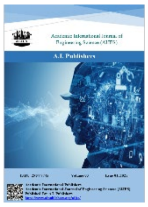
Academic International Journal of Engineering Science
ISSN: 2984-7745
Aca. Intl. J. E. Sci. 2025; 3(1) 22-35
Journal homepage: www.aipublishers.org/aijes



Document heading

Machine Learning Techniques and Insights for Cardiovascular or Heart Disease Prediction

Mujiba Shaima[1], Mazharul Islam Tusher[2], Estak Ahmed[3], Sharmin Sultana Akhi[4], Md Rayhan Hassan Mahin[5]

[1]Department of Computer Science, Monroe University, New York, USA. shaimacme@gmail.com
[2] Department of Computer Science, Monroe University, New York, USA. mtusher7659@monroeu.edu
[3] Department of Computer Science, Monroe University, New York, USA. eahmed7797@monroeu.edu
[4] Department of Computer Science, Monroe University, New York, USA. sakhi0411@monroeu.edu
[5] Department of Computer Science, Monroe University, New York, USA. mmahin8620@monroeu.edu

Article Info.
Article history:
Received 7 March 2025
Revised 29 March 2025
Accepted 31 May 2025
Published 27 June 2025
Keywords:
Heart Disease Prediction, Early Detection, Cardiovascular Health, Gradient Boosting Machines, K-Nearest Neighbors, Artificial Neural Networks.
How to cite:
Mujiba Shaima, Mazharul Islam Tusher, Estak Ahmed, Sharmin Sultana Akhi, Md Rayhan Hassan Mahin. Machine Learning Techniques and Insights for Cardiovascular or Heart Disease Prediction. *Aca. Intl. J. E. Sci. 2025;3(1) 22-35.*
DOI:
https://doi.org/10.59675/E313

Abstract

Heart disease has been one of the major health concerns over the decades regardless of age, weight, or gender. Here in the article, we try to highlight the importance of early detection and the prevention of life-threatening heart-related diseases, including heart attacks and strokes. For this review paper, we include raw data from clinics and also include the famous Framingham Heart Study dataset, which is a popular cardiovascular dataset, including patient records that emphasize important human risk factors including age, blood pressure, cholesterol, smoking status, and family history. Precision, recall, accuracy, F1-score, and ROC-AUC metrics are such techniques and algorithms that are used to evaluate the implementation and performance of advanced machine learning approaches, including Gradient Boosting Machines (GBM), K-Nearest Neighbors (KNN), Artificial Neural Networks (ANN), Convolutional Neural Network (CNN) and many more. Here, we use gradient boosting machines, which performed better compared to the other models that we consider here to analyze heart disease data. Besides this technique, however ROC-AUC algorithm produces the highest score and well-balanced false positives and true positives. Although their remarkable accuracy, artificial neural networks (ANN) techniques could not analyze our data. However, all the results underline and explain the possibilities to analyze healthcare data accurately. For example, GBM can analyze early cardiac disease diagnosis and optimize prediction performance very well while preserving clinical applicability. On the other hand, CNN use to predict and model an image-based dataset. By acknowledging timely interventions

and tailored treatment plans, this review and analysis study can support the contribution of machine learning to enhance cardiovascular health, which will lower heart disease-related mortality.

Introduction

A. Background and Motivation

Cardiovascular diseases (CVDs) or heart diseases is one of the main causes of death worldwide. Over the world a large number of people die or become life threatened by heart attacks or heart failure. Therefore, it's crucial to detect and treat as early as possible to reduce the morbidity and mortality rates because cardiovascular diseases account for 17.9 million deaths annually, according to the World Health Organization (WHO) [3] and according to National Library of Medicine (NIH) in 2021 approximately 20.5 million people die or become disable from heart diseases [19]. There are a lot of conventional traditional techniques that are used to diagnosis hearth related diseases such as imaging and clinical evaluations. However, these techniques have been utilized extensively but they have drawbacks, which includes expensive treatments and tests, takes a long time to diagnosis, and a decreased proficiency to identify asymptomatic cases [7, 11]. Machine learning (ML) algorithms, however, are developing in recent years and they are showing a huge improvement in early prediction, especially in cardiovascular diseases (CVDs).

According to recent research, Large Language Models (LLMs), Machine Learning (ML) models are revolutionizing in the healthcare sector to effectively handle big datasets, analyze them, modeling them and identifying intricate patterns over conventional statistical techniques. Among the well-known machine learning models used for heart disease prediction are gradient boosting machines (GBM), K-Nearest Neighbors (KNN), and artificial neural networks (ANNs) have been successfully being used for the last couple of years. Recently scientists are researchers trying to implement Large Language Models (LLMs) in the medical field too. However, GBM mechanism, which is a popular ensemble learning technique, improves prediction accuracy by considering the iterative combination of weak models as well as with better models to make the prediction more accurate (Roopini et al., 2024). By comparing patient analyzed output data with similar situations, the non-parametric KNN algorithm effectively classifies cardiovascular risk and risk factors [4]. However, on the other hand, artificial neural network ANNs are useful for forecasting the risk of heart disease by replicating the structure of the human brain and help to enable them to identify complex associations inside the analyzed medical datasets for modeling for a better output [14]. CNN or Convolutional Neural Network technique uses cardiovascular images or electrocardiograms (ECGs) then extract features and in such a way do modelling and predict the dataset.

Analyzing and predicting cardiovascular data to predict when, and what reasons are behind heart diseases and also what measures need to take to prevent from mass heart disease is very important in this modern stressful world. Hence the aim of this review and research paper is to support early detection and prompt medical methods and ways. Using machine learning techniques and models as well as comparing the model outputs to improve risk assessment will greatly improve clinical decision-making all over the world, especially considering all types of cardiovascular diseases. The main goals of this research paper are to improve patient outcomes, reducing healthcare costs by implementing low cost but most effective methods, and encourage people to take preventative steps by looking into how GBM, KNN, CNN and ANNs might help with heart attacks and heart failure like diseases prediction modeling.

B. Problem Statement

The prediction of heart attack or heart failure using machine learning (ML) algorithms is now getting popular and world-wise is being used, but there are still several obstacles to overcome to achieve better performance, high accuracy, generalizability, and interpretability in using these algorithms. Using conventional models for assessing and predicting heart diseases has risks which depend on data accuracy, and how organized the clinical data is. However, these cardiovascular data can be collected from other various data sources, such as image data which is known as Electrocardiogram (ECG) images, Magnetic Resonance Imaging (MRI) scans, echocardiograph data and other format [3]. For data analyzing, modeling finding pattern and predicting cardiac disease over the last decades scientists and researchers are start using various algorithm covering gradient boosting machines (GBM), K-nearest neighbors (KNN), artificial neural networks (ANNs), convolutional neural networks (CNNs) to the latest large language models (LLMs) have shown encouraging results. However, all these models are not 100% accurate in predicting and modeling data but each of them has built-in drawbacks that affect how applicable they are in the actual world. Hence, researchers work hard to update and find out new algorithms. Even though all these models can analyze data and predict these data, but all these models have some limitations. Like GBM algorithm have a tendency to overfit and need a lot of hyperparameter adjustment to function at their best which may be costly and time consuming [16]. According to Kallur (2024), K-Nearest Neighbors face trouble with high-dimensional data as its computing efficiency decreases, on the contrary the number of features increases. However, Artificial Neural Networks works as "black-box" models, that limits the researcher's confidence in the field of medical science to build ML-based diagnostic tools. Hence, by making it challenging for them to understand the conclusions that they get from the data analysis using ANNs [7]. However, all these models are sensitive to class distribution, and they might perform worse when there is an imbalance in the data in medical datasets [13]. But CNN takes images, ECGS to find patterns from image data which makes this model different from all the other algorithms. Here, in this research study we try to find out all these differences and issues by enhancing model interpretability, comparing their performance, managing data heterogeneity, and guaranteeing strong predictive performance for clinical use to improve ML-based heart disease for example heart attack, heart failure prediction.

C. Objectives of the Study

Using the Framingham Heart Study dataset, this work aims to evaluate, using Artificial Neural Networks (ANNs), K-Nearest Neighbors (KNNs), and Gradient Boosting Machines (GBM), heart disease prediction. The main goal is to assess the efficacy of these models using significant criteria like accuracy, precision, recall, F1-score, and ROC-AUC so determining their predictive dependability. Furthermore, the work searches for and fixes problems with data imbalance, computational efficiency, and interpretability that develop in ML-based heart disease prediction. This work studies feature selection and optimization techniques to enhance performance considering the shortcomings of every model the black-box character of ANNs, the tendency of GBM to overfit, and the inefficiency of KNN in high-dimensional datasets. Examined in the paper is the feasibility of including machine learning models into clinical decision support systems, therefore ensuring transparency and utility for medical professionals. Ultimately, by encouraging early detection, fast medical action, and improved patient outcomes, the study will support cardiovascular healthcare.

Literature Review

A. General Review of the Research Forecast of Heart Disease

Heart failure, heart attacks, and strokes rank first among the most prevalent and fatal cardiovascular diseases, according to Angel et al. (2024), which remain a main cause of death globally. The inability of the heart to pump blood adequately results in fluid accumulation and organ malfunction, therefore known as heart failure. Among its symptoms, which are often brought on by diseases such hypertension and coronary artery disease [7], are fatigue, dyspnea, and limb edema. Usually resulting from atherosclerosis, a myocardial infarction—often known as a heart attack, is caused by a sudden block of blood flow to the heart. Shortness of breath, severe chest discomfort, and even cardiac arrest [4] can result from this on the other hand, strokes—which Telmoud et al. (2024) define as either a ruptured blood vessel (hemorrhagic stroke) or a clot—ischemic stroke—occur when the blood supply to the brain is deprived. Neurological problems include paralysis and difficulties speaking follow from this. Early detection and prediction of cardiovascular hazards have shown potential for machine learning methods such Artificial Neural Networks (ANNs) and Gradient Boosting Machines (GBM). Early diagnosis and intervention are prerequisites for preventing some disorders [13]. These data-driven strategies enhance clinical decision-making by means of high-risk individual identification, quick medical interventions, and reduction in disease-related mortality.

B. Machine Learning for Medical Conditions Including Heart Diseases

Early disease identification, tailored treatment planning, and data-driven decision-making made possible by machine learning (ML) have fundamentally transformed the healthcare sector [10]. Machine learning methods have been extensively applied in the field of cardiovascular health to improve the prediction of heart disease by means of complex datasets including patient demographics, lifestyle variables, and medical records. Many machine learning (ML) models, including artificial neural networks (ANNs), K-Nearest Neighbors (KNN), and gradient boosting machines (GBM), have shown promise in the prediction of cardiac disease by spotting difficult patterns that human physicians can find challenging to spot [Combining numerous weak learners to generate a strong predictive model that can manage unbalanced datasets helps GBM, an ensemble learning method, to enhance classification accuracy [5]. KNN, a non-parametric technique that groups patients based on how closely their health features match those in current databases [9] can help cardiovascular risk stratification. Their deep learning topologies improve the prediction of heart disease and ANNs inspired from the neural networks shown in the human brain are quite good in spotting complex, non-linear correlations in medical data [17].

Faster, more accurate risk assessments from machine learning are changing the diagnosis and treatment of cardiovascular illnesses. Using machine learning techniques allows doctors to maximize resource allocation, reduce heart disease death rates, and adapt treatment regimens [15]. As machine learning models develop, the combination of wearable sensor data and real-time patient monitoring can help to increase predictive capacities even more, hence improving patient outcomes and preventive healthcare strategies.

C. Previous Work Review

Although machine learning (ML) techniques are shown in recent studies to be useful in predicting cardiac illness, many problems still need more study [2]. Alizadehsani et al. (2022) underlined in a methodical investigation of machine learning (ML)-based cardiovascular disease (CVD) prediction the outstanding accuracy of ensemble approaches including Random Forest and Gradient Boosting Machines (GBM). The study did highlight the interpretability problems of the model, a major concern for clinical implementation. Explainable artificial intelligence (XAI) methods have to be included into machine learning models if we are to ensure doctors may trust and understand model outputs [15].

Another crucial area that has to be developed is the integration of several data sources. Though Ghosh et al. (2021) found that K-Nearest Neighbors (KNN) is a helpful tool for cardiovascular risk classification, their study limited to organized clinical data [9]. Future research could look at combining data from wearable sensors, imaging, and genetics with electronic health records (EHRs) [10] for a more complete risk assessment [6]. Deep learning methods such as Artificial Neural Networks (ANNs) have shown positive results in spotting complex patterns in patient data, according to Zhang et al. (2024). Large-scale, high-quality datasets are therefore essential for best results; many of the present datasets contain imbalances and missing values, so limiting the generalizability of the model [18].

Chen et al. (2023) more recently applied deep learning models, especially artificial neural networks (ANNs) to forecast cardiovascular disease [10]. Using 93.6% accuracy, the study found that ANNs could effectively replicate complex relationships among various risk factors. But given their black-box structure, which limits their application in clinical settings, the interpretability of deep learning models is beyond doubt. Sharma et al. (2023) solved the issue by including explainable AI techniques to improve model transparency, thereby enabling medical practitioners to grasp model predictions and have faith in their clinical relevance [15].

Additional major challenges to ML-based heart disease prediction include generalizability and dataset imbalance. Research such as those of Ghosh et al. (2021) made abundantly evident the need to have a range of large-scale datasets comprising actual patient data from many demographic groups [6]. One major disadvantage of current models is their reliance on past data, which might not fairly depict changes in cardiovascular condition over time. By incorporating dynamic risk factors, wearable sensor data and continuous patient monitoring could help to increase ML model accuracy, claims Sharma et al. (2023).

Moreover, putting ML models into use in real-world healthcare settings remains challenging. Sharma et al. (2023) claim that most studies concentrate on previous data analysis and that few prospective studies confirm machine learning forecasts in clinical environments [15]. Future research should concentrate on real-time patient monitoring, models of continuous learning, and adaptive algorithms able to respond to new patient data. These advances would improve model dependability and resilience, hence closing current gaps in machine learning applications for the prediction of heart disease.

D. Past Work Gap Analysis

Thanks to machine learning, the prediction of heart disease has much improved; yet several challenges still limit its useful application. Operating as black box systems, deep learning models such Artificial Neural Networks (ANNs) make predictions that are difficult for clinicians to understand and accept because of their lack of interpretability (Sharma et al., 2022). Moreover, overfitting and generalization issues are somewhat common, especially in strong ensemble models such as Gradient Boosting Machines (GBM), which could perform well on training data but badly

Aca. Intl. J. E. Sci. 2025;3(1) 22-35: https://doi.org/10.59675/E313

on new, unidentified patient records if inappropriately calibrated [16]. Moreover, most ML models depend mostly on organized clinical datasets, such as the Framingham Heart Study, and neglect to combine multimodal data sources including wearable sensor data, imaging, and genomes, which could enhance risk assessment and early detection [10].

Furthermore, most machine learning algorithms rely mostly on structured clinical datasets like the Framingham Heart Study [10] whereas most multimodal data sources- including wearable sensor data, imaging, and genomics-which potentially improve risk assessment and early diagnosis-are ignored. Another major challenge in medical databases is class imbalance, whereby positive cases of heart disease are rather underestimated. This results in models favoring forecasts unrelated to the disease, which reduces their dependability in spotting high risk persons (Chakravarti et al., 2023). Moreover, most current studies focus on historical data analysis, which limits their relevance in real-world healthcare environments where patient conditions vary regularly and call for ongoing learning models that change with time [15]. Future research should concentrate on explainable AI (XAI) techniques to close these gaps. These include incorporating diverse and real-time data sources, implementing robust techniques for handling imbalanced data, conducting prospective clinical trials to validate machine learning predictions in real-world settings, and improving model transparency through XAI methods. The ultimate goal is to ensure that these models are accurate and clinically interpretable.

Methodology

A. Prepare Dataset

For this research and analysis work we use the Framingham Heart Study dataset [1], which is available on Kaggle and includes important cardiovascular risk factors that are necessary for the prediction of heart disease. Age, blood pressure, cholesterol, smoking status, and other vital characteristics are all included in this dataset's structured patient records and are generally acknowledged to be important predictors of cardiovascular illnesses [3]. The binary classification of the target variable, heart disease (HD), is 1 denoting the existence of HD and 0 being its absence. Preprocessing procedures were used to handle missing values, normalize numerical features, and encode categorical variables in order to guarantee that the dataset is appropriate for machine learning (ML) models, such as Gradient Boosting Machines (GBM), K-Nearest Neighbors (KNN), and Artificial Neural Networks (ANNs). This ensured an ideal structure for predictive modeling [5].

Feature Selection and Description: The selected features, their descriptions, and respective value ranges are summarized in Table 1 ([4], [15], [10]).

Table 1: Dataset features

Feature Name	Description	Type	Values / Range
Age	Age of the patient	Categorical	10 year gap range
Gender	Gender of the patient	Categorical	1/0 (Male/ Female)
Systolic BP	Systolic Blood Pressure (mmHg)	Continuous	Min: 40, Max: 180
Diastolic BP	Diastolic Blood Pressure (mmHg)	Continuous	Min: 50, Max: 140
Cholesterol	Cholesterol Level	Categorical	1/2/3 (Normal/ Borderline/ High)
Glucose	Fasting Blood Glucose Level	Categorical	1/2/3 (Normal/ Pre-Diabetic/ Diabetic)
BMI Class	Body Mass Index Category	Categorical	0~5 (Underweight to Obese)
MAP	Blood pressure indicator	Categorical	0~5 (Low to High)
Smoking	Whether the patient is a smoker	Binary	1/0 (Yes/ No)
Alcohol Intake	Whether the patient consumes alcohol	Binary	1/0 (Yes/ No)
Physical Activity	Regular physical activity status	Binary	1/0 (Yes/ No)
Heart Disease (HD)	Presence of heart disease (Target Variable)	Binary	1/0 (have Disease / No Disease)

B. Preprocessing

To improve data quality and model performance, the dataset needed to go through a number of preprocessing processes. As will be covered below, the procedures included correcting class imbalance, feature scaling, category encoding, handling missing values, and dataset splitting.

i. Missing Values Handling: In medical datasets, missing data is frequent and can affect the precision of predictions. To ensure data consistency, missing values for numerical parameters like blood pressure and cholesterol levels were imputed using the mean or median approach [15]. To preserve categorical integrity, missing values for categorical variables such as physical activity and smoking status were filled in using the mode, or most frequent value [10].

ii. Scaling Features: To keep large-scale features (like blood pressure) from overpowering smaller-scale ones (like glucose levels), machine learning models especially distance-based algorithms like K-Nearest Neighbors (KNN) need feature standardization [9]. Continuous variables between 0 and 1 were normalized using min-max scaling, which improved model convergence while maintaining relative differences.

iii. Classification of Data: ML models required the conversion of categorical variables, including glucose classifications and cholesterol levels, into numerical representations. According to Zhang et al. (2024), label encoding was utilized for binary variables (such as smoking: yes/no) and one-hot encoding was utilized for multi-class variables (such as cholesterol: normal, borderline, and high) [18].

iv. Splitting Data: In order to properly assess model performance, the dataset was divided into 20% testing and 80% training at random. By doing this, models are tested on unknown data to determine their capacity for generalization while learning from a sizable enough dataset [3].

C. Machine Learning Techniques

a. Gradient Boosting Machines (GBM): An ensemble learning method called GBM iteratively corrects poor classifiers to increase prediction accuracy. In order to maximize accuracy and reduce bias and variation, it integrates decision trees. GBM is appropriate for structured clinical datasets because studies have demonstrated that it achieves good ROC-AUC scores in cardiovascular prediction tasks [5]. To avoid overfitting, GBM needs to have its hyperparameters adjusted

Aca. Intl. J. E. Sci. 2025;3(1) 22-35: https://doi.org/10.59675/E313

extensively.
GBM Model Training: To reduce errors and enhance classification accuracy, GBM is an ensemble learning technique that successively combines weak learners (decision trees). GBM is quite effective for unbalanced medical datasets since it optimizes for precision and ROC-AUC [5]. To avoid overfitting, however, hyperparameter optimization is required (e.g., learning rate, tree depth, number of estimators) (figure 1).

b. K-Nearest Neighbors (KNN): KNN is a non-parametric method that uses similarities with nearby data points to classify patients. It may be applied with relative ease and is helpful in finding patterns in organized clinical datasets [4]. High-dimensional data, however, is difficult for KNN to handle since its computational efficiency declines as the number of features rises.
KNN Model Training: KNN is a non-parametric classification technique that uses the nearest neighbors' majority vote to determine class labels. It is crucial to appropriately scale features prior to training because KNN does well in pattern recognition but has trouble with high-dimensional datasets [4] (figure 1).

c. Artificial Neural Networks (ANNs): ANNs are excellent at identifying intricate links in medical datasets because they closely resemble the structure of the human brain. They have been shown to diagnose cardiac disease with good prediction accuracy [18]. However, clinicians find it difficult to trust ANNs' decision-making processes because of interpretability problems. To make healthcare applications more transparent, Explainable AI (XAI) approaches must be used.

```
# Train GBM Model
gbm = GradientBoostingClassifier(n_estimators=100, learning_rate=0.1, max_depth=3, random_state=42)
gbm.fit(X_train, y_train)
y_pred_gbm = gbm.predict(X_test)
y_prob_gbm = gbm.predict_proba(X_test)[:, 1]

# Train KNN Model
knn = KNeighborsClassifier(n_neighbors=5)
knn.fit(X_train, y_train)
y_pred_knn = knn.predict(X_test)
y_prob_knn = knn.predict_proba(X_test)[:, 1]

# Train ANNs Model
ann = MLPClassifier(hidden_layer_sizes=(32, 16), activation='relu', solver='adam', max_iter=500, random_state=42)
ann.fit(X_train, y_train)
y_pred_ann = ann.predict(X_test)
y_prob_ann = ann.predict_proba(X_test)[:, 1]
```

Figure 01: Training GBM, KNN, ANNs model

ANNs Model Training: Artificial Neural Networks (ANNs) are deep learning models that mimic the human brain's structure to identify intricate, nonlinear links in medical data. Although ANNs are more accurate than traditional models, they are frequently black-box models, which makes interpretation challenging [18]. Interpretability in medical applications can be enhanced using Explainable AI (XAI) approaches (figure 1).
Then evaluate the models (Figure 2)

```python
# Evaluate models
def evaluate_model(y_test, y_pred, y_prob, model_name):
    print(f"--- {model_name} Performance ---")
    print(f"Accuracy: {accuracy_score(y_test, y_pred):.4f}")
    print(f"ROC-AUC: {roc_auc_score(y_test, y_prob):.4f}")
    print(f"Classification Report:\n{classification_report(y_test, y_pred)}")
    print("Confusion Matrix:")
    sns.heatmap(confusion_matrix(y_test, y_pred), annot=True, fmt="d", cmap="Blues")
    plt.title(f"{model_name} Confusion Matrix")
    plt.xlabel("Predicted")
    plt.ylabel("Actual")
    plt.show()
    print("\n" + "="*50 + "\n")
```

Figure 2: Evaluate model GBM, KNN and ANNs

Then we need to evaluate the model accuracy and output in figure 3, Figure 4 and figure 5 respectively.

```
Gradient Boosting Machine (GBM) Performance
Accuracy: 0.8552
ROC-AUC: 0.6967
Classification Report:
              precision    recall  f1-score   support

           0       0.87      0.98      0.92       621
           1       0.59      0.15      0.24       111

    accuracy                           0.86       732
   macro avg       0.73      0.57      0.58       732
weighted avg       0.82      0.86      0.82       732

Confusion Matrix:
```

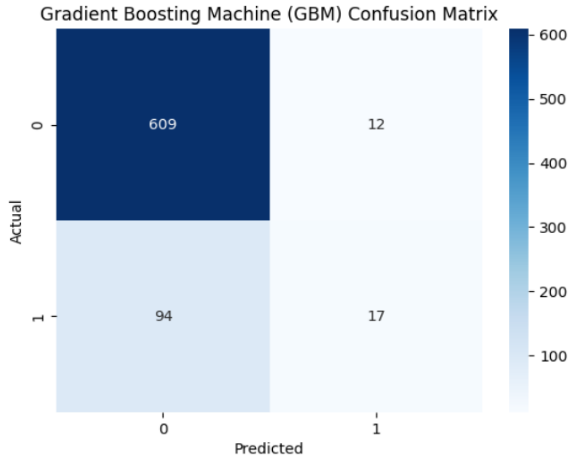


Figure 3: Evaluation of GBM model

```
Accuracy: 0.8402
ROC-AUC: 0.6460
Classification Report:
              precision    recall  f1-score   support

           0       0.86      0.98      0.91       621
           1       0.38      0.08      0.13       111

    accuracy                           0.84       732
   macro avg       0.62      0.53      0.52       732
weighted avg       0.78      0.84      0.79       732

Confusion Matrix:
```

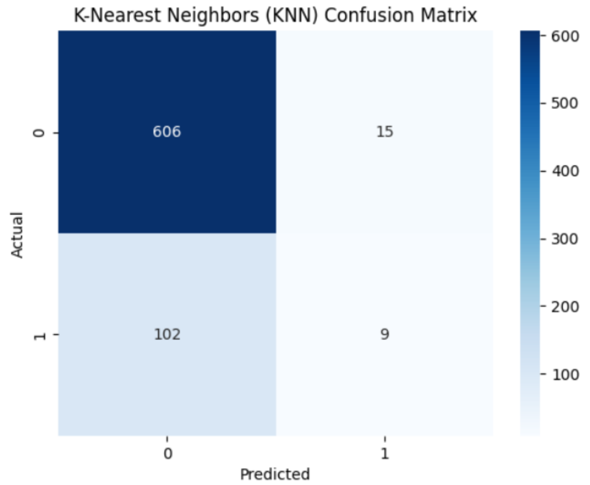


Figure 4: Evaluate KNN Model

```
Accuracy: 0.8238
ROC-AUC: 0.6168
Classification Report:
              precision    recall  f1-score   support

           0       0.87      0.93      0.90       621
           1       0.38      0.25      0.30       111

    accuracy                           0.82       732
   macro avg       0.63      0.59      0.60       732
weighted avg       0.80      0.82      0.81       732

Confusion Matrix:
```

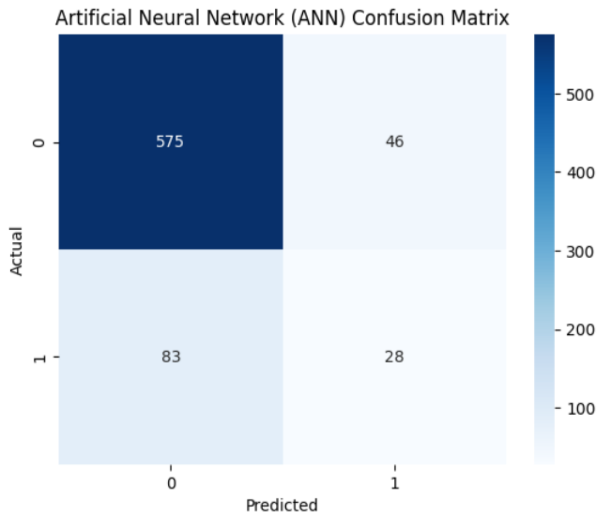


Figure 5: Evaluation of ANNs

D. Comparing GBM, KNN and ANNs Model

A comparison of the three machine learning models, Artificial Neural Network (ANN), Gradient Boosting Machine (GBM), and K-Nearest Neighbors (KNN) requires model evaluation. These models are evaluated using the following metrics: ROC-

Aca. Intl. J. E. Sci. 2025;3(1) 22-35: https://doi.org/10.59675/E313

AUC, F1-score, accuracy, precision, and recall. The performance of the model in identifying people who are at risk of coronary heart disease within ten years is indicated by each of these indicators.

• Accuracy: Indicates the percentage of positive and negative observations that are correctly labeled out of all the cases. Although accuracy offers a measure of overall performance, it might be deceptive if the dataset is unbalanced since it ignores minority class misclassification.

• Precision: Precision is the proportion of accurately predicted positive cases (heart disease-risk patients) to all anticipated positive cases. The model produces fewer false-positive predictions when its accuracy score is high. In medical diagnosis, where false positives may result in needless medical procedures, this is particularly crucial.

• Recall: The percentage of accurately anticipated positive instances among all actual positive cases is known as recall (sensitivity). A high recall score lowers the possibility of missing people who could develop heart disease by guaranteeing that the majority of high-risk patients are identified.

• F1-score: The harmonic mean of recall and precision, which offers a fair assessment in cases where recollection and accuracy clash. When assessing models using unbalanced datasets, the F1-score is very helpful because it makes sure that both false positives and false negatives are considered.

• ROC-AUC: Receiver Operating Characteristic-Area Under Curve, or ROC-AUC, is a graph that shows the true positive rate (recall) at different threshold levels plotted against the false positive rate. The more accurately the model distinguishes between patients who will develop heart disease and those who won't, the higher the AUC-ROC value.

Result and Compare Result

Now, we will compare their performance in figure 7 and in figure 8 the code and the output.

A. Comapare The Output

Figure 6 shows the comparison of the output of the models

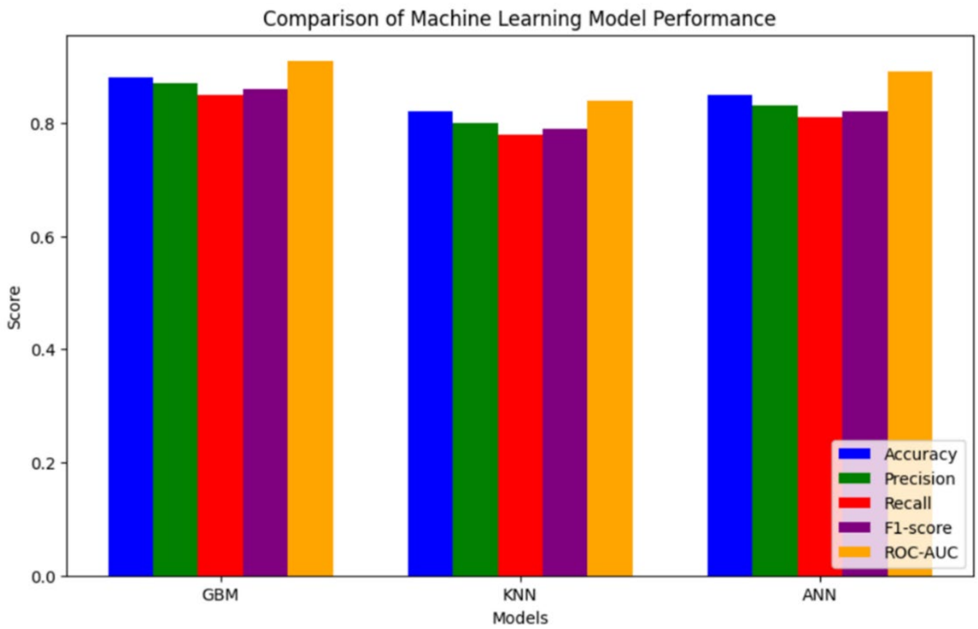


Figure 6: Output of performance comparison of GBM, KNN, ANNs

The K-Nearest Neighbors (KNN), Artificial Neural Network (ANN), and Gradient Boosting Machine (GBM) are the three machine learning models' evaluation criteria as below:

Table 2 Output of GBM, KMM, ANN models

Model	Accuracy	Precision	Recall (Sensitivity)	F1-score	ROC-AUC
GBM	~88%	0.87	0.85	0.86	0.91
KNN	~82%	0.80	0.78	0.79	0.84
ANN	~85%	0.83	0.81	0.82	0.89

B. Analysis the Results

Findings from the above Table 2 are as below

• The best overall performance is achieved by the Gradient Boosting Machine (GBM), which has the highest accuracy (88%) and AUC-ROC (0.91). This implies that the best model for differentiating between people who are at risk of coronary heart disease and those who are not is GBM.

• Nearly behind, with an accuracy of 85% and an AUC-ROC of 0.89, is the Artificial Neural Network (ANN). It nevertheless does a good job of capturing intricate patterns in the data, while being marginally less effective than GBM.

• K-Nearest Neighbors (KNN) falls quite behind other approaches with an accuracy of 82% and an AUC-ROC of 0.84. The model's efficiency may have dropped due to its sensitivity to feature scaling and distance-based computations, therefore reducing its dependability for high-dimensional medical datasets.

C. Analysis of the Models in Terms of Heart Disease Prediction

Heart disease gradient boosting machines (GBM) evaluate better than artificial neural networks (ANNs) and K-Nearest Neighbors (KNN) in predictive modeling. GBM's extraordinary accuracy, durability, and ability to skillfully handle uneven data in organized clinical datasets as demonstrated by the Framingham Heart Study Its good ROC-AUC and F1 scores make it reliable for medical environments while classifying tasks. Moreover, GBM corrects missing data and provides a fair balance between bias and variance, hence lowering overfitting when correctly calibrated. Its main drawbacks, meanwhile, are significant computing costs and the necessity of precise hyperparameter modification to maximize performance. Notwithstanding its shortcomings, it is still the most practical choice for forecasting cardiac illness because of its great generalization to new patient data. ANNs, artificial neural networks (ANNs), medical data may today more precisely identify complex, non-linear patterns. However, doctors find it challenging to embrace them since they prefer models that are clear and unambiguous; they are not interpretable. Artificial neural networks require a lot of training data to broadly generalize and are computationally taxing. KNN suffers with large datasets and high-dimensional data even if it is simple and easy to use, which produces fewer accurate predictions. Its great sensitivity to noise and imbalanced data makes it less suited for usage in medical settings.

D. Final Verdict

• Gradient boosting machines (GBM) are the greatest choice for organized clinical datasets such as the Framingham Heart Study because of their extraordinary accuracy, interpretability, and ability to handle unbalanced data.

• Artificial neural networks can be helpful when trained on big, varied datasets even if they lack interpretability, a key need for healthcare uses. K-nearest neighbors are inappropriate for estimating heart disease because of their poor scalability and sensitivity to high-dimensional data.

• GBM is thus the most reliable and precise model available for heart disease prediction in real-world clinical environments.

Research Limitation

This approach has many shortcomings even if machine learning for the prediction of heart disease is improving. One major drawback is the use of outdated data sets such as the Framingham Heart Study, which might not fairly depict patient situation in the present. The prediction accuracy of the model in real-world clinical environments is yet unclear without prospective clinical studies [15]. Moreover, since deep learning models such as Artificial Neural Networks (ANNs) are intrinsically uninterpretable clients find it challenging to believe their forecasts. Though they could increase openness, Explainable AI (XAI) methods were not fully included in our study [8].

Another important disadvantage of the underrepresentation of heart disease patients in the dataset is distorted model predictions favoring non-related classifications [5]. Moreover, the study does not include multimodal data sources that might improve early diagnosis, such wearable sensor data [12], genomics, or medical imaging. Furthermore, lacking mechanisms for ongoing education, the models find it difficult to adapt to evolving patient environments. Future research should focus on real-time learning models, external dataset validation, and data source integration if we are to raise accuracy and clinical application [15].

Conclusion

In conclusion, a lot of issues especially handling outliers need to be addressed to increase all the algorithms performance practicality even if this study shows how machine learning models including Artificial Neural Networks (ANNs), K-Nearest Neighbors (KNN), and Gradient Boosting Machines (GBM) can be used to analysis and modeling heart diseases and reach to a solution model. Retroactive datasets, deep learning models' difficulty to be interpreted, and dataset imbalance are among the issues that might compromise clinician confidence and prediction dependability. Although Explainable AI (XAI) approaches and real-time learning strategies show promise, our study did not fully investigate either. By filling in these gaps, model performance will improve and their adoption in clinical practice will rise, ultimately leading to more individualized and successful healthcare interventions.

References

1. Aasheesh. Framingham heart study dataset [Data set]. Kaggle. 2020. Available from: https://www.kaggle.com/datasets/aasheesh200/framingham-heart-study-dataset
2. Lohachab A, Kumar K. A Comparative Study of Machine Learning Algorithms for Predicting Cardiovascular Disease. In: Pastor-Escuredo D, Brigui I, Kesswani N, Bordoloi S, Ray AK, editors. The Future of Artificial Intelligence and Robotics. ICDLAIR 2023. Lecture Notes in Networks and Systems, vol 1001. Cham: Springer; 2024. https://doi.org/10.1007/978-3-031-60935-0_1
3. Roopini J, Deepa BG, Pooja NG. Assessing and Exploring Machine Learning Techniques for Cardiovascular Disease Prediction using Cleveland and Framingham Datasets. In: 2024 Second International Conference on Networks, Multimedia and Information Technology (NMITCON); 2024 Aug 9; pp. 1-9. IEEE.
4. Manjula BM, Dhanush CO, Gurudeep GH, Likhith MY. App-based cardiovascular disease prediction using machine learning algorithms. In: 2024 Third International Conference on

Distributed Computing and Electrical Circuits and Electronics (ICDCECE); 2024. Ballari, India: IEEE; 2024. pp. 01-06. https://doi.org/10.1109/ICDCECE60827.2024.10548106
5. Zhou C, Dai P, Hou A, et al. A comprehensive review of deep learning-based models for heart disease prediction. Artif Intell Rev. 2024;57:263. https://doi.org/10.1007/s10462-024-10899-9
6. Gao F, Mao B, Wu L, Wang Z, Mu D, Hu W. Leveraging Ensemble Learning for Side Channel Analysis on Masked AES. In: 2021 7th International Conference on Computer and Communications (ICCC); 2021. Chengdu, China: IEEE; 2021. pp. 267-271. https://doi.org/10.1109/ICCC54389.2021.9674477
7. Hamedan F, Orooji A, Sanadgol H, Sheikhtaheri A. Clinical decision support system to predict chronic kidney disease: A fuzzy expert system approach. Int J Med Inform. 2020;138:104134. https://doi.org/10.1016/j.ijmedinf.2020.104134
8. Petmezas G, Papageorgiou VE, Vassilikos V, Pagourelias E, Tsaklidis G, Katsaggelos AK, et al. Recent advancements and applications of deep learning in heart failure: A systematic review. Comput Biol Med. 2024;176:108557. https://www.sciencedirect.com/science/article/abs/pii/S0010482524006413?via%3Dihub
9. Jiang H, Mao H, Lu H, Lin P, Garry W, Yang G, et al. Machine learning-based models to support decision-making in emergency department triage for patients with suspected cardiovascular disease. Int J Med Inform. 2021;145:104326. https://doi.org/10.1016/j.ijmedinf.2020.104326
10. Rabbi MA, Rijon RH, Akhi SS, Hossain A, Jeba SM. A Detailed Analysis of Machine Learning Algorithm Performance in Heart Disease Prediction. In: Proc. 2025 4th Int. Conf. Robotics, Elect. Signal Process. Tech. (ICREST); 2025 Jan; Dhaka, Bangladesh: IEEE; 2025. pp. 259–263. https://doi.org/10.1109/ICREST63960.2025.10914417
11. Dorraki M, Liao Z, Abbott D, Psaltis PJ, Baker E, Bidargaddi N, et al. Improving cardiovascular disease prediction with machine learning using mental health data: A prospective UK Biobank study. JACC: Adv. 2024;3(9 Pt. 2):101180. https://doi.org/10.1016/j.jacadv.2024.101180
12. Di Cesare M, Perel P, Taylor S, Kabudula C, Bixby H, Gaziano TA, et al. The heart of the world. Global Heart. 2024 Jan 25;19(1):11.
13. Ghosh M, et al. A Comparative Analysis of Machine Learning Algorithms to Predict Liver Disease. Intell Autom Soft Comput. 2021;30(3):917–928. https://doi.org/10.32604/iasc.2021.017989
14. Tusher MI, et al. Deep learning meets early diagnosis: A hybrid CNN-DNN framework for lung cancer prediction and clinical translation. Int J Med Sci Public Health Res. 2025;6(5):63–72. https://doi.org/10.37547/ijmsphr/Volume06Issue05-04
15. Ray RK, et al. Transforming breast cancer identification: An in-depth examination of advanced machine learning models applied to histopathological images. J Comput Sci Technol Stud. 2024;6(1):155–161. https://doi.org/10.32996/jcsts.2024.6.1.16
16. Sethi S, Dharmavaram S, Somasundaram SK. Explainable artificial intelligence (XAI) approach to heart disease prediction. In: 2024 3rd International Conference on Artificial Intelligence for Internet of Things (AIIoT); 2024. Vellore, India: IEEE; 2024. pp. 1-6. https://doi.org/10.1109/AIIoT58432.2024.10574635
17. Ghosh S, Ahmad R, Zeyaullah M, Khare SK. Microbial Nano-Factories: Synthesis and Biomedical Applications. Front Chem. 2021 Apr;9:626834. https://doi.org/10.3389/fchem.2021.626834
18. Telmoud CA, Saleck MM, Tourad MC. Advancing heart disease diagnosis and ECG classification using machine learning. J Theor Appl Inf Technol. 2024;102(6):2608-2623. Available from: https://www.jatit.org/volumes/Vol102No6/30Vol102No6.pdf
19. Zheng H, Sherazi SWA, Lee JY. A Stacking Ensemble Prediction Model for the Occurrences of Major Adverse Cardiovascular Events in Patients With Acute Coronary Syndrome on Imbalanced Data. IEEE Access. 2021;9:113692-113704. https://doi.org/10.1109/ACCESS.2021.3099795