

## Clinical Support System for Cardiovascular Disease Forecasting Using ECG

Mohammed Altaf Ahmed<sup>1</sup>, Q S Tameem Naz<sup>2</sup>, Raghav Agarwal<sup>\*,3</sup>, Mannava Yesubabu<sup>4</sup> and Rajesh Tulasi<sup>5</sup>

<sup>1</sup>Department of Computer Engineering, College of Computer Engineering & Sciences, Prince Sattam Bin Abdulaziz University, Alkharj-11942, Saudi Arabia

<sup>2</sup>Shadan College of Engineering & Technology, JNTU, Hyderabad, India,

<sup>3</sup>School of Computer Science & Engineering (SCOPE), VIT-AP University, Amaravati, Andhra Pradesh, India

<sup>4</sup>Department of Computer Science & Engineering, Vardhaman College of Engineering, Hyderabad, India

<sup>5</sup>Department of Computer Science and Engineering, Koneru Lakshmaiah Education Foundation, Vaddeswaram, Guntur Dist., Andhra Pradesh - 522302, India

### Abstract

**INTRODUCTION:** Heart failure is a chronic condition that affects many people worldwide. Regrettably, it is now the biggest cause of mortality globally, and it is becoming more common. Before a cardiac event, early diagnosis of heart disease is challenging. Although healthcare institutions like hospitals and clinics have access to a wealth of heart disease data, it is rarely used to uncover underlying trends.

**OBJECTIVES:** Algorithms for machine learning (ML) can turn this medical data into insightful information. These methods are used to create decision support systems (DSS) that can gain knowledge from the past and advance. It is essential to use an effective ML-based technique to identify early heart failure and take preventive action to address this worldwide issue. Accurately identifying heart illness is our main goal in this study.

**METHODS:** For this work, we benchmark different datasets on heart illness, and we use feature engineering approaches to pick the most pertinent qualities for improved performance. Additionally, we assess nine ML methods using critical parameters including precision, f-measure, sensitivity, specificity, and accuracy.

**RESULTS:** Iterative tests are carried out to evaluate the efficacy of different algorithms. With a flawless cross-validation accuracy score of 99.51% and 100% in all other metrics, our suggested Decision Tree approach performs better than other ML models and cutting-edge studies.

**CONCLUSION:** Each methodology used in our study is validated using cross-validation techniques. The medical community benefits greatly from this research study.

**Keywords:** Machine learning, heart failure, decision support system, ensemble classifiers, cross-validation

Received on 20 December 2023, accepted on 11 March 2024, published on 18 March 2024

Copyright © 2024 M. A. Ahmed *et al.*, licensed to EAI. This is an open access article distributed under the terms of the [CC BY-NC-SA 4.0](#), which permits copying, redistributing, remixing, transformation, and building upon the material in any medium so long as the original work is properly cited.

doi: 10.4108/eetpht.10.5455

\*Corresponding author. Email: [raghav.g2106@gmail.com](mailto:raghav.g2106@gmail.com)

### 1. Introduction

A medical illness known as heart failure is characterized by the heart's inability to sufficiently pump blood to fulfil the demands of the body [1]. The most prevalent cause of mortality in the world is cardiovascular disease, sometimes

known as heart disease, and its prevalence is rising. It poses a serious threat to the health of populations on a massive scale, endangering global public health. A common and dangerous condition affecting millions of people worldwide is heart failure. According to recent estimates, there are 26 million people who have heart failure [2]. Heart failure has two primary categories of causes: those connected to the

structure of the heart, such as a prior heart attack, and those connected to the function of the heart, like excessive blood pressure. Heart disease can cause several symptoms, such as weariness, perspiration, and chest discomfort, yet many people don't show any signs until they have a heart attack [3]. The absence of symptoms in many patients until they have a sudden cardiac arrest has made it difficult for medical practitioners to diagnose heart disease [4], [5].

To identify heart disease in patients, medical practitioners use several medical tests, physical examinations, medical history analyses, and other diagnostic procedures. Usually, clinical history, a chest x-ray, and a physical examination are used to evaluate suspected heart disease cases. However, specific signs and particular symptoms might not always match up with these scientific methods. The fact that many people who suffered heart attacks and strokes were not considered to be "at risk" by doctors and experts must be emphasized. In actuality, medical professionals misdiagnosed about one-third of cases.

Medical experts have investigated a variety of disease management strategies, including nursing-based therapies [6] and technological interventions [7], to enhance the health outcomes of patients with heart disease. These methods' overall success is, however, constrained by the difficulties and high costs associated with their implementation.

The use of the Cardiac Biomarkers test provides a practical means of assisting in the early diagnosis of heart disease. Many doctors decide to examine the biomarkers found in the blood of patients with heart disease. Proteins, enzymes, and hormones are all examples of cardiac biomarkers. Several indicators become visible in the circulation when a patient's heart does not get enough oxygen, signaling a probable cardiac event.

To correctly identify heart illness, however, further tests could be necessary if the interpretation of biomarkers is murky in some cases. Additionally, because it implies a linear link between each risk factor and the onset of heart disease, the present risk assessment model is judged insufficient. Risk factors for heart disease include things like smoking, diabetes, high cholesterol, and hypertension. Because of this, healthcare professionals require a high level of competence to evaluate people appropriately, and the intricacy of some situations may make it more difficult to diagnose cardiac disease [8].

Within clinics, hospitals, and other healthcare facilities, there is a lot of medical data available. However, it is currently ineffective to use this data to derive useful insights and information. By using efficient ML algorithms for illness detection, it is possible to eliminate such medical mistakes [9] significantly.

In the field of healthcare and medical diagnostics, ML is essential [10]. ML algorithms may find trends and carry out prediction assessments by using massive medical information. ML has several benefits over conventional medical methods, including a reduction in time and expense and better diagnosis. These cutting-edge computer algorithms excel at differentiating between persons with heart illness and those who are healthy, offering distinct distinctions. Advanced ML techniques are used by clinical decision

support systems (CDSS) to help identify cardiac problems. It's crucial to keep in mind that CDSS depends on solid data-driven knowledge to allow precise analysis. The information required for illness prediction is provided by these DSS [11]. Implementing ML-based CDSS can improve outcomes by increasing the precision with which cardiac diseases are identified in the medical community.

The development of an ML model for the treatment of heart failure and the improvement of patient health is the main goal of this research. Our work significantly improves heart failure detection by applying ML algorithms, emphasizing the following important factors:

- **Machine Learning Models:** We compare the ability to predict heart failure using Nine Advanced ML Models with selected features.
- **Dataset Evaluation:** Using a dataset composed of four recognized heart illness datasets, we evaluate the efficacy of our suggested strategy.
- **Performance Validation:** Using the k-fold cross-validation approach, the performance of the ML models used is confirmed.

There are different sections in the article. The present research on heart failure is covered in depth in Section II. Section III describes the study methodology and process. To show the research's scientific validity, Section IV presents experimental data and analysis. Section VI concludes the study piece towards the end with future remarks.

## 2. Literature Review

This section offers a thorough examination of studies that explicitly use ML and deep learning methods to develop clinical DSS for the early diagnosis of heart disease.

By combining the Oppositional Firefly with BAT (OFBAT) and Rule-Based Fuzzy Logic (RBFL) methods, a hybrid system for heart disease prediction was created [12]. To create a classification model based on fuzzy logic, the Locality Preserving Projection (LRP) approach was used for feature selection. To find pertinent rules based on fuzzy rules generated from sample data, the OFBAT approach was used. Then, a fuzzy categorization system was built using these fuzzy rules and membership functions. Utilizing publicly accessible UCI datasets from Switzerland, Cleveland, and Hungary, the system's performance was assessed. According to the findings, RBFL has a 78% accuracy rate for the UCI datasets. Deep learning approaches, however, have the potential to enhance the suggested process.

[13] suggested a method for identifying cardiovascular illness. The study tested several ML techniques using a publicly accessible UCI dataset with 303 items. The authors separated their research into six independent cases to examine various facets of the categorization process. Classifiers were assessed in the first case without any feature reduction. The second example included using feature reduction techniques to narrow down the initial 14 features to only seven crucial ones. In the final illustration, the accuracy was assessed by omitting generic traits like gender, age, and resting blood

pressure. The fourth instance entailed utilizing the Weka tool to resample the datasets and calculate accuracy based on the seven crucial attributes. Re-sampling was used with all 14 features in the fifth illustration. The sixth example was the last to compute accuracy using the Weka tool and SMOTE. The KNN algorithm produced the most encouraging results of the studied methods, attaining an accuracy of over 80%. However, it's vital to remember that the suggested method has limits because it only takes into account the integration of one dataset.

[14] offered a DSS for aiding medical professionals in the diagnosis of cardiac disease. The Learning Vector Quantization (LVQ) approach was used in the suggested model as the primary strategy for prediction and analysis, and it made use of an existing dataset to assist in the diagnostic process. 13 pertinent characteristics were first chosen from the dataset. Artificial neural networks were then used to classify heart illness based on data that was explicitly trained on these chosen criteria. This work led to the creation of a simple method for identifying heart disease. ROC curve analysis was used to assess the performance of the proposed model, and the results showed an accuracy of 80%. The model may also benefit from text mining techniques to significantly improve its performance if textual data were provided. To make better predictions and manage unstructured data, text mining has the potential to handle heart illness diagnostics.

The Nave Bayes classifier was used in their study [15] to improve the prediction accuracy for the early identification of heart disease. Using discretization and genetic search methods, redundant and unwanted characteristics were removed from the dataset to make it more efficient. The least-rated qualities from the dataset were eliminated using the Genetic Algorithm (GA), which was used to select features. Next, the effectiveness of the Nave Bayes classifier was contrasted with that of other methods. The accuracy of the suggested method for the heart disease dataset was 86.29%, demonstrating its efficacy in early identification.

The researchers sought to raise the previously recorded 87.27% degree of heart failure detection accuracy in their study [16]. For the diagnosis of heart failure in patients, earlier studies used techniques including SVM, random forest, decision tree, logistic regression, and naive Bayes classifier. The decision tree approach had an accuracy of 93.19% when the results were compared, demonstrating a high degree of detection for heart failure in the particular dataset under examination.

Additionally, it is essential to fill in the study gaps that have been found in earlier studies, which indicate that performance accuracy has not yet peaked. To close this gap, we thoroughly assess the performance analysis of the prior study in this part. The data are summarized in this part, and the efficacy of the preceding models is evaluated. Previous studies have shown that various model types still produce variable prediction scores, highlighting the need for more study and advancement in this area.

### 3. Methodology

The proposed methodology is staged into five steps and the proposed framework is shown in Figure 1.

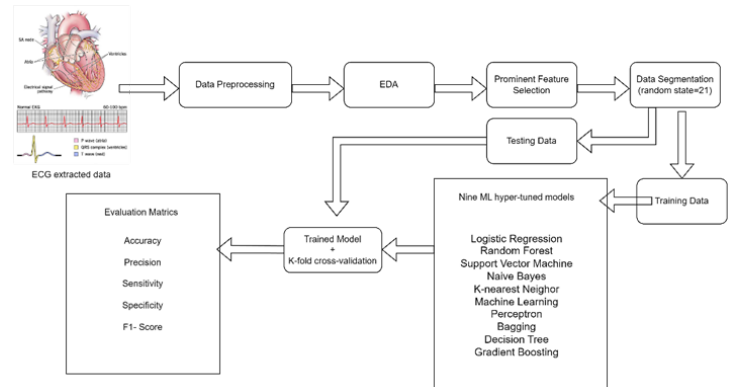


Figure 1. Proposed Framework of the Methodology

#### 3.1. Data Acquisition and Overview

The Kaggle repository was utilized to get the heart failure dataset that was used in this work [17]. The collection consists of 1025 patient records from four different databases: Long Beach, Switzerland, Hungary, and Cleveland. These records, comprising 713 male samples and 312 female samples, provide data on both heart failure patients and healthy persons. The dataset initially had 14 characteristics. A summary of these characteristics is shown in Table 1. One characteristic, slope, was of the float data type, but 13 of the dataset's properties were of the integer data type, as we saw. To locate any missing values in the dataset, we also ran an analysis. The analysis's findings showed that there were no missing values, confirming the validity of the following investigations and assessments.

Table 1. Data characteristics of the features

| Column                  | Non-Null Count | Dtype   |
|-------------------------|----------------|---------|
| age                     | 1025 non-null  | int64   |
| sex                     | 1025 non-null  | int64   |
| chest_pain_type         | 1025 non-null  | int64   |
| resting_blood_pressure  | 1025 non-null  | int64   |
| cholesterol             | 1025 non-null  | int64   |
| fasting_blood_sugar     | 1025 non-null  | int64   |
| rest_ecg                | 1025 non-null  | int64   |
| max_heart_rate_achieved | 1025 non-null  | int64   |
| exercise_induced_angina | 1025 non-null  | int64   |
| st_depression           | 1025 non-null  | float64 |
| st_slope                | 1025 non-null  | int64   |
| num_major_vessels       | 1025 non-null  | int64   |
| thalassemia             | 1025 non-null  | int64   |
| target                  | 1025 non-null  | int64   |

#### 3.2. Data Acquisition and Overview

Several data preparation procedures were used to guarantee the consistency of the dataset. These preparation stages included addressing missing data by replacing them with

suitable values, carrying out normalization and standardization, finding and deleting outliers, and removing duplicate entries. The dataset was organized and standardized using normalization and standardization techniques for consistency reasons, notably by using the Standard Scaler preprocessing module from the Sklearn package. These steps were essential in getting the dataset ready for additional analysis and modeling.

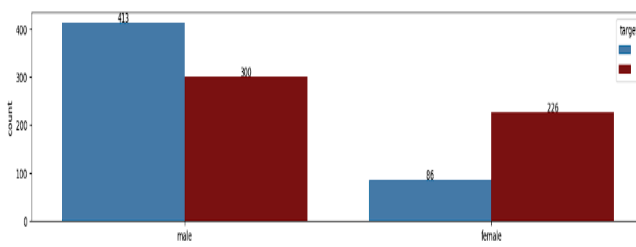


Figure 2. Dataset's distribution on sex based on the target variable

### 3.3. Exploratory Data Analysis

To comprehend the data patterns and causes affecting heart failure better, an exploratory data analysis (EDA) was carried out. The EDA's goal was to draw insightful facts and statistics from the research dataset. Charts and heatmap graphs were used as visual aids to show the patterns and relationships found in the dataset. A bar graph depicting the dataset's distribution based on the target variable is shown in Figure 2. It demonstrates that the sample consisted of 526 heart failure patients and 499 healthy people. Additionally, there are 226 female and 300 male cases of heart failure. The dataset is usually balanced, according to this study, which helps create ML models. Figure 3 shows the correlation analysis between the heart failure dataset's properties. The findings show that most traits have substantial correlations, with just a few factors showing strong negative correlations. These results shed light on the features of the dataset and the interrelationships of its variables.

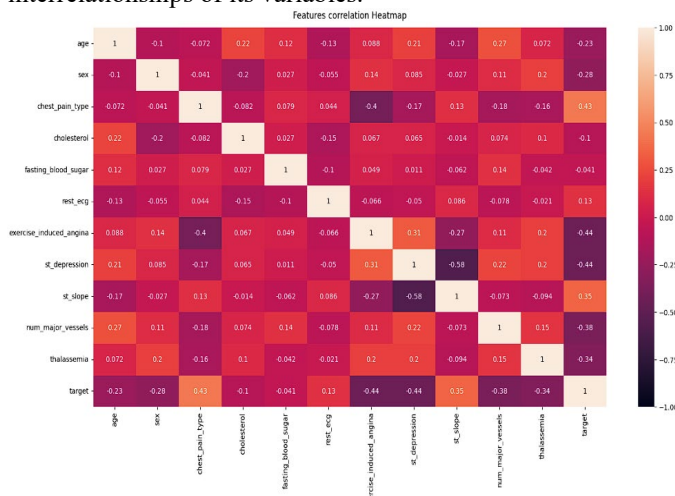


Figure 3. Feature Correlation Heatmap

### 3.4. Feature Selection and Data Splitting

To improve accuracy and optimize the dataset, a revised feature set was created, yielding a dataset with twelve carefully chosen properties. These recently created characteristics demonstrated astounding accuracy rates when it came to heart failure prediction. We dropped the resting\_blood\_pressure and max\_heart\_rate\_achieved columns for our new featured dataset.

The dataset was split into training and testing sets to guarantee consistent model performance and reduce overfitting. By using this partitioning technique, ML classifiers may be trained on a portion of the data while having their performance assessed on test data that has not yet been viewed. The heart failure dataset used in this work was divided into two parts: training, and testing. With the help of this method, classifiers can accurately anticipate the outcomes of new data. The classifiers used in this study are famous for their outstanding accuracy, producing results that are dependable and solid.

### 3.5. Applied Machine Learning Techniques

In this part, we look at different ML techniques that are used to forecast cardiac failure. We clarify the key ideas and operational procedures that serve as these frameworks' cornerstones. Our research focuses on evaluating the performance of 9 reliable ML models created particularly for heart failure prediction.

**Logistic Regression (LR).** A popular supervised ML technique for regression and classification applications is LR. It uses probability to predict the labels of categorical data. To aid learning and prediction, LR calculates the probability of binary classifications. The target variables in our study dataset must be binarily categorized for LR to be performed, and in our instance, they do so by having two unique binary values. In the dataset, 1 denotes potential heart patients, whereas 0 denotes those who have no chance of getting heart failure.

$$Prob(target = 1 | inputs) = \frac{1}{1 + e^{-z}}$$

$$where z = \alpha_0 + \alpha_1x_1 + \dots + \alpha_nx_n \tag{1}$$

The probability that the binary outcome variable, y, will have a value of 1, given the predictor variable(s) denoted as "inputs," is represented by the formula "Prob(target = 1 | inputs)". Any real integer, z, is transformed into the range between 0 and 1 using the logistic function, abbreviated as  $e^{-z}$ , making sure the resultant value accurately represents a probability.

**Decision Tree (DT).** A DT is frequently used for data pattern recognition and data cleaning activities. Some features in a dataset with many features may be more or less important than others. The Gini Index is a useful tool for locating and removing superfluous or tangentially related variables. Every node of a DT is chosen based on the calculation of conditional probabilities. The selection of characteristics with lower Gini Index values and the



development of rules based on these features. DTs are used in the field of bioinformatics for illness diagnosis and prediction. The formula listed below is used to calculate the Gini Index.

$$Gini - ind = \sum_{k=1}^j prob(k) * (1 - prob(k)) \quad (2)$$

The variable 'j' in the above formula stands for the class number, and 'prob(k)' stands for the probability attached to each class.

**Random Forest (RF).** A popular supervised ML technique for solving classification and regression issues is called RF. It makes use of a collection of DTs to handle difficult problems and improve efficiency. The outputs of several DTs trained on various subsets of the dataset are averaged by the RF classifier to increase predicted accuracy. The required degree of precision dictates the ideal number of trees in an RF model. Overfitting is successfully reduced by RF, which also generates accurate forecasts with good performance.

$$m = \frac{1}{n} \sum_{p=1}^n f_p(o) \quad (3)$$

The outputs of 'n' DTs are averaged to create predictions inside the RF framework. Here, "m" stands for the expected result for a certain input vector "o." These outputs, indicated as 'f<sub>p</sub>(o)', come from the various predictions made by each DT and are averaged to provide the final forecast.

**Support Vector Machine (SVM).** A well-liked supervised learning method for data categorization and prediction is the SVM. Finding the best choice limits is its main goal. SVM does this by using a hyperplane, an ideal decision boundary, to divide the n-dimensional space. The hyperplane may be used by SVM to precisely assign fresh data points to their appropriate categories. It is called an SVM because choosing extreme support vectors is a necessary step in building the hyperplane.

$$Weight * input + constant = 0 \quad (4)$$

The weight vector connected to the SVM model is referred to as "Weight" in this context. The input vector is referred to as the "input" in the SVM formulation, and the "constant" is either the bias term or the intercept term.

**Naïve Bayes (NB).** A technique for supervised ML used to address classification issues is the NB model. It is a straightforward yet efficient technique for producing precise forecasts. Using probabilistic concepts, the NB approach enables the classifier to produce predictions based on the probability of variables in the dataset. To predict the target class for each data record, NB makes use of the values of independent variables. The NB model has several uses in the field of medicine.

$$= \frac{Prob(Class_x | input) * Prob(Class_x)}{Prob(input)} \quad (5)$$

The probability that a sample will belong to class x given the input attributes in the current situation is represented by the expression "Prob(Class<sub>x</sub> | input)". The chance of seeing

the input characteristics given the class x is denoted by the term "P(input | Class<sub>x</sub>)" on the other hand.

**K-Nearest Neighbor (KNN).** A supervised ML technique called KNN is used for classification and regression tasks. Since the KNN approach is non-parametric, no assumptions are made on the distribution of the data. Instead, it assigns a new instance to a class depending on how similar it is to instances of that class already in existence. In KNN, the Euclidean distance metric is frequently employed to calculate similarity and extract useful data from nearby samples.

$$dist(l, m) = \sqrt{\sum_{k=1}^n (m_k - l_k)^2} \quad (6)$$

**Gradient Boost (GB).** A popular ensemble learning technique in ML is GB. It is a useful method for producing predictive analytics in scenarios including regression and classification. By merging predictions from many different learner models, GB creates a final predictive model that produces precise predictions. The fundamental goal of GB is to improve weak models iteratively in order to build a robust model. Each base classifier is trained individually while a model is being built progressively using the GB technique.

**Multilayer Perceptron (MLP).** A feed-forward artificial neural network called an MLP converts a collection of inputs into equivalent outputs. The input layer, hidden layer(s), and output layer are only a few of the layers that make up the MLP. The network's hidden layer processes the data once it is received by the input layer, which then produces the results. For training MLP networks, back-propagation is a popular learning method.

**Bagging.** An ML ensemble approach, known as Bootstrap Aggregation or Bagging, improves model performance by aggregating the results of many learners. It's frequently used to lower variance in noisy data. Multiple data samples are created, and separately trained using various ML models, and portions of the training data are selected at random. By combining the results of different models, the final forecast is achieved. Classification issues are frequently addressed by ensemble learning using DTs and NB. It can be applied to regression approaches, although classification issues benefit most from it. Using the above equation, ensemble learning calculations are made.

## 4. Experimental Evaluation & Measures

This part focuses on the results of our investigation and the validity of our conclusions. We used the Python computer language and the skit-learn library module to build the machine algorithms. We used several performance measures in our evaluation of the suggested framework's efficacy. Figure 4 represents the confusion matrix originated with the applied ML models and the corresponding tabulated evaluated matrix is shown in Table 2.

### 4.1. Evaluation Metrics

**ACCURACY.** The percentage of correct classifications produced by a classification model to the total number of evaluations performed on the test data is known as accuracy.

$$Accuracy = \frac{True\_Negatives + True\_Positives}{Total\_Positives + Total\_Negatives} \quad (7)$$

**PRECISION.** The percentage of accurate assessments determined by a classification model is known as the accuracy percentage. The following formula may be used to compute it:

$$Precision = \frac{True\_Positives}{Total\_Positives} \quad (8)$$

**SENSITIVITY.** It is crucial to correctly identify those who suffer from a certain condition. The ratio of genuine positive predictions to the total number of real positive samples may be used to measure this accuracy. The formula below may be used to calculate this ratio:

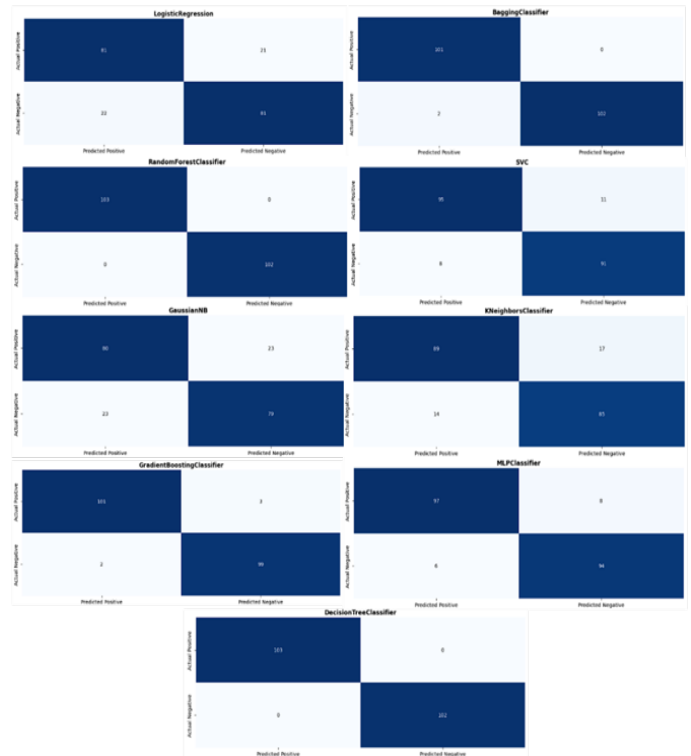
$$Sensitivity = \frac{True\_Positives}{True\_Positives + False\_Negatives} \quad (9)$$

**SPECIFICITY.** This metric's objective is to precisely pinpoint those who do not have the disease.

$$Specificity = \frac{True\_Negatives}{False\_Positives + True\_Negatives} \quad (10)$$

**F1 SCORE.** Precise class partitioning is the process of accurately classifying data into the proper classes while taking into account various classifications. The harmonic mean of recall and accuracy, which yields the F1 score, measures the proportion of erroneous positive and false negative classifications. This equation may be used to calculate it:

$$F1\_Score = \frac{2 \times (Precision + Sensitivity)}{Precision + Sensitivity} \quad (11)$$



**Figure 4.** Confusion Matrix originated with the applied ML models

### 4.2. K-fold Cross-Validation Comparative Analysis

We use a 10-fold cross-validation approach to divide each dataset into training and test sets to ensure accurate performance evaluation. According to this method, one subset from each fold is used for testing while the other nine are used for training. We determine the final accuracy score, which reflects how well the trained classifier performed, by averaging the accuracy over these folds. For each dataset and classifier used in our investigation, this process is repeated.

We use K-fold cross-validation with 10 folds on our ML models to combat the issue of overfitting. Notably, based on the cross-validation findings from our investigation, our proposed DT technique gets an amazing accuracy score of 99.51%. However, when put through cross-validation, the NB, LR, and KNN models perform poorly.

A comparison of the performance of the ML models used in testing and cross-validation situations is shown in Table 2. The results demonstrate our suggested DT model's superior and all-around performance. No matter whether the performance of the models was assessed using testing or cross-validation methods, this comparison analysis offers validation for their efficacy.

Table 2. Comparison analysis of applied ML algorithms

| Models    | Accuracy (%) | Precision (%) | Sensitivity (%) | Specificity (%) | F1 score (%) | Cross-validation |
|-----------|--------------|---------------|-----------------|-----------------|--------------|------------------|
| LR        | 79.02        | 79.41         | 78.64           | 79.41           | 79.02        | 83.2927          |
| <b>DT</b> | <b>100</b>   | <b>100</b>    | <b>100</b>      | <b>100</b>      | <b>100</b>   | <b>99.5122</b>   |
| RF        | 100          | 100           | 100             | 100             | 100          | 99.3902          |
| SVM       | 90.7317      | 89.62         | 92.23           | 89.22           | 90.91        | 91.5854          |
| NB        | 77.5610      | 77.67         | 77.67           | 77.45           | 77.67        | 82.3171          |
| KNN       | 84.8780      | 83.96         | 86.41           | 83.33           | 85.17        | 84.7561          |
| GB        | 97.56        | 97.12         | 98.06           | 97.06           | 97.58        | 97.5610          |
| MLP       | 93.17        | 92.38         | 94.17           | 92.16           | 93.27        | 94.1463          |
| Bagging   | 99.02        | 100           | 98.06           | 100             | 99.02        | 99.0244          |

### 4.3. Comparison with State-of-the-Art Studies

A thorough comparison of the outcomes of earlier research projects and our analysis of the dataset is provided in Table 3. The table contains information on each study, including its proposed strategy, and its accuracy metric. Our recommended DT model beats the previously offered techniques, according to the findings of our study. Our study yields the most encouraging results for heart failure prediction, especially by adding certain feature selection algorithms [25].

Table 3. Comparison Analysis with state-of-the-art models

| Reference | Proposed Strategy   | Accuracy               |
|-----------|---|------------------------|
| [18]      | ANN+SVM   | 99%                    |
| [19]      | Ensemble Techniques   | 90%                    |
| [20]      | Gaussian Naive Bayes+SVM+ RF+KNN +XGBoost   | 88.52%                 |
| [21]      | ANN + Logistic Regression   | 85%                    |
| [22]      | Hybrid CNN-GRU  | 94%                    |
| [23]      | Decision Tree   | 98%                    |
| [24]      | Decision Tree   | 98%                    |
| Proposed  | <b>Decision Tree with prominent feature selection<br/>Accuracy and all metrics<br/>Cross-validation</b> | <b>100%<br/>99.51%</b> |

## 5. Conclusion and future remarks

This study suggests using ML methods to anticipate cardiac failure. 1025 patient records make up the dataset utilized in this study, which was used to create the prediction models. We discover and pick the top twelve criteria to improve performance. We examine several ML techniques, including NB, KNNs, LR, RF, SVM, DT, KNN, MLP, Bagging, and GB [26]. Each learning method undergoes a ten-fold cross-validation to evaluate and validate the performance of the models. Notably, the cross-validation accuracy of our proposed DT technique is remarkable at 99.51%. This shows the generalizability of our suggested method in correctly detecting heart failure and outperforms the current state-of-the-art research.

Comparing the suggested model to earlier techniques, it performed better in terms of accuracy, sensitivity, and specificity. Future studies will include imaging information gathered from heart disease patient images to further improve this strategy. Laboratory tests and several imaging techniques will be used to get this imaging data. Additionally, we want to effectively diagnose heart illness using convolutional neural networks (CNNs) capabilities. The primary benefit of using CNNs with the supplied

picture data is their capacity to automatically identify the most important characteristics.

The PR curve and the ROC curve are two more performance indicators that we will use to analyze the model. These metrics offer a thorough analysis of the model's effectiveness at various thresholds. Additionally, the CNN model will be evaluated as part of our research utilizing both structured and unstructured data. Utilizing the potential advantages of merging structured and unstructured data, this extensive review intends to improve the model's performance and accuracy in forecasting heart illness.

### Acknowledgements.

This study is supported via funding from Prince Sattam bin Abdulaziz University project number (PSAU/2023/R/1444)

### References

- [1] M. Gjoreski, M. Simjanoska, A. Gradišek, A. Peterlin, M. Gams, and G. Poglajen, "Chronic heart failure detection from heart sounds using a stack of machine-learning classifiers," in Proc. Int. Conf. Intell. Environments (IE), Aug. 2017, pp. 14–19.
- [2] G. Savarese and L.H. Lund, "Global public health burden of heart failure," Cardiac Failure Rev., vol. 3, no. 1, p. 7, 2017.

- [3] N.-S. Tomov and S. Tomov, "On deep neural networks for detecting heart disease," 2018, arXiv:1808.07168.
- [4] C. Sowmiya and P. Sumitra, "Analytical study of heart disease diagnosis using classification techniques," in Proc. IEEE Int. Conf. Intell. Techn. Control, Optim. Signal Process. (INCOS), Mar. 2017, pp. 1–5.
- [5] A. Kumar, P. Kumar, A. Srivastava, V. D. A. Kumar, K. Vengatesan, and A. Singhal, "Comparative analysis of data mining techniques to predict heart disease for diabetic patients," in Proc. Int. Conf. Adv. Comput. Data Sci. Singapore: Springer, 2020, pp. 507–518.
- [6] Y.-J. Huang, M. Parry, Y. Zeng, Y. Luo, J. Yang, and G.-P. He, "Examination of a nurse-led community-based education and coaching intervention for coronary heart disease high-risk individuals in China," *Asian Nursing Res.*, vol. 11, no. 3, pp. 187–193, Sep. 2017.
- [7] S. D. Pande and R. Agarwal, "Multi-Class Kidney Abnormalities Detecting Novel System Through Computed Tomography," in IEEE Access, doi: 10.1109/ACCESS.2024.3351181.
- [8] S. Khan and S. T. Rasool, "Current use of cardiac biomarkers in various heart conditions," *Endocrine, Metabolic Immune Disorders-Drug Targets*, vol. 21, no. 6, pp. 980–993, Jun. 2021.
- [9] R. Agarwal, S. D. Pande, S. N. Mohanty and S. K. Panda, "A Novel Hybrid System of Detecting Brain Tumors in MRI," in IEEE Access, vol. 11, pp. 118372-118385, 2023, doi: 10.1109/ACCESS.2023.3326447.
- [10] A. Qayyum, J. Qadir, M. Bilal, and A. Al-Fuqaha, "Secure and robust machine learning for healthcare: A survey," *IEEE Rev. Biomed. Eng.*, vol. 14, pp. 156–180, 2021.
- [11] R. Agarwal and D. Godavarthi, "Skin Disease Classification Using CNN Algorithms", *EAI Endorsed Trans Perv Health Tech*, vol. 9, Oct. 2023.
- [12] G.T. Reddy and N. Khare, "An efficient system for heart disease prediction using hybrid OFBAT with rule-based fuzzy logic model," *J. Circuits, Syst. Comput.*, vol. 26, no. 4, Apr. 2017, Art. no. 1750061.
- [13] N. Khateeb and M. Usman, "Efficient heart disease prediction system using K-nearest neighbor classification technique," in Proc. Int. Conf. Big Data Internet Thing, Dec. 2017, pp. 21–26.
- [14] A. H. Chen, S. Y. Huang, P. S. Hong, C. H. Cheng, and E. J. Lin, "HDPS: Heart disease prediction system," in Proc. Comput. Cardiol., Sep. 2011, pp. 557–560.
- [15] M. A. Jabbar and S. Samreen, "Heart disease prediction system based on hidden naïve Bayes classifier," in Proc. Int. Conf. Circuits, Controls, Commun. Comput. (I4C), Oct. 2016, pp. 1–5.
- [16] F.S. Alotaibi, "Implementation of machine learning model to predict heart failure disease," *Int. J. Adv. Comput. Sci. Appl.*, vol. 10, no. 6, pp. 1–8, 2019.
- [17] Heart Disease Dataset|Kaggle, DAVID LAPP, Atlanta, Georgia, 1988.
- [18] Vijeta Sharma, Shrinkhala Yadav, Manjari Gupta, "Heart Disease Prediction using Machine Learning Techniques", 2020 2nd International Conference on Advances in Computing, Communication Control and Networking(ICACCCN), ©2020 IEEE, ICACCCN51052 .2020.9362842.
- [19] Rahma Atallah, Amjed Al-Mousa, "Heart Disease Detection Using Machine Learning Majority Voting Ensemble Method," 2019 2nd International Conference on new Trends in Computing Sciences (ICTCS).
- [20] Shaik Farzana, Duggineni Veeraiah, "Dynamic Heart Disease Prediction using Multi-Machine Learning Techniques", 2020 5th International Conference on Computing, Communication and Security (ICCCS).
- [21] Ghosh, S. (2017). Application of Various Data Mining Techniques To Classify Heart Diseases. [online].
- [22] A.A. Ali, H.S. Hassan, and E.M. Anwar, "Heart diseases diagnosis based on a novel convolution neural network and gate recurrent unit technique," in Proc. 12th Int. Conf. Electr. Eng. (ICEENG), Jul. 2020, pp. 145–150.
- [23] O. E. Taylor, P. S. Ezekiel, and F. B. D. Okuchaba, "A model to detect heart disease using machine learning algorithm," *Int. J. Comput. Sci. Eng.*, vol. 7, no. 11, pp. 1–5, Nov. 2019.
- [24] D.K. Chohan and D.C. Dobhal, "A comparison-based study of supervised machine learning algorithms for prediction of heart disease," in Proc. Int. Conf. Comput. Intell. Sustain. Eng. Solutions (CISES), May 2022, pp. 372–375.
- [25] R. Agarwal, A. S. Sathwik, D. Godavarthi, and J. V. Naga Ramesh, "Comparative Analysis of Deep Learning Models for Multiclass Alzheimer's Disease Classification", *EAI Endorsed Trans Perv Health Tech*, vol. 9, Nov. 2023.
- [26] R. Agarwal, J. Suthar, S. K. Panda, and S. N. Mohanty, "Fuzzy and Machine Learning based Multi-Criteria Decision Making for Selecting Electronics Product", *EAI Endorsed Scal Inf Syst*, vol. 10, no. 5, Jul. 2023.