



Available online at www.sciencedirect.com

ScienceDirect

Procedia Computer Science 235 (2024) 937–947

Procedia
Computer Science

www.elsevier.com/locate/procedia

International Conference on Machine Learning and Data Engineering (ICMLDE 2023)

Heart Disease Detection Using Machine Learning Models

Amrit Singh^a, Harisankar Mahapatra^b, Anil Kumar Biswal^{c*}, Madhumita Mahapatra^d, Debabrata Singh^e, Milan Samantaray^f

^{a,b}Department of Computer Science and Engineering, NIST Institute of Science and Technology (Autonomous), Berhampur, India

^cDept. of Computer Science, Udayan Autonomous College of Science and Technology, Cuttack, India

^dDept. of Computer Science and Engineering, Delhi Technical Campus, Greater Noida, UP, India

^eDepartment of Computer Application, Institute of Technical Education and Research (ITER), Siksha 'O' Anusandhan (Deemed to be University), Bhubaneswar, India

^fDepartment of Computer Science and Engineering, Silicon Institute of Technology, Bhubaneswar, India

Abstract

An enormous number of deaths occur every year as a result of heart disease, making it a major concern in world health. Improving patient outcomes and lowering death rates, early detection and correct diagnosis of cardiac disease play a key role. When the heart's arteries become blocked, oxygen-poor blood cannot reach the heart properly, resulting in coronary heart disease. Early detection of cardiac disease is viable because it reduces medical costs and potentially saves the patient's life. Recently presented methods have improved heart failure detection accuracy on testing data without sacrificing accuracy on training data, yet most of these algorithms are suffering from the issue of overfitting. Models that were created end up fitting the test data too well. In this study, we create a novel diagnostic system to address this issue, and the resulting system demonstrates high intelligence and excellent performance on both training and testing data. Machine learning (ML) algorithms have demonstrated promising potential in assisting healthcare professionals with timely and accurate diagnosis. In this paper, is based on supervised machine learning methods are decision tree (DT), random forest (RF), Support vector Machine (SVM), Principal Component Analysis(PCA). We compare their accuracy with each other by using bar plot.

© 2024 The Authors. Published by ELSEVIER B.V.

This is an open access article under the CC BY-NC-ND license (<https://creativecommons.org/licenses/by-nc-nd/4.0>)

Peer-review under responsibility of the scientific committee of the International Conference on Machine Learning and Data Engineering

Keywords: Coronary Heart Disease; Blood Pressure; Health Care; Machine Learning; Imbalanced Data; Decision Tree; Random Forest; SVM; PCA;

1. Introduction

Heart disease, encompassing conditions like coronary artery disease, heart failure, and arrhythmias, continues to be one of the primary causes of death worldwide. The World Heart Federation (WHF) has released a report estimating that the number of deaths caused by cardiovascular disease (CVD) over the world will increase from 20.5

million in 2021 to 12.1 million in 1990 [1].

* Corresponding author.
E-mail address: anil.biswal123@gmail.com

Early detection and precise diagnosis of cardiac disease improve patient outcomes and reduce mortality by offering timely intervention and care. Timely identification enables timely intervention, often preventing disease progression and complications. It empowers healthcare providers to tailor treatment plans, address risk factors, and enhance long-term disease management. Moreover, it contributes to cost savings and has a positive impact on public health by lowering the overall burden of cardiovascular morbidity and mortality. Traditionally, heart disease diagnosis heavily relied on clinical assessment, medical history, physical examinations, and invasive procedures. While these methods have proven to be valuable, they are often subjective, time-consuming, and may require specialized expertise [2, 3]. Furthermore, the complexity and diversity of heart disease patterns make accurate diagnosis a challenging task for healthcare professionals. Recent advances in ML and huge healthcare datasets have helped doctors detect and diagnose cardiac illness. ML algorithms automatically extract key information and predict based on data patterns and correlations. These algorithms can find heart disease trends in massive volumes of medical, imaging, and genetic data [4]. The objective of this paper is to provide a comprehensive review of the existing literature on heart disease detection using machine learning techniques. This study intends to contribute to the understanding and advancement of this important area of research by assessing and summarising the methodology, datasets, performance indicators, and obstacles faced in constructing reliable prediction models [5, 6].

In this paper, we will first present a review of relevant literature on heart disease detection using ML techniques. Here the dataset is Framingham. Various ML techniques employed in previous studies will be discussed, including both supervised [7]. We will also examine this dataset used for training and evaluation, considering the diversity and representativeness of the data [8]. Furthermore, we will delve into the data collection and pre-processing methods used in heart disease detection studies. Handling missing data, outliers, feature selection, and considerations for data augmentation and balancing class distribution will be explored, as these steps significantly impact the performance and generalizability of ML models [9, 10].

Performance evaluation plays a crucial role in the development of robust heart disease detection models. In our discussion, we will go through a number of criteria that are used to rate these models' efficacy. The area under the receiver operating characteristic curve (AUC-ROC) is one of these metrics, along with accuracy, confusion matrix, sensitivity, specificity, precision, recall, F1-score, and F-score. Additionally, we will explore cross-validation techniques and model selection strategies to ensure reliable and accurate predictions.

1.1 Motivation

The motivation behind this paper stems from the urgent need to improve heart disease detection and diagnosis using advanced technologies such as machine learning. Heart disease continues to be a major public health concern, causing substantial morbidity and mortality worldwide. Improving patient outcomes and decreasing the strain on healthcare systems are both greatly aided by early identification. However, accurately diagnosing heart disease remains a complex and challenging task due to the intricate nature of the disease and the variability in its manifestations.

In numerous fields, including healthcare, machine learning techniques have demonstrated tremendous promise. Machine learning algorithms can learn from enormous datasets, recognise complicated patterns, and generate accurate predictions, which could improve heart disease detection. Clinicians can benefit from automated decision-support systems that can promptly process diverse patient data and provide valuable insights for diagnosis and treatment planning by leveraging these techniques.

The remaining five portions are as follows. In Section 2, we give a comprehensive literature review of text localization techniques and their applications. In section 3, we briefly introduce the proposed system and dataset describe the employed; in part 4, we detail the processes of the methodologies. The experimental result analysis and conclusions are addressed in sections 5 and 6.

2. Literature Survey

Predicting Cardiovascular and Hepatic Disease with a Hybrid Machine Learning Algorithm Based on Particle Swarm Optimisation and Support Vector Machines, Behera et al. [11]. Techniques such as particle swarm optimization (PSO) and feature selection methods have also been explored to enhance classification accuracy. The proposed hybrid model aims to improve classification accuracy compared to existing approaches. Classification accuracy, error, correctness, recall, and F1 score are only few of the metrics used to evaluate the hybrid model's efficacy in this study, which compiles datasets from the UCI Machine Learning Repository.

Machine learning classifier assembly and feature engineering based on entropy for cardiovascular disease prediction, Rajkamal Rajendran et al. [12]. The experimental results show that the suggested ensemble model (logistic regression + naive Bayes) performs better than the state-of-the-art results when using the new pipeline. AUC, accuracy, specificity, precision, and F1 score are just few of the measures that show the ensemble model's ability for accurate identification of cardiac disorders.

Systematic literature evaluation of machine learning for diagnosing cardiovascular illness by Md Manjurul Ahsan et al. [13]. This research investigates how machine learning algorithms fare when given an imbalanced dataset with a dominant class. It highlights the significance of trend, technique, gap, and opportunity identification in ML-based cardiac illness detection.

Predictive analysis of heart diseases with ML approaches describe by Ramesh TR et al. [14]. The process of training and testing ML models is explained, along with the use of high-dimensional data and the curse of dimensionality. Multiple ML methods are utilized in the classification and prediction of cardiovascular diseases with the objective of identifying high-risk individuals. Factors such as blood pressure, diabetes, cholesterol levels, obesity, tobacco use, and family history are mentioned as potential contributors to heart disease.

A novel Stacking Ensemble Learner-based machine learning model for predicting cardiac patients' need for readmission to the hospital in an emergency, Alireza Ghasemieh et al. [15]. The model used in this investigation is an XgBoost-based stacking ensemble learner (SEL) that employs ensemble learning approaches. The objective is to achieve the highest possible level of detection efficiency. The study's contributions include a unique classification scheme for emergency readmission based on behavioural traits, the use of a large, confidential dataset from the MIT Laboratory for Computational Physiology, and the design of the SEL model. In order to classify heart sounds, a ConvNet model has been suggested and trained using Hybrid Constant-Q Transform (HCQT) by Tiwari et al. [19]. Each phonocardiogram signal's acoustic properties, including the dominating MFCC features, are fed into five-layer regularised ConvNets, where the CQT, Variable-Q Transform (VQT), and HCQT are extracted. Based on a review of related literature, it can be concluded that this is the first instance of using HCQT for PCG signals.

3. Proposed Model

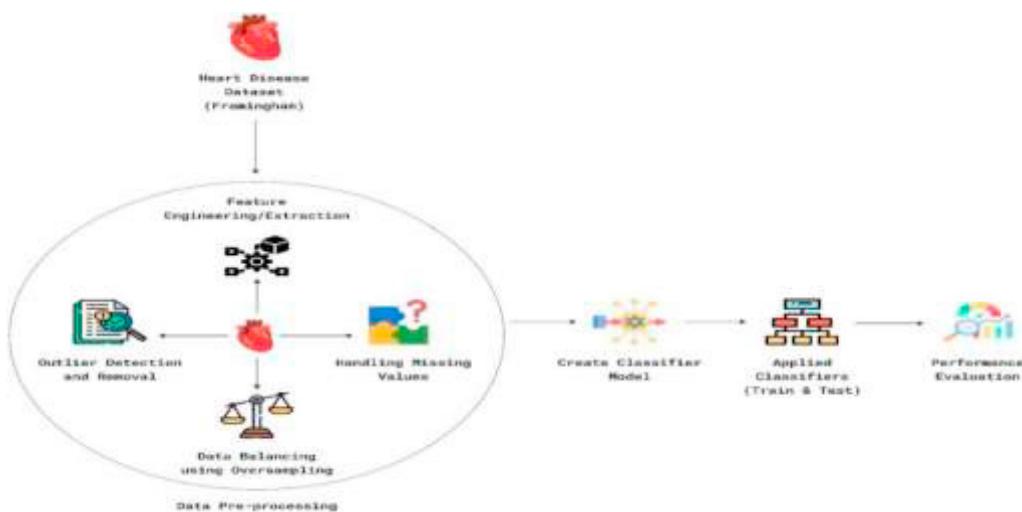


Figure 1: Experimental Methodology

In this section proposed model is a discussion for Heart disease prediction. We apply various machine algorithms like Decision tree, PCA, SVM and Random forest. Among them, a random forest gives the best accuracy of 97%. In the Framingham dataset we do various machine learning steps to get the final output. The Figure 1 represents the machine life cycle of heart disease prediction. Data collection is carried out by gathering relevant medical records, including patient demographics, symptoms, and diagnostic test results. Next, the data is pre-processed by cleaning, normalizing, and handling missing values. Feature selection and engineering techniques are then applied to extract meaningful predictors. The prepared dataset is split into training and testing sets. The training data is used to teach ML algorithms like DT, RF, and SVM, and the testing data is used to assess how well the algorithms performed.

3.1 Dataset Description

Framingham Dataset predicts heart disease risk over the following decade. This dataset is used to assess all aspects of applicants to determine whether they are going to suffer heart disease. This dataset describes the medical details of patients [16]. The task involves predicting whether a customer will repay a loan or credit. The "Framingham" heart disease dataset comprises 4,240 records with 16 columns and 15 attributes. The objective of the data collection is to determine if the patient has a 10-year chance of developing CHD. There is a missing value present in this dataset [17, 18].

The number of classes of this dataset is Two (0, 1). Each row represents an individual with various attributes such as age, education, smoking status, cholesterol levels, blood pressure, BMI, and other factors that are known to be associated with heart disease. The last column, "TenYearCHD," indicates whether the individual had a coronary heart disease event within the next ten years (1 for yes, 0 for no) [20,21,22]. A majority class refers to a situation in which a dataset exhibits a significantly higher number of instances in one class compared to the other classes.

4. Methodology

Methods for comparing how well various cardiac disease classification strategies work are outlined below. This study examines and discusses a variety of algorithms for detecting heart disease using machine learning. Here is an explanation of some commonly used algorithms: Decision tree, Principal component Analysis (PCA), Support Vector Machines (SVM), Random Forest, Gradient Boosting. An automated machine learning (ML) model is introduced in this research to predict and classify heart disease.

4.1 Data Collection (Framingham Dataset)

Data collection is the initial step machine learning model. It contains 4,240 records, 16 columns and 15 attributes. There are some missing values and unimportant features are present, so we need data pre-processing here.

4.2 Data Pre-processing

Data pre-processing, a crucial and necessary first stage in the creation of a ML model, prepares and makes suitable the raw data for use in the model. Data pre-processing is necessary because the data must be correctly cleansed and formatted. The task may involve evaluating data, finding relevant properties, and transforming them into a model-friendly format. It increases the model's precision. Its objective is to repair data that is flawed in some manner, such as being incomplete, noisy, or otherwise inconsistent. We use random over sampling for data pre-processing, and our strategy is minority.

A correlation heatmap is a graphical depiction that showcases the correlation between variables within a dataset, that depicts in Figure 2. It offers a visual summary of the magnitude and direction of the linear relationship between variable pairs. Each cell in the heatmap displays a correlation coefficient between two variables, and the intensity of the correlation is represented by the cell's colour. Outlier detection and removing unwanted data is very important.

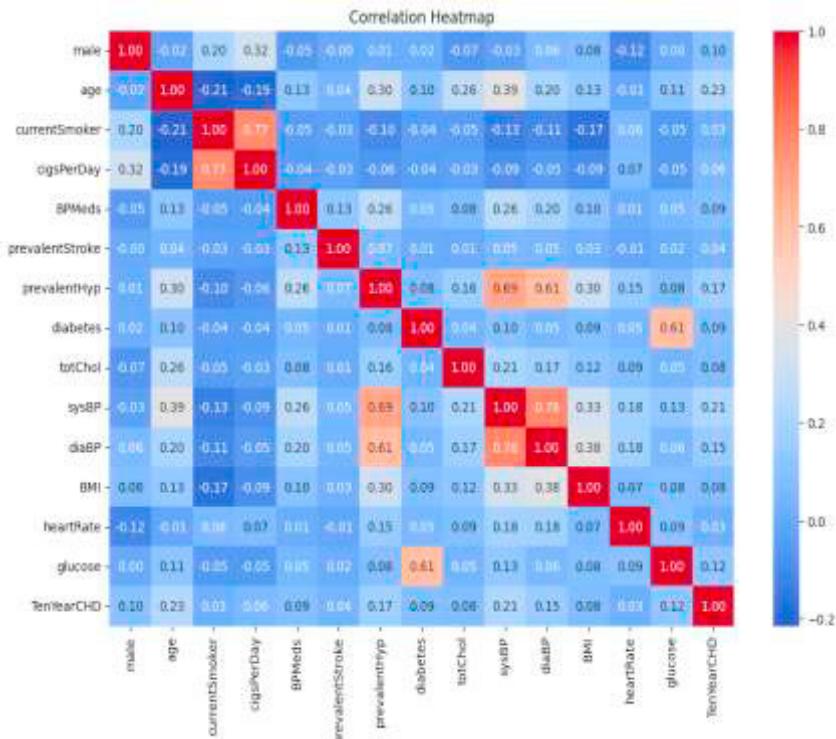


Figure 2: Correlation Heatmap for Framingham Dataset

4.3 Model Training

Machine language model training involves feeding a ML algorithm processed data to develop a model. A training dataset comprising sample output data and its corresponding set of input data should be prepared before feeding.

4.4 Create Model

Several machine learning algorithms have been put to use in this investigation of detecting cardiac disease. The selected algorithms are supervised machine algorithm including include decision tree, random forest, principal component analysis (PCA), and gradient boosting. Each of these algorithms brings unique strengths and characteristics to the task of heart disease detection. Among them, a random forest gives the best accuracy.

4.4.1 Decision Trees

A tree is constructed by the decision tree classifier using observations from data items. Decision trees are hierarchical structures that utilize a sequence of if-else conditions to generate predictions. In heart disease identification, decision trees can identify the most relevant features and establish rules for identifying patient symptoms as heart disease or not. Classes and Regression Trees (CART) and C4.5 are popular decision tree methods. The class labels are displayed in the leaves, while the branches represent the features that are pivotal in determining the class labels for the respective classes. Decision trees are employed to identify data objects by leveraging their visual components. In the Framingham dataset, the decision tree algorithm determines the optimal splits at each node and assigns class labels or values to the leaf nodes, that is shown in Figure 3.

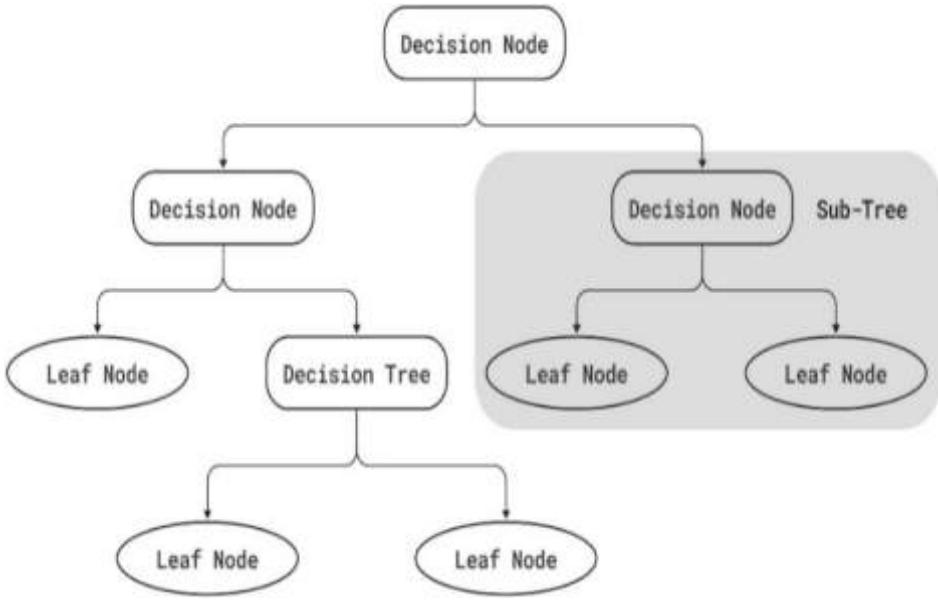


Figure 3: Graphical representation of Decision Tree

4.4.2 Gini Impurity

Gini impurity is a metric used to assess the impurity or heterogeneity of a node within a decision tree. Mathematically, it is calculated as the sum of the squared probabilities of each class label in the node. Greater the value of the index, higher would be the inequality. The Gini impurity at a node can be expressed as:

$$G = \sum_{i=1}^C p(i) * (1 - p(i)) \quad (1)$$

4.4.3 Information Gain

The information gain is a measure of how much uncertainty is eliminated when data is partitioned on the basis of some characteristic. Entropy is calculated using the class label distribution in the node. Mathematically, information gain is the difference between the entropy before and after the split. The information gain at a node can be expressed as:

$$\text{Gain} = E_{\text{parent}} - E_{\text{children}} \quad (2)$$

4.4.4 Entropy

Entropy is used to calculate information gain, which is the difference between the entropy before and after a split. At each node, we select the best split based on the information gain it generates, with the goal of increasing the quality of the resulting subsets as much as possible. Decision tree algorithms are able to make educated decisions on how to split data in order to minimise entropy and increase the purity of the resulting subsets because of the idea of entropy, which offers a measure of the disorder or uncertainty in a set of data. The entropy at a node can be expressed as:

$$E = - \sum_{i=1}^N p_i \log_2 p_i \quad (3)$$

4.4.5 Principal Component Analysis (PCA)

PCA is a useful dimension reduction method in ML and data analysis. It helps computers learn and predict by simplifying and organising data. Its objective is to convert a dataset with high dimensions into a lower-dimensional representation while retaining the most crucial information or patterns inherent in the data. The principal components of a Framingham dataset of heart disease are identified using the PCA algorithm. These principal components are new variables that represent linear combinations of the original features. The first component

captures the most data variation, and the second collects the most variance not accounted for by the first (in an orthogonal direction). Performing PCA on a specific dataset involves calculating the covariance matrix of the entire dataset and computing the corresponding eigenvectors and eigenvalues.

$$\text{cov}(A, B) = \frac{1}{m} \sum_{j=1}^m (a - \bar{a})(b - \bar{b}) \quad (4)$$

Here,

$\text{cov}(A, B)$:	covariance between A & B variables
$a & b$:	members of A & B variables
$\bar{a} & \bar{b}$:	mean of A & B variables
m	:	number of members

4.4.6 Support Vector Machines (SVM)

An optimal hyperplane is sought by the SVM, which is a supervised learning algorithm utilized to classify data points into distinct classes. Mapping patient data into a high-dimensional feature space and identifying the optimum hyperplane to maximally differentiate heart disease patients from others can be done with SVMs. Using kernel functions, Support Vector Machines (SVMs) can handle linearly separable and non-linearly separable data. A separator between categories is found, and the data are converted to generate a hyperplane.

4.4.7 Random Forests

In ensemble learning, random forests blend many decision trees to create predictions. A random forest trains each decision tree with a distinct data and feature sample. Random forests use the forecasts of numerous trees to create accurate and reliable heart disease detection models. Random forests can handle high-dimensional data and complex feature interactions. Once all the decision trees have been grown, predictions are made for new instances by aggregating the predictions of each tree. In classification problems, each tree "votes" for a class label, and the class label with most votes becomes the final prediction. In regression problems, the average of the predictions from all the trees is taken as the final prediction. The potential for overfitting can be minimised, as can the amount of time spent training, that depicts in Figure 4.

Steps involving Random Forest Algorithm:

- Step-1: Take shuffled representations of the data or training set at hand.
- Step-2: For every bit of training data, this algorithm will create a decision tree.
- Step-3: The voting process will be carried out by averaging the tree of decisions.
- Step-4: Finally, select the most voted prediction result as the final prediction result.

The below diagram represents the working of Ensemble Technique (Random Forest).

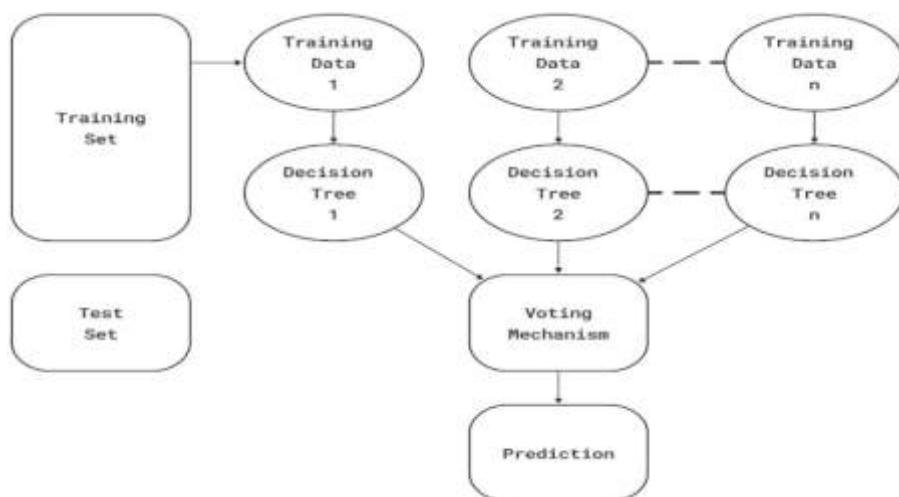


Figure 4: Graphical representation Random Forest

5. Result Analysis and Description

Model testing is a critical step in the ML workflow, where the performance and generalization capabilities of a trained model are evaluated using unseen data. The purpose of model testing is to assess how well the model can make accurate predictions on new, previously unseen instances. Effective model testing is essential to ensure reliable and accurate predictions when deploying the model in real-world applications.

5.1 Model Evaluation

Evaluation of a Heart Disease prediction model is crucial to its medical industry viability. It helps identify heart disease risk. It guarantees model accuracy, reliability, and performance in many settings. The model's accuracy, precision, recall, f1 score, AUC-ROC curve, and gain coefficient are needed here. The confusion matrix for the used model is shown in Figure 5.

5.2 Performance Evaluation and Metrics

Performance evaluation is a crucial aspect of developing accurate and reliable ML models for heart disease detection. It enables researchers and healthcare professionals to assess the effectiveness of the models, compare different algorithms or techniques, and make informed decisions regarding their implementation in clinical practice. There are a number of standard evaluation criteria and approaches used to assess the efficacy of ML models for diagnosing cardiac disease. Table 1 shows 84% accuracy in Decision Tree. Table 2 shows 97% accuracy in Random Forest. Table 3 shows 79% accuracy in PCA. Table 4 shows 68% accuracy in SVM. Thus the Random Forest gives best accuracy than other machine learning algorithms.

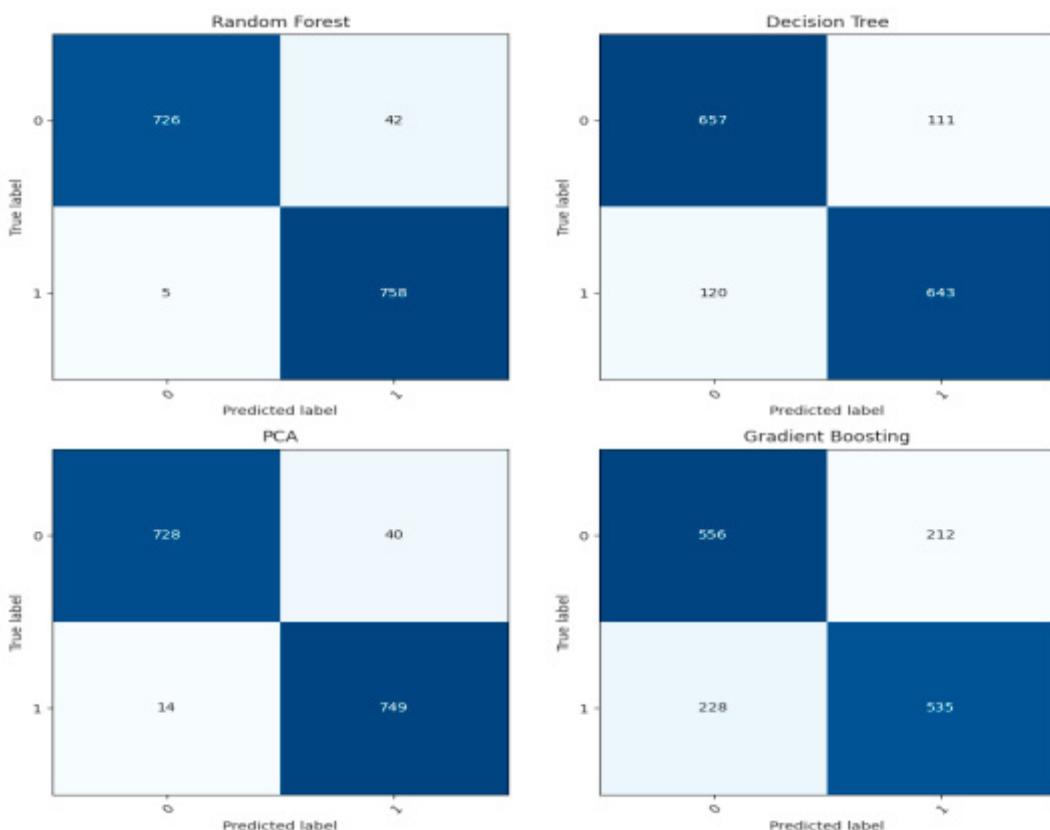


Figure 5: Confusion matrix of Different algorithms

Table 1: Decision Tree Results

Metrics	Prec.	Recall	F1-sc.	Supp.
weighted avg	0.84	0.84	0.84	1531
macro avg	0.84	0.84	0.84	1531
accuracy			0.84	1531

Table 2: Random Forest Results

Metrics	Prec.	Recall	F1-sc.	Supp.
weighted avg	0.97	0.97	0.97	1531
macro avg	0.97	0.97	0.97	1531
accuracy			0.97	1531

Table 3: PCA Results

Metrics	Prec.	Recall	F1-sc.	Supp.
weighted avg	0.80	0.80	0.79	1531
macro avg	0.80	0.80	0.79	1531
accuracy			0.79	1531

Table 4: SVM Results

Metrics	Prec.	Recall	F1-sc.	Supp.
weighted avg	0.68	0.68	0.68	1531
macro avg	0.68	0.68	0.68	1531
accuracy			0.6832	1531

Accuracy measures the percentage of positive cases predicted correctly. It evaluates the model's false positive reduction. Recall, or sensitivity, measures the ratio of accurately anticipated positive instances to total positive instances. Precision and memory are crucial in healthcare, where erroneous positives and negatives can have serious consequences. Figure 6 shows heart disease prediction ML model precision and recall.

5.2.1 Model Selection

Model selection techniques, such as k-fold cross-validation, can be employed when comparing multiple machine learning algorithms or models. The best-performing model can be identified based on evaluation metrics using these techniques. In addition to hyper parameter optimisation, ensemble methods and performance comparison against baseline models are additional factors to consider when choosing the most suitable model. Noteworthy is the fact that performance evaluation should be conducted using diverse and representative data sets. To ensure impartial evaluation and prevent disparities in model performance, pre-processing of data, handling of missing values, imbalanced class distribution, and potential biases should be carefully considered.

In summary, a comprehensive analysis of various metrics such as accuracy, sensitivity, specificity, precision, recall, F1-score, and AUC-ROC is required for performance evaluation in heart disease detection using ML techniques. Employing cross-validation techniques and considering model selection methodologies helps ensure robust and reliable performance assessment. By carefully evaluating the performance of machine learning models, researchers and healthcare professionals can make informed decisions about their implementation and potential impact on heart disease detection and patient care. Figure 7 represents the accuracy comparison between models

used for heart disease prediction.

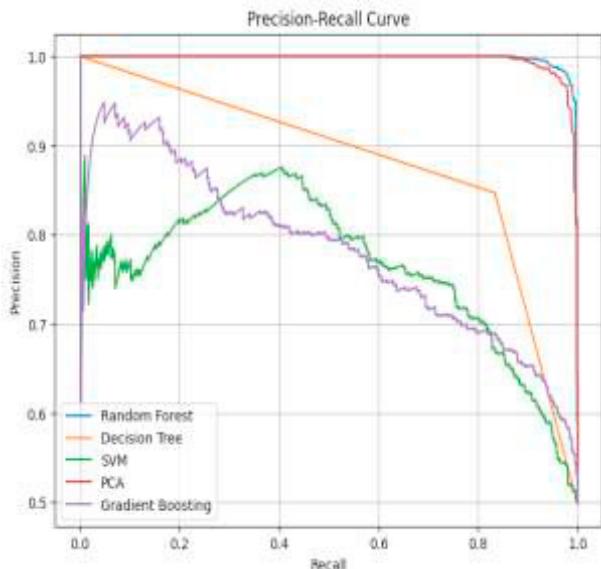


Figure 6: Precision-Recall curve for each model

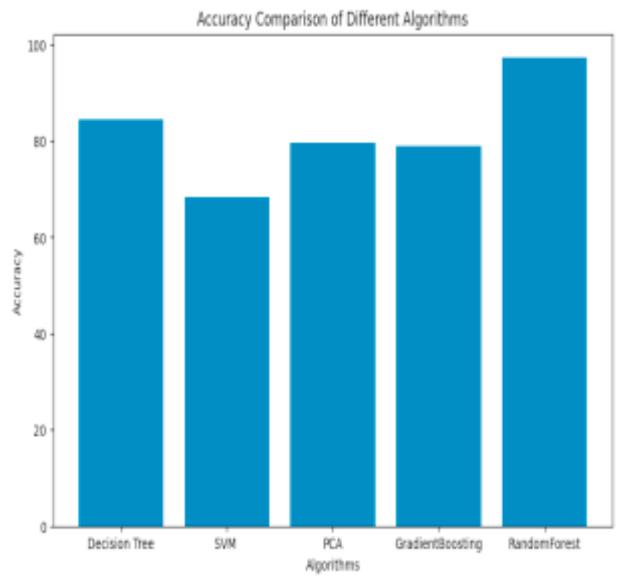


Figure 7: Accuracy Comparison of Different Algorithms

6. Conclusion

In conclusion, this comprehensive review has provided valuable insights into the application of machine learning algorithms in heart disease detection. Machine learning's potential in healthcare lies in its ability to enhance diagnostic precision, streamline resource allocation, pave the way for personalised therapy, and fuel progress in this vital sector. The performance evaluation metrics discussed, including accuracy, sensitivity, specificity, precision, recall, F1-score, and AUC-ROC, provide a robust framework for assessing the effectiveness of machine learning models in heart disease detection. These metrics, along with cross-validation techniques and model selection strategies, ensure reliable and accurate predictions. By providing a comprehensive overview of the existing literature, this paper aims to serve as a valuable resource for researchers, clinicians, and policymakers interested in leveraging machine learning techniques to advance heart disease detection. Ultimately, the integration of machine learning algorithms in healthcare has the potential to transform the diagnosis and management of heart disease, leading to improved patient outcomes, reduced mortality rates, and more efficient healthcare systems. The suggested learning system outperforms eleven recently proposed heart failure diagnosis methods and other well-known machine learning models. It was also discovered that by decreasing the amount of features, the suggested system speeds up the development of machine learning models. The experimental findings suggest that the proposed learning system can assist clinicians in better detecting heart failure.

References

- [1]. Coffey, S., Roberts-Thomson, R., Brown, A., Carapetis, J., Chen, M., Enriquez-Sarano, M., ... & Prendergast, B. D. (2021). Global epidemiology of valvular heart disease. *Nature Reviews Cardiology*, 18(12), 853-864.
- [2]. De Hert, M., Detraux, J., & Vancampfort, D. (2022). The intriguing relationship between coronary heart disease and mental disorders. *Dialogues in clinical neuroscience*.
- [3]. Tsao, C. W., Aday, A. W., Almarzooq, Z. I., Alonso, A., Beaton, A. Z., Bittencourt, M. S., ... & American Heart Association Council on Epidemiology and Prevention Statistics Committee and Stroke Statistics Subcommittee. (2022). Heart disease and stroke statistics—2022 update: a report from the American Heart Association. *Circulation*, 145(8), e153-e639.
- [4]. Katarya, R., & Srinivas, P. (2020, July). Predicting heart disease at early stages using machine learning: A survey. In *2020 International Conference on Electronics and Sustainable Communication Systems (ICESC)* (pp. 302-305). IEEE.
- [5]. Liu, A., Diller, G. P., Moons, P., Daniels, C. J., Jenkins, K. J., & Marelli, A. (2023). Changing epidemiology of congenital heart disease: effect on outcomes and quality of care in adults. *Nature Reviews Cardiology*, 20(2), 126-137.
- [6]. Bhatt, C. M., Patel, P., Ghetaia, T., & Mazzeo, P. L. (2023). Effective heart disease prediction using machine learning techniques. *Algorithms*, 16(2), 88.

- [7]. Aluru, J. S., Barsouk, A., Saginala, K., Rawla, P., & Barsouk, A. (2022). Valvular heart disease epidemiology. *Medical Sciences*, 10(2), 32.
- [8]. Ahsan, M. M., & Siddique, Z. (2022). Machine learning-based heart disease diagnosis: A systematic literature review. *Artificial Intelligence in Medicine*, 128, 102289.
- [9]. Sekar, J., Aruchamy, P., Sulaima Lebbe Abdul, H., Mohammed, A. S., & Khamruddeen, S. (2022). An efficient clinical support system for heart disease prediction using TANFIS classifier. *Computational Intelligence*, 38(2), 610-640.
- [10]. Zhou, Y., Zhu, X., Cui, H., Shi, J., Yuan, G., Shi, S., & Hu, Y. (2021). The role of the VEGF family in coronary heart disease. *Frontiers in cardiovascular medicine*, 8, 738325.
- [11]. Behera, M. P., Sarangi, A., Mishra, D., & Sarangi, S. K. (2023). A Hybrid Machine Learning algorithm for Heart and Liver Disease Prediction Using Modified Particle Swarm Optimization with Support Vector Machine. *Procedia Computer Science*, 218, 818-827.
- [12]. Rajendran, R., & Karthi, A. (2022). Heart disease prediction using entropy based feature engineering and ensembling of machine learning classifiers. *Expert Systems with Applications*, 207, 117882.
- [13]. Ahsan, M. M., & Siddique, Z. (2022). Machine learning-based heart disease diagnosis: A systematic literature review. *Artificial Intelligence in Medicine*, 128, 102289.
- [14]. Ramesh, T. R., Lilhore, U. K., Poongodi, M., Simaiya, S., Kaur, A., & Hamdi, M. (2022). Predictive analysis of heart diseases with machine learning approaches. *Malaysian Journal of Computer Science*, 132-148.
- [15]. Ghasemieh, A., Lloyed, A., Bahrami, P., Vajar, P., & Kashef, R. (2023). A novel machine learning model with Stacking Ensemble Learner for predicting emergency readmission of heart-disease patients. *Decision Analytics Journal*, 7, 100242.
- [16]. Pathan, M. S., Nag, A., Pathan, M. M., & Dev, S. (2022). Analyzing the impact of feature selection on the accuracy of heart disease prediction. *Healthcare Analytics*, 2, 100060.
- [17]. Wang, J., Rao, C., Goh, M., & Xiao, X. (2023). Risk assessment of coronary heart disease based on cloud-random forest. *Artificial Intelligence Review*, 56(1), 203-232.
- [18]. Emmons-Bell, S., Johnson, C., & Roth, G. (2022). Prevalence, incidence and survival of heart failure: a systematic review. *Heart*, 108(17), 1351-1360.
- [19]. Tiwari, S., Jain, A., Sharma, A. K., & Almustafa, K. M. (2021). Phonocardiogram signal based multi-class cardiac diagnostic decision support system. *IEEE Access*, 9, 110710-110722.
- [20]. Singh, D., Prusty, S. K., Sarangi, S. S., Sahoo, S., & Biswal, A. K. (2020). Attitude, anxiety, psychological effects and prevention during COVID-19 in India. *Indian J Pharm Educ Res*, 54(3s), s400-10.
- [21]. Jena, R., Biswal, A. K., & Singh, D. (2022). A Novel Approach for an IoT-Based U-Healthcare System. In *Handbook of Research on Mathematical Modeling for Smart Healthcare Systems* (pp. 247-260). IGI Global.
- [22]. Malla, S., Sahu, P. K., Patnaik, S., & Biswal, A. K. (2023). Obstacle Detection and Assistance for Visually Impaired Individuals Using an IoT-Enabled Smart Blind Stick. *Revue d'Intelligence Artificielle*, 37(3).