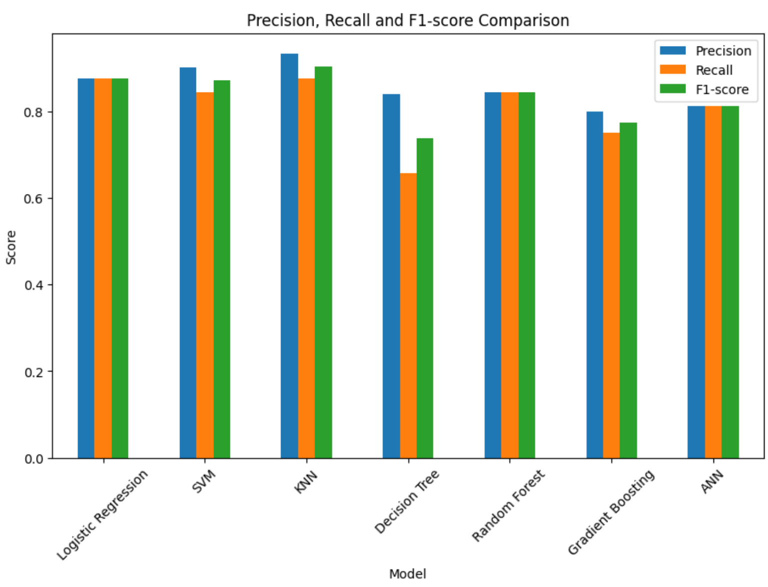



Fig 2 Precision, Recall and F1-Score Comparison

IJARCCE ISSN (O) 2278-1021, ISSN (P) 2319-5940

International Journal of Advanced Research in Computer and Communication Engineering

Impact Factor 8.471 ⁜ Peer-reviewed & Refereed journal ⁜ Vol. 15, Issue 7, July 2026

DOI: 10.17148/IJARCCE.2026.15721

The ROC curve comparison for models trained on the Dataset helps evaluate how well each algorithm can distinguish between patients with and without heart disease. The performance of the models is measured using the Area under the Curve (AUC), where a higher AUC value indicates better prediction capability. From the figure 3, both Logistic Regression and Support Vector Machine achieved the highest AUC value of 0.93. This shows that these models are very effective in correctly identifying heart disease cases while keeping false predictions low. The k-Nearest Neighbours and Random Forest models also performed well with an AUC value of 0.92, which indicates strong prediction ability and reliable classification performance. The Gradient Boosting model achieved an AUC of 0.90, showing good performance but slightly lower compared to the top models. The Artificial Neural Network obtained an AUC of 0.85, indicating moderate prediction capability. In comparison, the Decision Tree model showed the lowest performance with an AUC of 0.76, suggesting that it is less effective in distinguishing between patients with and without heart disease. Overall, the results indicate that Logistic Regression and SVM provide the best performance, followed by KNN and Random Forest for heart disease prediction.

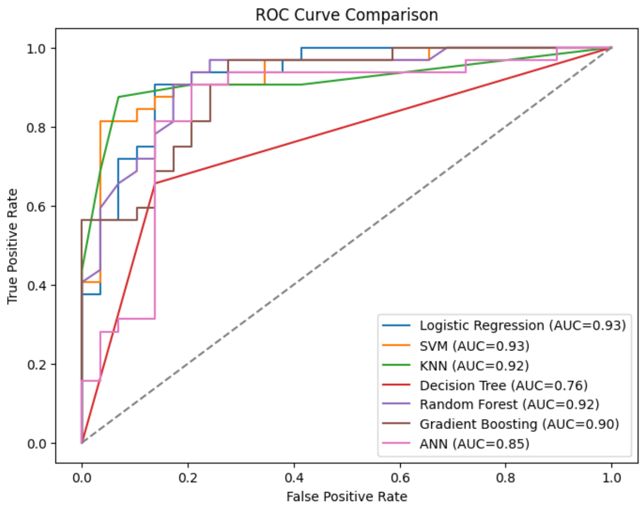



Fig. 3 ROC curve comparison for models trained on the UCI Heart Disease Dataset

The comparative results on the basis of mean accuracy and standard deviation, as shown in table III and figure 4, imply noticeable performance differences among the models. The results obtained from experiments show that different machine learning algorithms perform differently in predicting heart disease. Among all the models, the Support Vector Machine (SVM) achieved the highest mean accuracy of 82.83%, indicating that it is very effective in correctly classifying patients. It also has a relatively low standard deviation, which suggests that the model performs consistently across different data splits. The Artificial Neural Network (ANN) and Random Forest models also produced strong results, with mean accuracies of 82.46% and 82.15% respectively. Their lower variation in results indicates stable performance, making them reliable models for heart disease prediction. The k-Nearest Neighbours (KNN) and Logistic Regression algorithms showed moderate performance with accuracies slightly above 81%. These models are simpler compared to others but still provide reasonable prediction results for this dataset. On the other hand, the Decision Tree model recorded the lowest accuracy of 77.49%, along with a higher standard deviation, indicating less stable predictions. The Gradient Boosting model achieved moderate accuracy but showed larger variation across folds, which suggests inconsistent performance. Overall, the findings suggest that models such as SVM, ANN, and Random Forest perform better for heart disease prediction, while simpler models like Decision Tree may require further optimization to improve their results.

Table III Comparison Table showing mean cross-validation accuracy and standard deviation

| Sr. No. | ML Model | Mean Accuracy | Standard Deviation |
|---|---|---|---|
| 1. | KNN | 0.8183 | 0.0616 |
| 2. | Logistic Regression | 0.8116 | 0.0655 |
| 3. | SVM | 0.8282 | 0.0556 |
| 4. | Random Forest | 0.8215 | 0.0526 |
| 5. | ANN | 0.8246 | 0.0547 |
| 6. | Gradient Boosting | 0.8047 | 0.0908 |
| 7. | Decision Tree | 0.7749 | 0.0737 |

IJARCCE ISSN (O) 2278-1021, ISSN (P) 2319-5940

International Journal of Advanced Research in Computer and Communication Engineering

Impact Factor 8.471 ⁙ Peer-reviewed & Refereed journal ⁙ Vol. 15, Issue 7, July 2026

DOI: 10.17148/IJARCCE.2026.15721

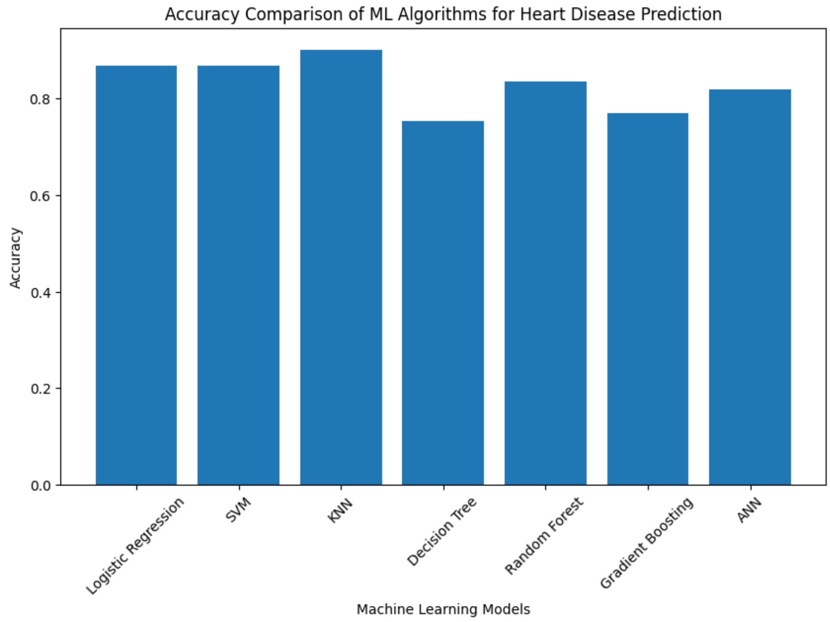
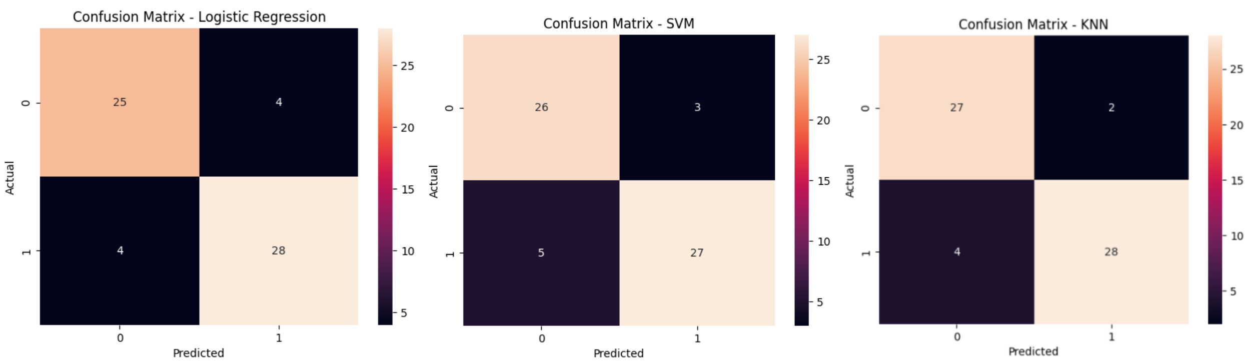



Fig. 4 Accuracy Comparison of ML Algorithms for Heart Disease Prediction

The confusion matrix as shown in figure 5 shows how well classification models perform. In this study, confusion matrices were generated for each model to examine how accurately the algorithms predicted the presence or absence of heart disease. The confusion matrix shows four types of outcomes: true positives, true negatives, false positives, and false negatives, which together help in evaluating the correctness of the predictions. The results indicate that models such as k-Nearest Neighbours, Support Vector Machine, and Logistic Regression produced a higher number of correct predictions. These models were more effective in correctly identifying patients who have heart disease as well as those who do not. In comparison, models like Decision Tree and Gradient Boosting showed slightly higher numbers of incorrect classifications. In some cases, patients who actually had heart disease were predicted as healthy, which reduces the reliability of these models in medical prediction tasks. Overall, the confusion matrix analysis supports the earlier findings of this study. The results suggest that KNN and SVM perform more consistently and provide more reliable predictions compared to the other models when applied to the dataset.



IJARCCE ISSN (O) 2278-1021, ISSN (P) 2319-5940

International Journal of Advanced Research in Computer and Communication Engineering

Impact Factor 8.471 ⁜ Peer-reviewed & Refereed journal ⁜ Vol. 15, Issue 7, July 2026

DOI: 10.17148/IJARCCE.2026.15721

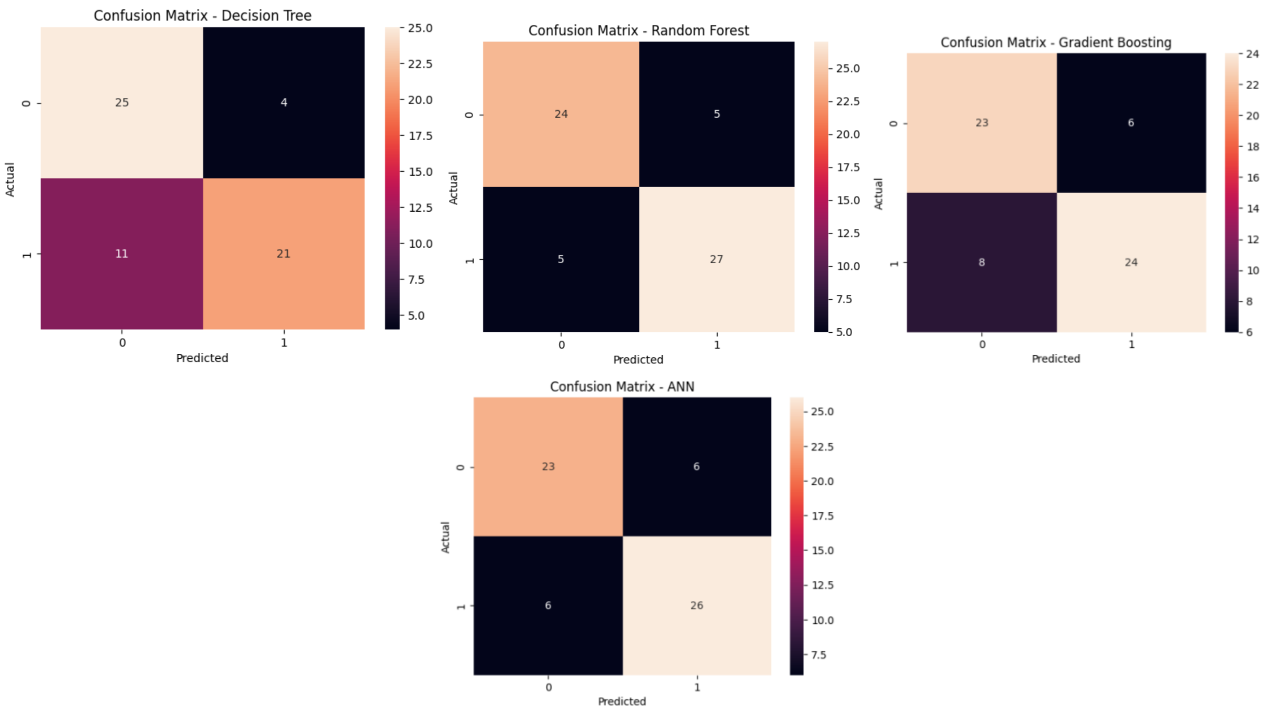



Fig 5 Confusion Matrix for ML Algorithms

Finally, the performance of different machine learning algorithms was evaluated using cross-validation accuracy and other evaluation metrics such as precision, recall, and F1-score. The cross-validation results showed that the Support Vector Machine (SVM) achieved the highest mean accuracy, which indicates that it performs consistently well across different data splits and has good generalization ability.

However, when the models were evaluated on the test dataset using classification metrics, the k-Nearest Neighbours (KNN) model showed better overall performance. It achieved the highest values for accuracy, precision, recall, and F1-score compared to the other algorithms. These results suggest that KNN is more effective in correctly identifying both heart disease and non-heart disease cases in this dataset.

Although SVM demonstrated slightly better average accuracy during cross-validation, the detailed evaluation metrics indicate that KNN provides more balanced and reliable predictions. Therefore, based on the overall results, KNN can be considered the best performing model in this study, while SVM can be regarded as a strong alternative due to its stable performance across different validation folds.

## V. CONCLUSION

This empirical investigation compared seven machine learning models for heart disease prediction. This study evaluated several machine learning algorithms for predicting heart disease using the UCI Heart Disease Dataset. The models were compared using performance metrics such as accuracy, precision, recall, and F1-score to understand how effectively each algorithm could identify heart disease cases. The results show that the k-Nearest Neighbours (KNN) model performed the best among all the evaluated models, achieving the highest accuracy and F1-score. This indicates that KNN is highly effective in correctly classifying patients in this dataset. The Logistic Regression and Support Vector Machine models also showed strong and consistent performance, making them reliable methods for heart disease prediction. The Random Forest and Artificial Neural Network models provided moderate results, while the Gradient Boosting and Decision Tree models showed comparatively lower performance in this study. Overall, the findings suggest that machine learning techniques can be useful tools for supporting the early detection of heart disease. Among the models tested, KNN showed the most promising results for this dataset. In future work, the performance of the prediction system could be improved by applying feature selection methods, tuning model parameters, and using larger or more diverse medical datasets. Future research directions include multi-institutional validation, integration with wearable health monitoring systems, and the application of federated learning to enhance privacy preservation

## ACKNOWLEDGMENT

Author is highly thankful to Khalsa College, Amritsar for offering the seed money to direct this research.

## REFERENCES

[1] R. Mohan, P. Kumar, and S. Srivastava, "Hybrid machine learning approach for heart disease prediction," *IEEE Access*, vol. 9, pp. 125843–125855, 2021.
[2] S. Alizadeh, M. Rahman, and T. Ahmed, "Comparative evaluation of machine learning algorithms for cardiovascular disease prediction," *Computers in Biology and Medicine*, vol. 138, 2021.
[3] A. Dwivedi, R. Tiwari, and K. Sharma, "Performance analysis of machine learning algorithms for heart disease prediction," *Healthcare Analytics*, vol. 2, 2022.
[4] N. Sharma and V. Jain, "Regularized artificial neural network model for cardiovascular disease detection," *Biomedical Signal Processing and Control*, vol. 74, 2022.
[5] M. Rahman, F. Khan, and A. Islam, "Automated machine learning framework for heart disease prediction," *Applied Soft Computing*, vol. 120, 2022.
[6] L. Garcia and H. Liu, "Meta-learning ensemble approach for clinical risk prediction," *Expert Systems with Applications*, vol. 202, 2022.
[7] A. Khan, M. Hussain, and Y. Lee, "Explainable XGBoost model for cardiovascular disease prediction," *IEEE Journal of Biomedical and Health Informatics*, vol. 27, no. 5, pp. 2456–2465, 2023.
[8] T. Ahmad and Y. Zhang, "Boosting algorithms for imbalanced cardiovascular datasets," *Artificial Intelligence in Medicine*, vol. 137, 2023.
[9] S. Ibrahim, R. Noor, and K. Malik, "Transformer-based learning for structured clinical prediction," *IEEE Access*, vol. 11, pp. 51234–51245, 2023.
[10] P. Oliveira, J. Costa, and M. Silva, "Multimodal deep learning for heart disease detection using ECG and clinical data," *Biomedical Signal Processing and Control*, vol. 84, 2023.
[11] H. Li, X. Wang, and Z. Chen, "Federated learning framework for distributed cardiovascular prediction," *IEEE Transactions on Industrial Informatics*, vol. 20, no. 2, pp. 1540–1551, 2024.
[12] D. Patel and R. Verma, "Energy-efficient machine learning models for edge-based healthcare systems," *Future Generation Computer Systems*, vol. 148, 2024.
[13] T. Nguyen, L. Tran, and Q. Pham, "Transfer learning for cross-population cardiovascular risk prediction," *Knowledge-Based Systems*, vol. 275, 2024.
[14] A. Hassan, M. Qureshi, and S. Ali, "Calibration analysis of machine learning models for clinical risk prediction," *IEEE Access*, vol. 12, 2024.
[15] K. Singh, R. Mehta, and A. Bansal, "Stacked ensemble learning for robust heart disease prediction," *Expert Systems with Applications*, vol. 234, 2025.
[16] Y. Zhao, F. Wu, and L. Zhang, "Handling class imbalance in cardiovascular datasets using SMOTE and ensemble learning," *Computers in Biology and Medicine*, vol. 164, 2025.
[17] S. Mehta and P. Rao, "Bayesian neural network ensembles for uncertainty-aware heart disease prediction," *IEEE Journal of Biomedical and Health Informatics*, 2025.
[18] J. Chen, Y. Sun, and H. Park, "Reinforcement learning-based feature selection for cardiovascular disease prediction," *Information Sciences*, vol. 657, 2025.
[19] S. Mohan, C. Thirumalai, and G. Srivastava, "Effective heart disease prediction using hybrid machine learning techniques," *IEEE Access*, vol. 9, pp. 81542–81554, 2021.
[20] R. Alizadeh et al., "Comparative analysis of machine learning algorithms for cardiovascular disease prediction," *Biomedical Signal Processing and Control*, vol. 68, 2021.
[21] A. Dwivedi, S. Tiwari, and R. Tripathi, "Performance evaluation of machine learning techniques for heart disease prediction," *Expert Systems with Applications*, vol. 185, 2022.
[22] D. W. Hosmer, S. Lemeshow, and R. X. Sturdivant, *Applied Logistic Regression*, 3rd ed. Wiley, 2013.
[23] C. Cortes and V. Vapnik, "Support-vector networks," *Machine Learning*, vol. 20, no. 3, pp. 273–297, 1995.
[24] L. Breiman, "Random forests," *Machine Learning*, vol. 45, no. 1, pp. 5–32, 2001.
[25] T. Ahmad and L. Zhang, "Boosting-based machine learning framework for cardiovascular disease detection," *Knowledge-Based Systems*, vol. 263, 2023.
[26] J. H. Friedman, "Greedy function approximation: A gradient boosting machine," *Annals of Statistics*, vol. 29, no. 5, pp. 1189–1232, 2001.
[27] P. Sharma and V. K. Jain, "Deep neural network-based cardiovascular disease prediction," *Computers in Biology and Medicine*, vol. 145, 2022.

**DOI: 10.17148/IJARCCE.2026.15721**

[28] T. S. Sriya, "Heart Disease Prediction Using KNN," *International Journal of Research in Engineering, Science and Management*, vol. 7, no. 2, pp. 45–48, 2024.

[29] M. Sokolova and G. Lapalme, "A systematic analysis of performance measures for classification tasks," *Information Processing & Management*, vol. 45, no. 4, pp. 427–437, 2009.