JCBI
Journal of Computing & Biomedical Informatics
ISSN: 2710 - 1606
Volume 08 Issue 01
2024
Research Article
https://doi.org/10.56979/801/2024

Artificial Intelligence-Enhanced Risk Stratification and Prediction of Cardiovascular Disease through Machine Learning

Faiza Shabbir[1], Nimra Tariq[1*], Sidra Siddiqui[1], Nimra Tanveer[1], and Zille Huma[1]

[1]Department of Basic Sciences, The Superior University Lahore, Pakistan.
[*]Corresponding Author: Nimra Tariq. Email: nimra.tariq@superior.edu.pk

Received: March 21, 2024 **Accepted:** September 21, 2024

Abstract: Since cardiac illnesses provide substantial health concerns to patients, an accurate diagnosis is crucial. Due to the involvement of multiple factors, including smoking, high blood pressure, high blood sugar, excessive cholesterol, and environmental influences, it is difficult to diagnose cardiac abnormalities based only on symptoms. To tackle this, we use innovative machine learning algorithms to evaluate large volumes of medical data, find hidden patterns, and forecast the course of disease. Risk stratification and illness prediction comprise the two main aspects of our study. We evaluate and forecast cardiac anomalies using state-of-the-art algorithms such as Decision Trees, AdaBoost, and Extra Tree classifiers. The main objective of this research is to decrease errors and increase forecast accuracy by merging two datasets. For the old heart disease dataset, we obtained different accuracies for the three classifiers (Decision Tree, 81.5%), Extra Tree, (89%), and AdaBoost, 85.3%). On the newly established cardiovascular disease dataset, the accuracy of AdaBoost (98%), Decision Tree (96.5%), and Extra Tree (99%) was significantly higher. During GridSearchCV optimization, the accuracy rose, demonstrating the robustness of our models. This study shows how individuals who are at a high risk of cardiac events in the future can be identifying using machine learning.

Keywords: Cardiac Disease; Machine Learning; Classification; AdaBoost; Decision Tree; Extra Tree Classifier.

1. Introduction

Heart disease continues to be the leading cause of death worldwide, affecting millions of people. The cost of effective diagnostic tests can be reducing by utilizing computer technology to support accurate, trustworthy, and competent medical diagnosis. Machine learning is a vital and important field that could help with cardiac illness diagnosis for both patients and medical professionals. [1-3].

Heart and blood vessel disorders, such as heart disease and stroke, are included in the category of cardiovascular disease. Cardiovascular disease risk stratification is the procedure of determining an individual's probability of suffering from cardiovascular events, such as heart attacks, strokes, or heart failure, over a given period of time [4, 5].

It is imperative to evaluate an extensive array of risk variables, as illustrated in Figure 1.1, to categories an individual's overall risk assessment and ascertain the risk categories they belong to concerning their likelihood of getting cardiovascular disease. Numerous factors can lead to heart disease. This study predicts cardiac disease using categorization approaches. Despite the abundance of patient data found in hospital and clinic records, there is a dearth of published material on this quickly evolving topic [6].

To preserve patient privacy, more hospitals should be encouraged to submit high-quality datasets, as this has a significant impact on prediction accuracy. Supporting this would give researchers dependable resources to help with model development and yield useful results [7, 8].

There are four main types of cardiovascular diseases, as shown in Figure 1. The main cause of coronary heart disease is a restriction of blood flow to the heart muscle. Angina, heart attacks, and heart failure are among the symptoms brought on by this blockage, which also raises cardiac strain. The second group consists of transient ischemic attacks (TIAs) and stroke-related TIAs, which are reason by blood clots that block the brain and shortly stop blood flow.

Severe leg pain, recurrent ulceration, limb weakening, and limb and foot hair loss characterize the third type, known as tangential artery infection. The final category, which affects the main blood vessel, the aorta, called aortic disease. Despite being a potentially lethal illness, it does not exhibit signs or symptoms [9].

One popular topic in data science is machine learning, which is a subfield of artificial intelligence study. A wide range of activities, including classification, decision-making, and prediction, can accomplished with machine learning techniques. The necessary is training of data for understanding machine learning algorithms [10-13].

Cardiovascular disease prediction sallow users suffering from cardiac ailments receive a prediction result from a machine learning system. Machine learning algorithms have evolved because of recent technological advancements [14, 15].

The Random Forest Algorithm developed for use because of the precision and effectiveness of this suggested remedy. To improve accuracy, additional algorithms might be use. These algorithms perform better when they employ greater numbers of parameters [16-18].

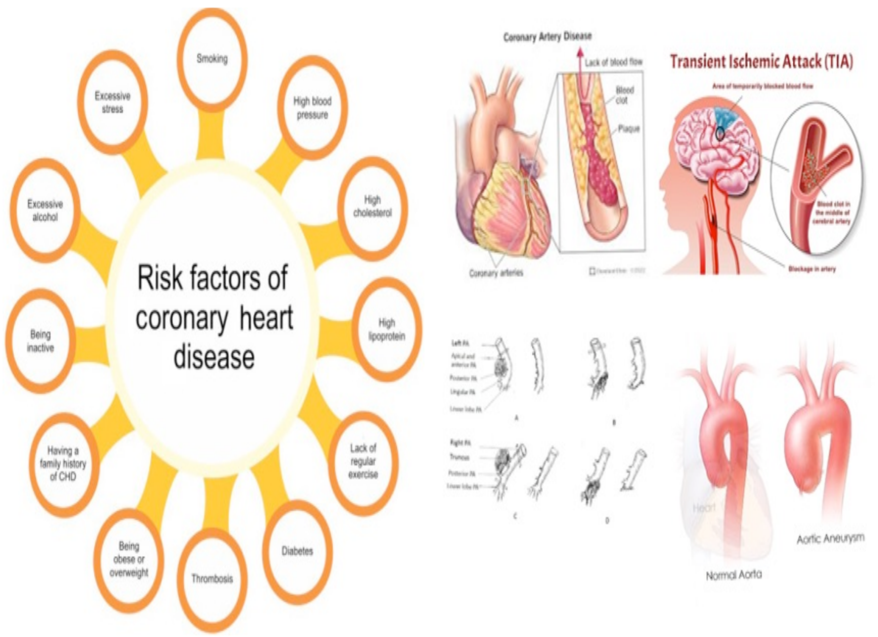

Figure 1. Risk factors of cardiovascular disease and four types of cardiovascular diseases [10]

According to a survey, Pakistan's death rate from heart disease currently at 15.36%, and it could rise to about 23 million deaths yearly by 2030 [20, 21].

One of a physical corpse's most basic organs is the heart. Greater precision and accuracy are need when identifying heart disorders [22].

Cardiovascular diseases might not be identifying in their early stages in real time. This needs further investigation. The proposed work presents an early-stage, accurate cardiovascular disease prediction using a dataset of heart ailments. Numerous ML techniques are required for the approach under discussion [23].

As seen in Figure 1.3, our goal is to apply machine learning algorithms and risk stratification to do prediction on cardiac disorders. It's also critical to examine the different strategies used in order to ascertain which machine learning algorithms are most effective [24].

In the last few years, machine-learning approaches thoroughly studied in the field of cardiovascular disease, with an emphasis on early disease identification and prevention. The topic has been the subject of several noteworthy studies that have enabled this improvement. The procedures, conclusions, and limitations of these investigations are covered in this particular study. The majority of research has examined how classification algorithms can accurately identify cases of cardiac disease [25].

All of these research efforts have advanced significantly by using machine-learning methods to anticipate cardiac problems. So, to calculate and predict cardiac issues, different algorithms like K-NN Gaussian Naive Bayes and Decision Tree etc. have been implemented. These algorithms can handle intricate linkages within the data, enhancing the precision of risk assessments and predictions.

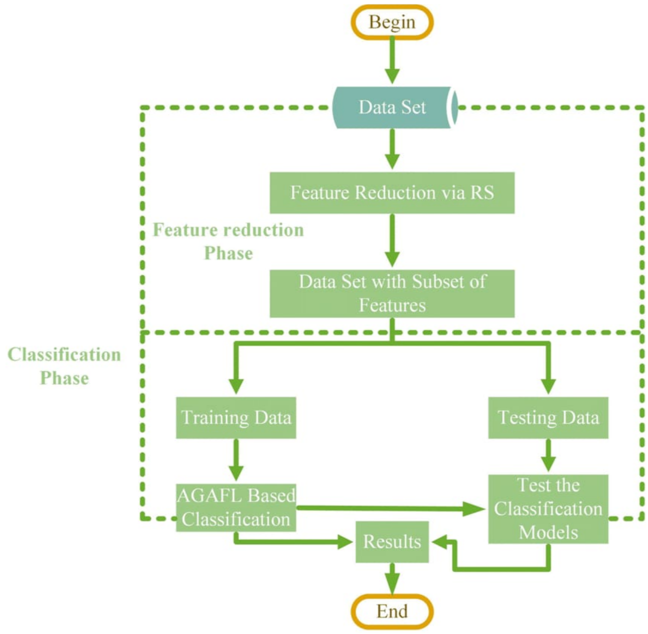

Figure 2. Disease prediction using machine learning

2. Methodology

A risk stratification model was developed and used to classify the patients on their health status. We used statistical methods and machine learning algorithms to determine critical features in premature risk estimations. The collected data goes to the data processing module, which eliminate repeating values and remove unwanted noise, in order to retrieve the data. The Figure 3 shows the proposed model that was implemented. We used two different datasets from Cleveland heart disease Database, a common source of Kaggle then visualized it to perform risk stratification, and preprocessed it to remove duplicate or nearly constant rows as well as missing values. Next, we fed our ML models with our "clean" dataset that kept in Comma Separated Value file. In order to extract the features and determine the labels, we first divided our both datasets into training and testing subsets using a Python script. The prepared data then fed to the three-machine learning algorithms AdaBoost, Decision Tree and Extra Tree Classifier for training. The steps followed in proposed model are explained below.

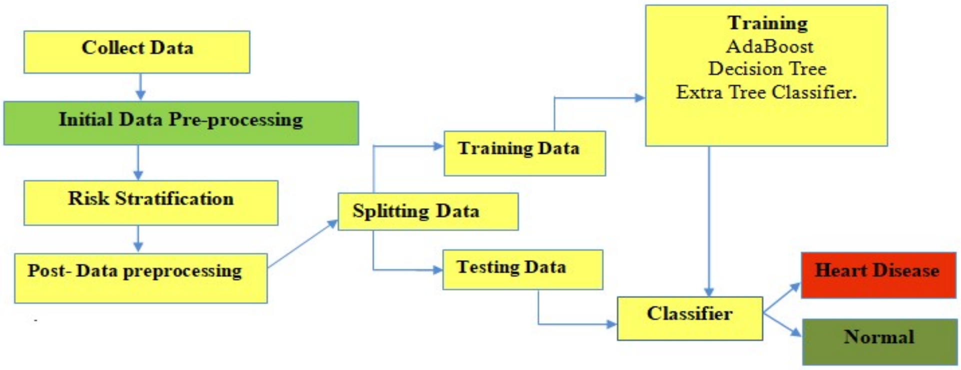

Figure 3. Proposed Model to predict cardiovascular disease

2.1. Data collection

The attributes of both used datasets are given in table 1 and table 2.

Table 1. Attributes of 1st Dataset

Column Number	Attribute	Description
0	Age	Age of patients in years
1	Sex	Patient's Gender (M = Male, F = Female)
2	Chest Pain Type	Chest pain type =Typical angina (TA), Atypical angina (ATA),

		Non-angina pain (NAP) and Asymptomatic (ASY)
3	Resting BP	Measure Resting blood pressure in millimeter Hg at hospital admission.
4	Cholester ol	Level of serum cholesterol expressed in mg/dl.
5	Fasting BS	Blood sugar level after fasting > 120 mg/dl (1 = true, 0 = false)
6	Resting ECG	Electrocardiographic results at rest (Normal, ST = abnormality ST-T wave, LVH = Left ventricular hypertrophy according to Estes' criteria.)
7	Max HR	During exercise, the maximum heart rate achieved.
8	Exercise Angina	(Y = Yes, N = No) Exercise-induced angina
9	Old peak	Exercise induced ST depression comparative to rest
10	ST-Slope	Slopes of the peak exercise ST section (Up = Up sloping, Flat = Flat, Down = Down sloping)
11	Heart Disease	Heart disease Diagnosis (0 indicates No heart disease problems, 1 indicates heart disease problems)

Table 2. Attributes of 2nd Dataset

Column Number	Attribute	Description
0	Patient id	Unique identifier for each patient
1	Age	Age of patients in years
2	Gender	Patient’s Gender (1 = male, 0 = female)
3	Chest pain	Chest pain type practiced by patients=Typical angina (0), Atypical angina (1), non-angina pain (2) and Asymptomatic (3).
4	Resting BP	Measure Resting blood pressure in millimeter Hg at hospital admission.
5	Serum cholesterol	Level of serum cholesterol expressed in mg/dl.
6	Fasting blood sugar	Blood sugar level after fasting > 120 mg/dl (1 = true, 0 = false)

7	Resting electro	Electrocardiographic results at rest (Normal=0, ST-T wave abnormality=1, Left ventricular hypertrophy=2)
8	Max heart rate	During exercise, the maximum heart rate achieved.
9	Exercise angina	Exercise-induced angina (1 = Yes, 0 = No)
10	Old peak	Exercise induced ST depression comparative to rest
11	Slope	Slopes of the peak exercise ST section (Up sloping=0, Flat=1, Down sloping=2)
12	No of major vessels	Number of major vessels colored by fluoroscopy (0-3)
13	Target	Heart disease Diagnosis (0 indicates No heart disease problems, 1 indicates heart disease problems)

2.2. Data Preprocessing

Once the data gathered, some preliminary pre-processing done to ensure they are appropriate for analysis i.e. (Data Cleaning, Generating Dummy Variable, Exploratory Data Analysis). Data undergo additional pre-processing after initial pre-processing in order to be refined for model education and testing i.e. (Handling Class Imbalance, Feature Engineering).

2.3. Data Splitting

The data set then separated into training and testing subsets in order to preserve the target variable distribution between the two sets. Generally, 80% of the data use for training the models and residual 20% data use for testing of the data.

2.4. Training Data

The model was then trained using 80% of the chosen datasets, employing machine learning algorithms i.e. AdaBoost, Decision Tree and Extra Trees Classifier. These algorithms were chosen for their ability to handle classification tasks effectively.

2.5. Testing Data

The remaining 20% of the datasets were used as a text sets to validate the model's performance.

2.6. Classifier Implementation

The classifier was then implemented by selecting and fine-tuning machine learning models to accurately predict the cardiovascular disease. Every algorithm was carefully evaluated to make a reliable system for early detection and prevention.

3. Results and Discussion

Machine learning models were used to examine, visualize and predict cardiovascular disease, provided that insights into their predictive capabilities. A comprehensive discussion of the analysis results and their consequences was presented using two datasets.

3.1. Distribution Analysis

We analyzed and compared the distributions of key health indicators in both datasets such as age, resting blood pressure, cholesterol, maximum heart rate, and workout-induced ST despair (oldpeak). Using KDE plots and histograms, we visualized these distributions from both datasets, highlighting the differences between patients with and without cardiovascular disease, as shown in figures (4 & 5).

Both datasets indicated that middle-aged people (ages 50 to 60) were at a higher risk for coronary heart disease, with the second dataset having a broader age distribution. The second dataset exhibits more stable resting blood pressure of about 140-180 mmHg, while the first dataset has more variability. The first dataset contains a wide range of cholesterol values, but the second dataset has a maximum of 400 mg/dl and a significant drop at 200 mg/dl. Heart rates in the second dataset were more spread out, with peaks between 120-160 bpm, while the first peaked at 140-160 bpm. The second dataset also had a more consistent spread of oldpeak values, peaking between 2-4 units.

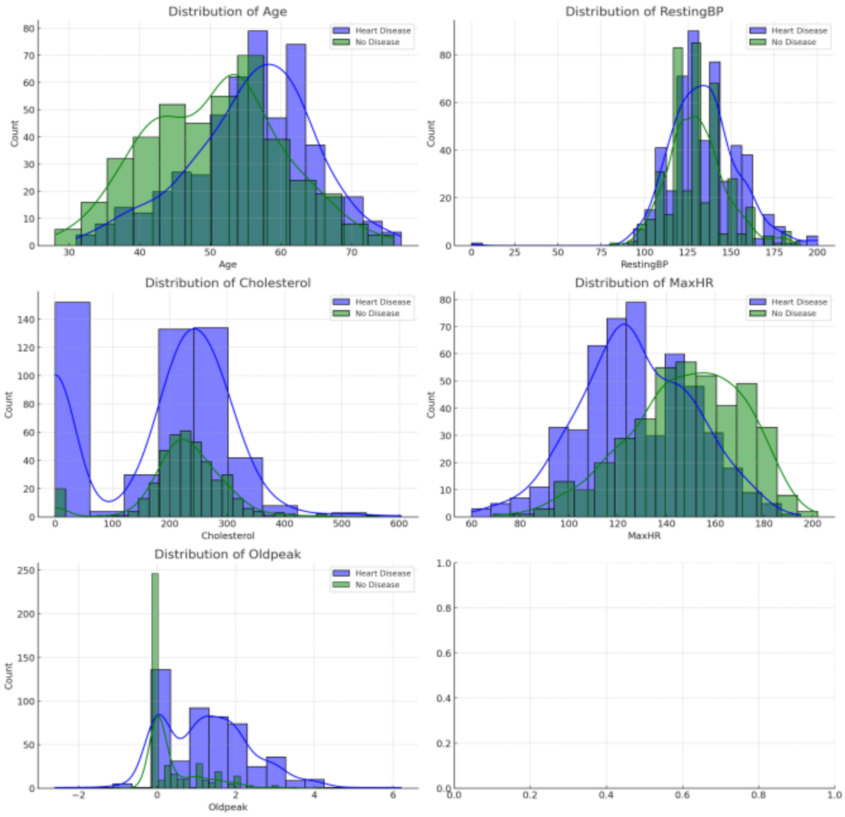

Figure 4. Distribution Analysis for 1st Dataset

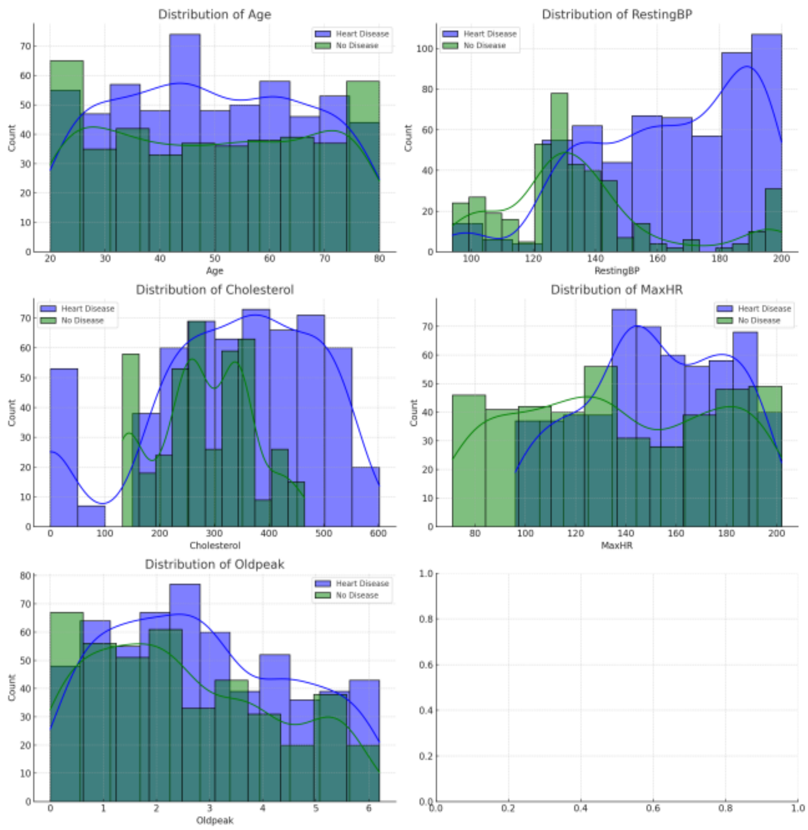

Figure 5. Distribution Analysis for 2ndDataset

3.2. Correlation Analysis

Correlation analysis was conducted to discover the relationship between key features in both datasets and their influence on the likelihood of coronary heart disease

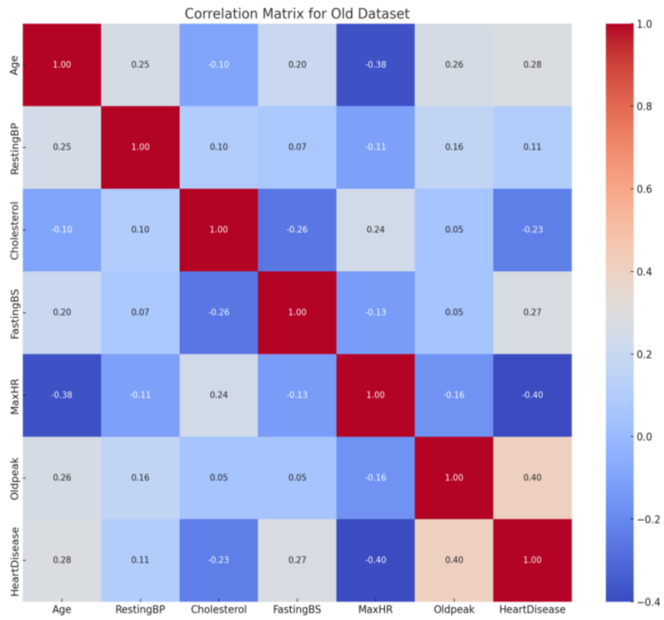

Figure 6. Correlation matrix for 1st Dataset

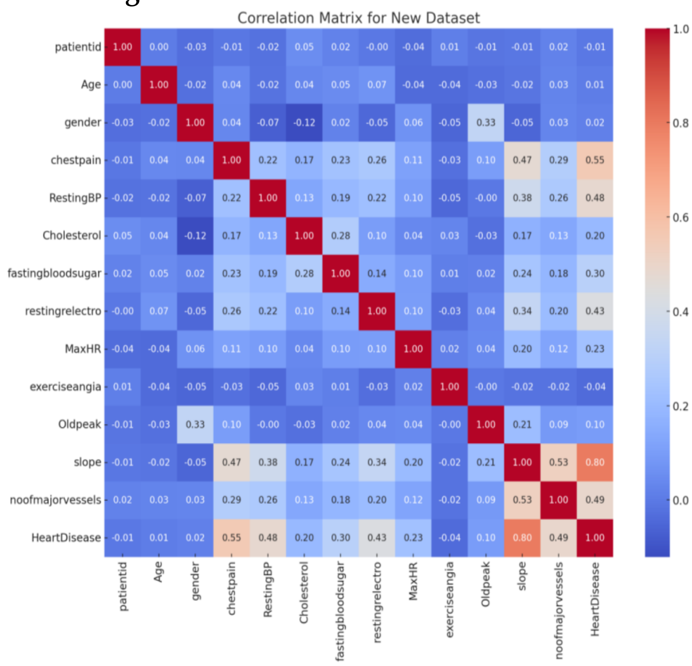

Figure 7. Correlation matrix for 2nd Dataset

The 1stdataset's "Age," "Max Heart Rate," and "Old Peak" were crucial factors for "heart disease" A poor relationship (-0.40) with "Max Heart Rate" indicated that lower charges are linked to coronary heart disease. The 2nddataset shows a significant relationship between "heart disease" and "Chest Pain,"

"Resting Blood Pressure," "Cholesterol," "Fasting Blood Sugar," and "Resting Electro" . The greatest association (0.55) is observed between "heart disease" and "Chest Pain."

3.3. Confusion Matrix for Model Evaluation

To analyze the performance of the classification models, confusion matrices were created for the Decision Tree, AdaBoost and Extra Tree classifiers. These matrices provide a visual representation of each model's performance, showing the count of true positives, true negatives, false positives and false negatives. This permits for a comprehensive evaluation of how well each model predicts a target class and detects areas for improvement.

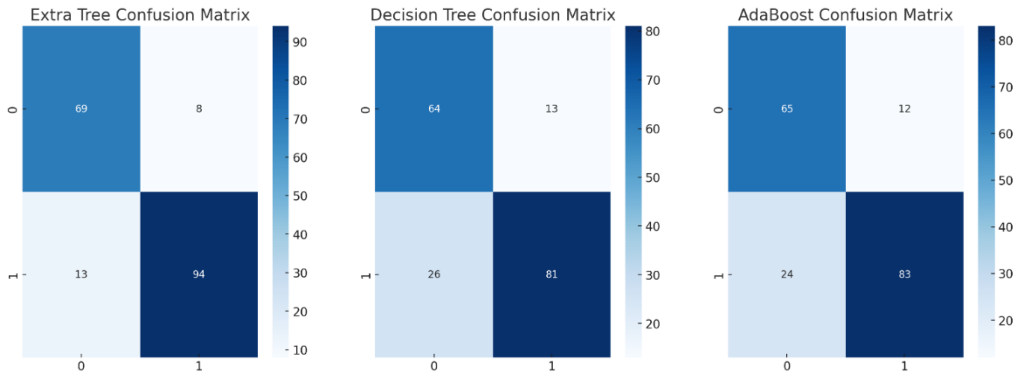

Figure 8. Confusion matrix for old Dataset

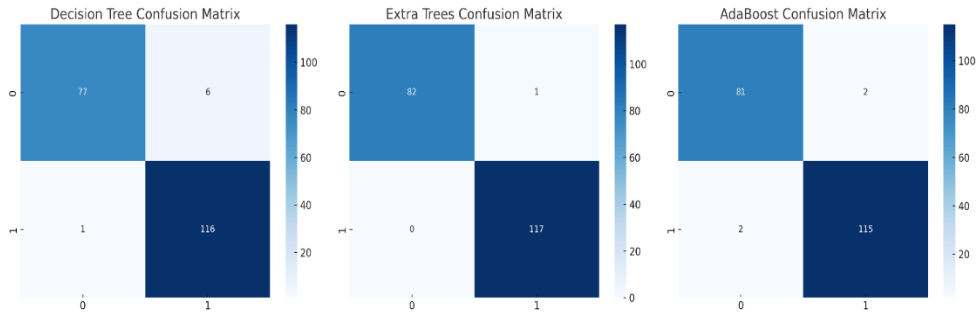

Figure 9. Confusion matrix for New Dataset

Table 3. Comparison of the Confusion matrix of both datasets:

Classifier	Datasets	True Positive	True Negative	False Positive	False Negative
AdaBoost	1st	83	65	12	24
	2nd	115	81	2	2
Decision Tree	1st	81	64	13	26
	2nd	116	77	6	1
Extra Tree	1st	94	69	8	13
	2nd	117	82	1	0

True Positive (TP): Data points where both the actual and predicted class are true
True Negative (TN): Data points where both the actual and predicted class are false
False Positive (FP): Data points where the actual class is false, but the predicted class is true
False Negative (FN): Data points where the actual class is true, but the predicted class is false

These confusion matrix metrics were used to evaluate the machine learning model with respect to key performance aspects, including its accuracy, precision, recall and overall effectiveness in predicting the target classes.

$$Accuracy = \frac{TP + TN}{TP + TF + FP + FN}$$

$$Precision = \frac{TP}{TP + FP}$$

$$Recall = \frac{TP}{TP + FN}$$

$$F1 - Score = 2.\frac{precision.Recall}{precision + Recall}$$

The results presented here as shown in Figures below, how three-machine learning models (DT) Decision Tree, Extra Trees and AdaBoost perform on two different datasets, Cardio and Heart. In the Cardio dataset, the Extra Trees model outperforms all other models with nearly perfect scores in Accuracy, Precision, F1-Score, and Recall. AdaBoost and Decision Tree models trail Extra Trees in the Heart dataset by a slight margin, performing equally. In terms of recall and F1-Score, AdaBoost and Extra Trees both perform better than the Decision Tree model.

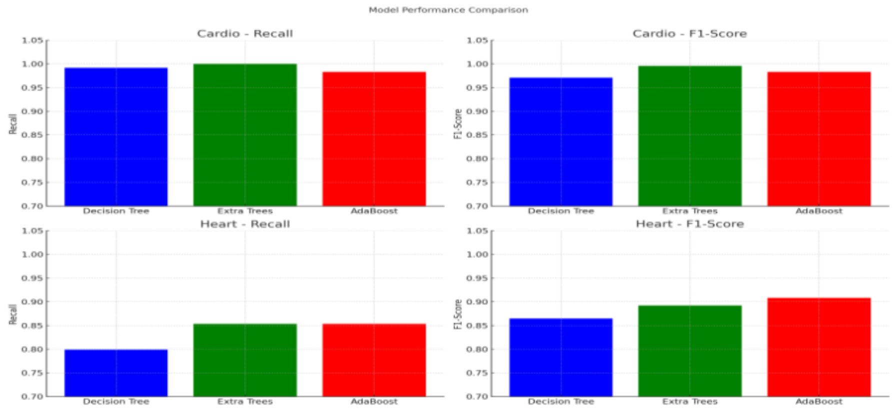

Figure 10. Model performance (Recall and F1 Score) comparison of both Dataset

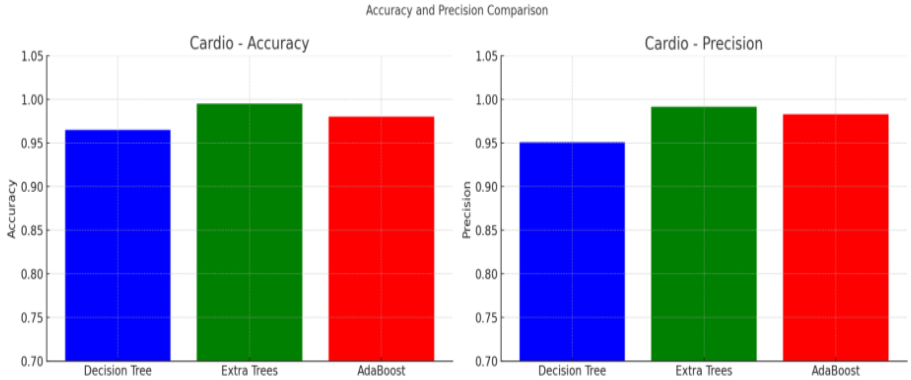

Figure 11. Model performance (accuracy and precision) comparison of 1stDataset

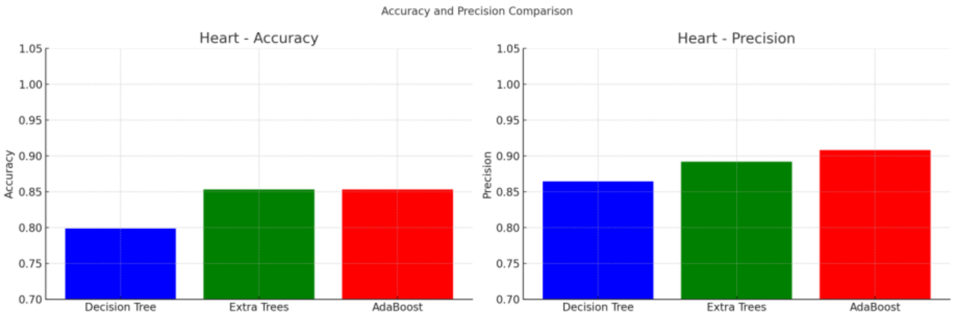

Figure 12. Model performance (accuracy and precision) comparison 2ndDataset

3.4. Receiver Operating Characteristic (ROC) Performance evolution of models

The ROC curves demonstrated that the Extra Trees and AdaBoost models consistently outperform the others. On the 1stDataset as shown in figure 11, Extra Trees obtained an AUC of 0.94, AdaBoost 0.91, and Decision Tree 0.81. For the 2nd dataset as shown in figure 12, both Extra Trees and AdaBoost each received top AUCs of 1.00, while Decision Tree achieved 0.95. Despite these findings, the Decision Tree variant is notably aggressive across datasets.

3.5. Overall Model Comparison for Both Datasets

On comparing models for the both datasets, Extra Trees and AdaBoost consistently perform better than Decision Trees (Tables 4 and 5). In Table 3.2 of the 1stdataset, Extra Trees and AdaBoost attain almost perfect AUCs of 0.99, with 99.0% accuracy and 99.1% precision. The Decision Tree trails with 96.5% accuracy and an AUC of 0.961. Table 3.3shown that Extra Trees had an AUC of 0.942, 89.7% accuracy, and 92.3% precision in the 2nddataset. The decision tree showed and AUC of 0.819 and 81.5% accuracy, while AdaBoost performed slightly better with an AUC of 0.915 and 85.3% accuracy. In all datasets combined, Extra Trees has the best overall performance.

3.6. Risk Stratification and Uncertainty Quantifications

We also calculated risk stratification and Uncertainty quantifications for better predictions analysis. Risk stratifications help us to identify and rank high-chance groups for closer monitoring, while uncertainty quantifications involve evaluating and running the inherent uncertainty in predictions.

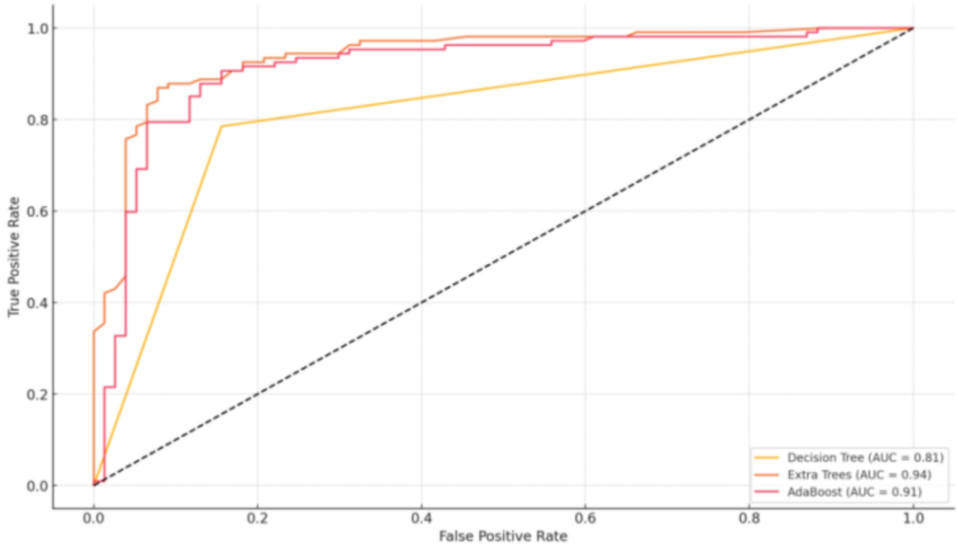

Figure 13. Receiver Operating Characteristic curve for 1stDataset

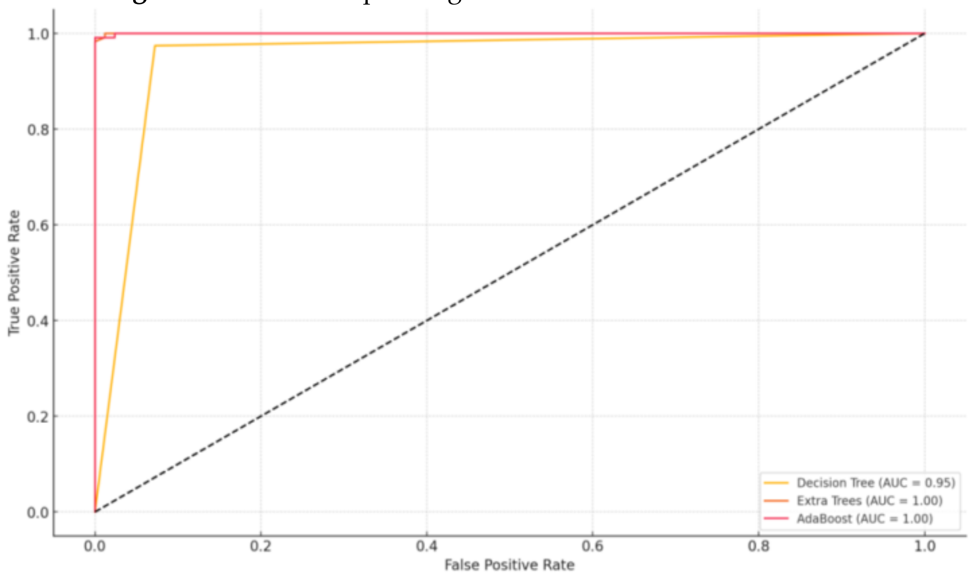

Figure 14. Receiver Operating Characteristic curve for 2ndDataset

Table 4. Model Comparison for Cardiovascular dataset (New):

Model	Accuracy	Precision	Recall	F1 Score	AUC
Decision Tree	0.965	0.958333	0.982906	0.970464	0.961333
Extra Trees	0.995	0.991453	0.991453	0.991453	0.999382
AdaBoost	0.980	0.981453	0.981453	0.981453	0.989794

Table 5. Model Comparison for Heart Datasets (Old)

Model	Accuracy	Precision	Recall	F1 Score	AUC
Decision Tree	0.815217	0.876289	0.794393	0.833333	0.819274
Extra Trees	0.896739	0.923077	0.897196	0.909953	0.941558
AdaBoost	0.853261	0.908163	0.831776	0.868293	0.914735

Table 6. Table of Risk Stratification

Index	AdaBoost	Decision Tree	Extra Trees
0	1	0	0
1	2	3	3
2	1	0	0
3	1	0	0

4	2	3	3
...	...	...	...
379	2	3	3
380	2	3	3
381	2	0	1
382	1	0	0
383	2	3	3

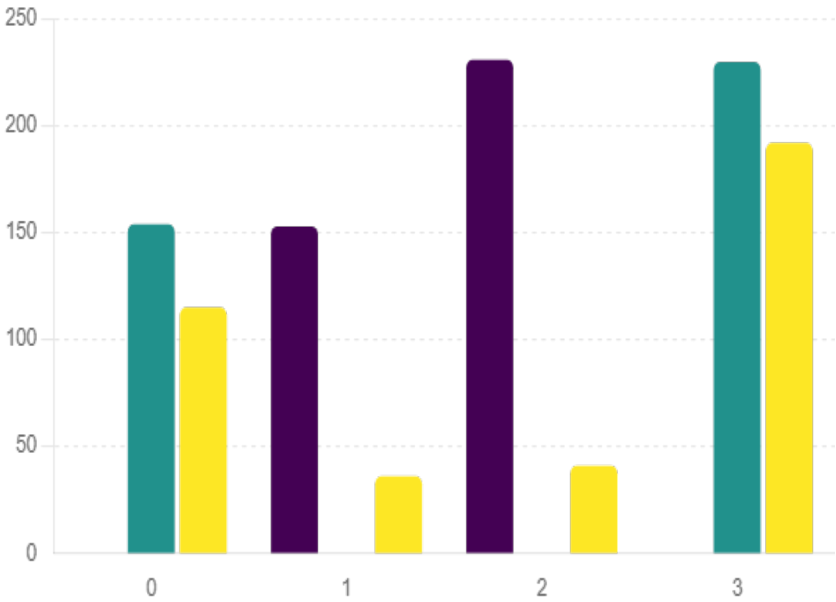

Figure 15. Graph of Risk Stratification

The risk stratification graph as shown in figure 3.12 compares the classifiers i.e. Decision Tree, AdaBoost, and Extra Trees to estimate patient risk:

3.6.1. AdaBoost Classifier:

- Low Risk (0): Few patients; broad disease spectrum
- Medium-Low Risk (1): Significant portion of patients
- Medium-High Risk (2): Fewer patients than Medium-Low

3.6.2. Decision Tree Classifier:

- Low Risk (0): Majority of patients; indicating overconfidence
- Medium-Low Risk (1): Current patient group
- Medium-High Risk (2): Patients in this category are quite uncommon
- High Risk (3): Patients in high-risk category

3.6.3. Extra Tree Classifier

- Low Risk (0): Most patients in this category similar to Decision Tree
- Medium-Low Risk (1): Substantial portion of patients
- Medium-High Risk (2): More patients than Decision Tree
- High Risk (3): Reasonable assessment, with a notable portion in high-risk

Table 7. Table of Uncertainty quantifications for classifiers

AdaBoost	Decision Tree	Extra Trees
0.6142116	-1.00E-09	-1.00E-09
0.6930867	-1.00E-09	0.471393485
0.6918657	-1.00E-09	0.198515241
0.6927268	-1.00E-09	0.098039111
0.6910662	-1.00E-09	-1.00E-09
...	...	...
0.6928059	-1.00E-09	0.253638945

0.6901559	-1.00E-09	0.198515241
0.6928886	-1.00E-09	0.673011665
0.6925867	-1.00E-09	0.6108643
0.6931453	-1.00E-09	0.598945906

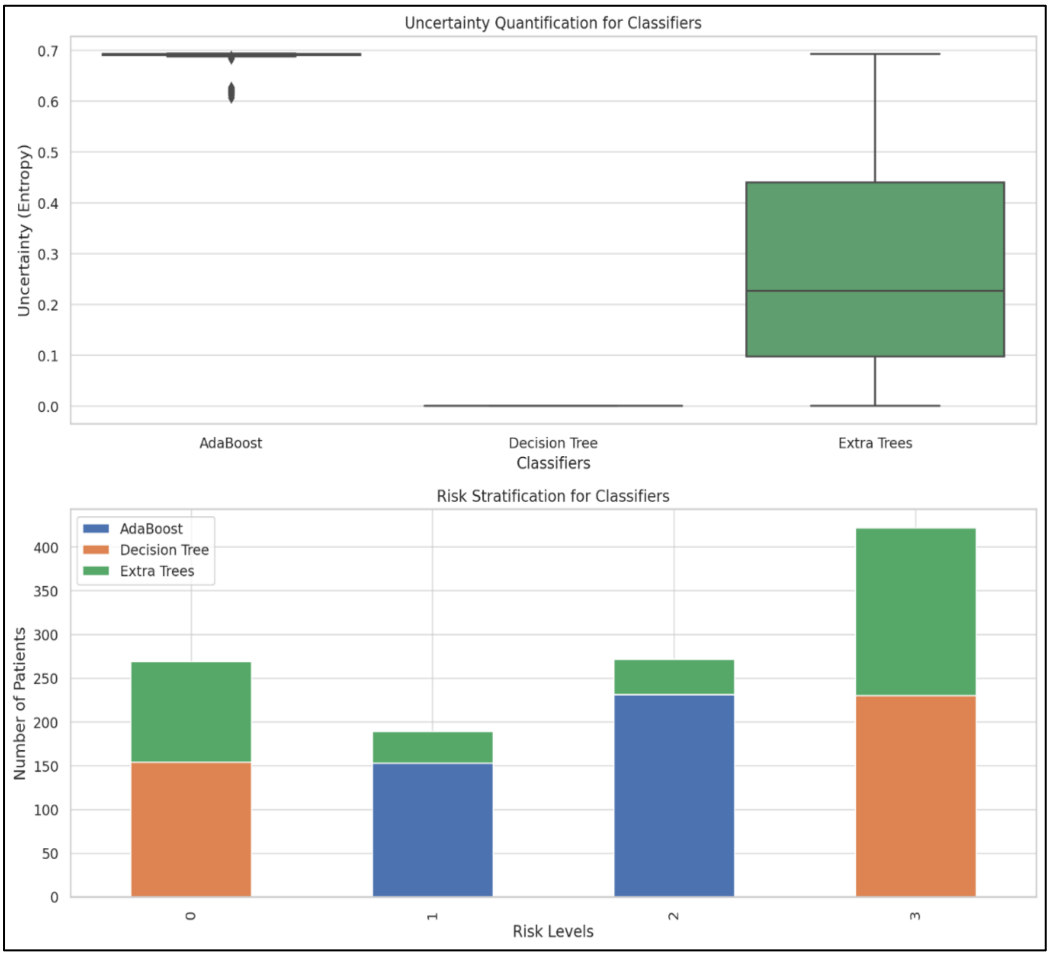

Figure 16. Uncertainty quantifications and Risk Stratification for classifiers

The uncertainty quantification graph as shown in Figure14 had shown the distribution of prediction uncertainty (calculated as entropy) for three classifiers: AdaBoost, (DT) Decision Tree, and Extra Trees. Low values of entropy mean better assurance in predictions, while higher entropy values mean more uncertainty.

- **AdaBoost Classifier:** Shown moderate uncertainty with substantial entropy in some predictions, an instable spread, and a few high-uncertainty outliers, demonstrating occasional lower confidence in predictions
- **Decision Tree classifier:** Exhibited a wide range of uncertainty with very low and very high values, high confidence in most predictions, but infrequent unpredictable predictions with high uncertainty
- **Extra Tree Classifier:** Demonstrated balanced and consistent uncertainty with a slight spread and low median, representing stable confidence across predictions and fewer plain outliers

3.7. Optimization

GridSearchCV optimization greatly increased the prediction accuracy of classifiers like AdaBoost, Decision Tree, and Extra Trees. For instance, the Extra Trees classifier obtained 99% accuracy on a new dataset of cardiovascular disorders following optimization. However, the improved models also shown ability over fitting, with slightly less generalization on older heart disease records. This highlighted the necessity of stabilizing model complexity through generalization in order to prevent accuracy improvements from endangering the model's overall performance on unidentified data.

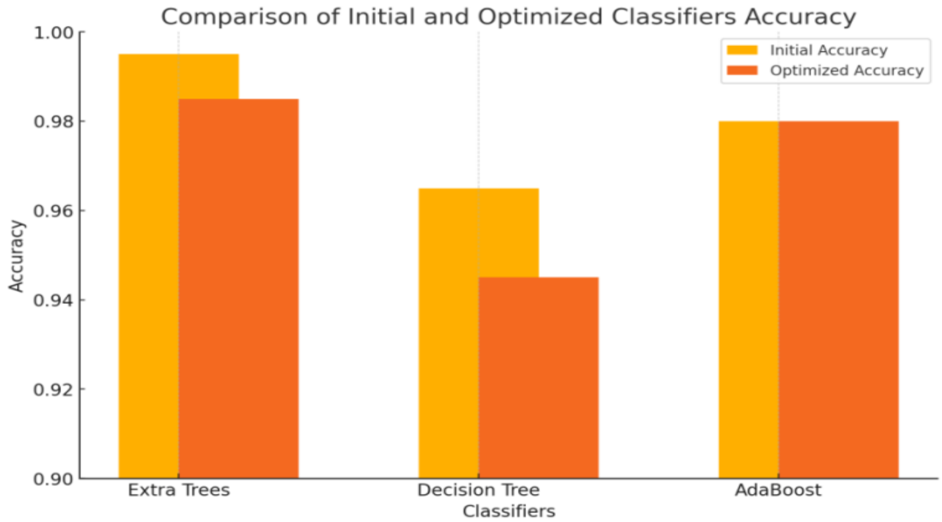

Figure 17. Comparison of Initial and optimized classifiers Accuracy for 1st dataset

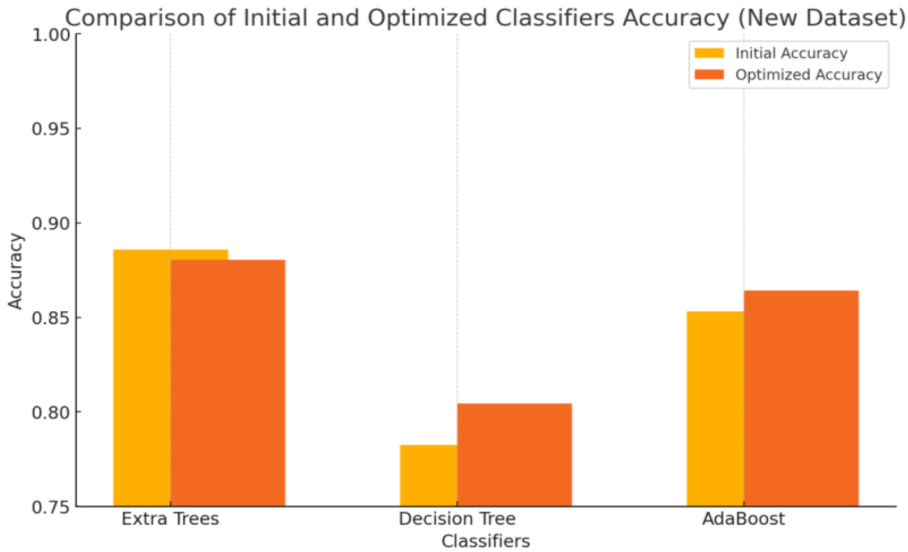

Figure 18. Comparison of Initial and optimized classifiers Accuracy 2nd dataset

The comparison of initial and optimized classifier revealed different patterns between both datasets. In the 1st dataset figure 17, extra trees have a slight decrease in accuracy after optimization, perhaps due to over fitting correction, while decision trees have a more significant decline. However, accuracy of AdaBoost increases after optimization. In the 2nd dataset, all classifiers, Decision Tree and AdaBoost, in particular significantly enhance put up-optimization (Figure 18), highlighting the importance of optimization for better performance on difficult datasets.

Table 8. Comparison of Initial and optimized Classifiers:

Classifier	Initial Accuracy	Optimized Accuracy
Extra Trees Classifier	0.995	0.985
Decision Tree Classifier	0.965	0.945
AdaBoost Classifier	0.98	0.98

The ROC Curves as shown in (Figures 19 & 20), demonstrated that, with small optimization improvements and high AUC values, the models' performance had improved on the 1st dataset. While optimization produces small increases, particularly for the Decision Tree, performance decreases on the 2nd dataset. This highlights the significant impact of datasets characteristics and the limited effect of optimization on model’s performance.

ROC Curves for Initial and Optimized Classifiers for both datasets:

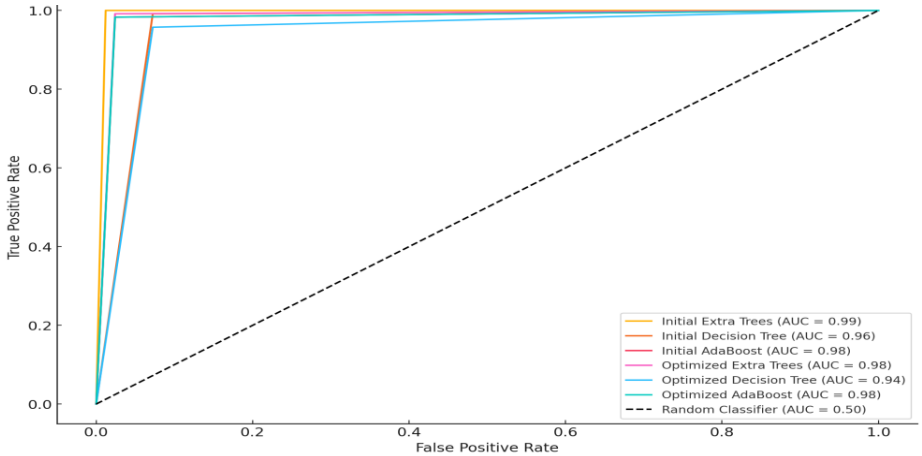

Figure 19. ROC Curve for Initial and Optimized Classifiers for 1stdataset

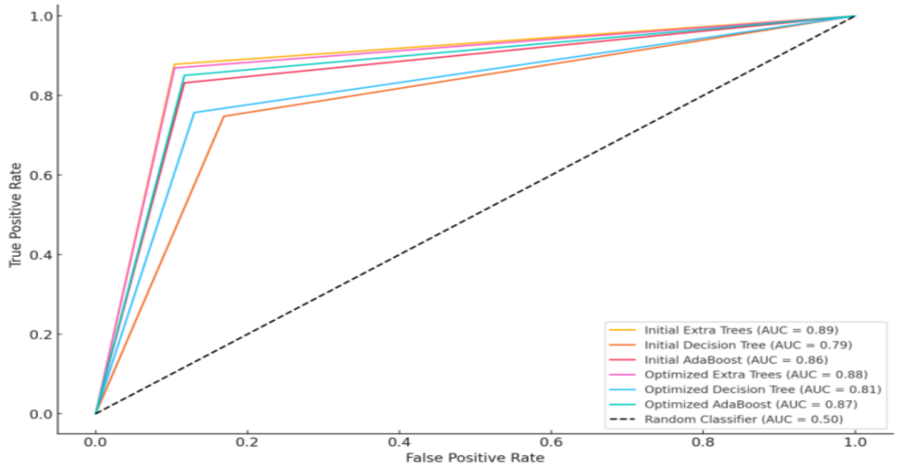

Figure 20. ROC Curve for Initial and Optimized Classifiers for 2nddataset

3.8. Sensitivity Analysis

This analysis using permutation importance revealed that 'Slope' and 'ST_Slope' were key parameters for predicting coronary heart disease and cardiovascular disease across AdaBoost, Extra Trees, and Decision Tree models. In the 1stdataset, 'ST_Slope' emerged as the most critical feature, with 'Fasting Blood Sugar' and 'Cholesterol' also showing significant influence. Other features like Chest Pain Type and Old Peak were moderately important, while Age, Sex, and Resting Blood Pressure were less relevant. In the 2nd dataset, 'Slope' was the most crucial feature, particularly in the Extra Trees model, followed by 'Resting Electrocardiographic Results' and 'Chest Pain Type'. Figures 3.18 and 3.19 illustrated these findings, highlighting the importance of these features in enhancing model predictions.

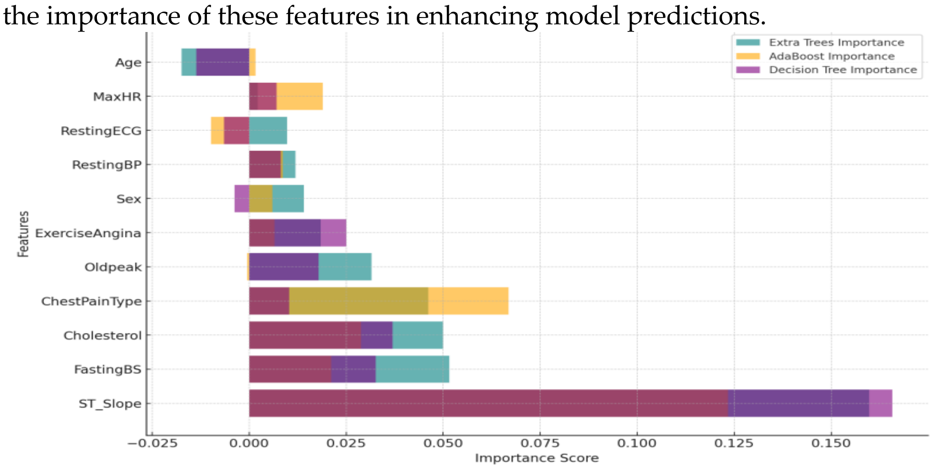

Figure 21. Sensitivity Analysis: Features importance Across Models (1st Dataset)

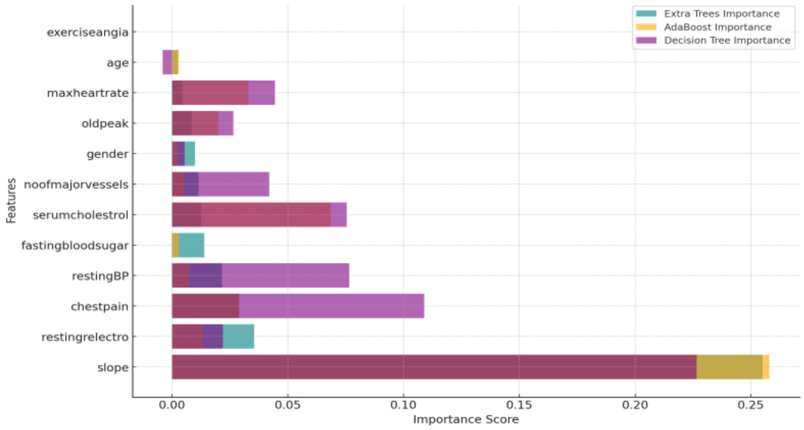

Figure 22. Sensitivity Analysis: Features importance Across Models (2nd Dataset)

4. Conclusions

This study highlights how machine-learning techniques can be used to better assess risk and predict cardiac abnormalities. By employing state-of-the-art methods such as AdaBoost, Decision Trees, and Extra Trees and optimizing them with GridSearchCV, we were able to achieve a considerable boost in prediction accuracy. The accuracy of diagnosis has significantly improved as a result of using machine learning to predict cardiac issues. Our models demonstrated robust performance on two distinct datasets, proving the versatility and reliability of our machine learning model for deployment in medical hospitals and clinics. On 1st dataset, the Decision Tree received 81.5%, the Extra Tree received 89%, and AdaBoost received 85.3%. On the 2nddataset on cardiovascular sickness, AdaBoost scored 98%, Decision Tree scored 96.5%, and Extra Tree scored 99%, demonstrating even higher accuracy. These results indicate how machine learning can be applied to identify individuals who have a higher tendency than others to have heart problems in the future, offering an effective means of early identification and treatment.

References

1. Nashif, S. Heart disease detection by using machine learning algorithm and a real-time cardiovascular health monitoring system. World J. Eng. Technol. 2018, 6, 88-650.
2. Kohli, P.S.; Arora, S. Application of machine learning in diseases prediction. In Proceedings of the 4th International Conference on Computing Communication and Automation (ICCCA), Greater Noida, India, 14-15 December 2018; IEEE: Piscataway, NJ, USA, 2018.
3. Chicco, D.; Jurman, G. Machine learning can predict survival of patients with heart failure from serum creatinine and ejection fraction alone. BMC Med. Inform. Decis. Mak. 2020, 20(16), 1-16.
4. Yazdani, A.; Varathan, K.D.; Chiam, Y.K. A novel approach for heart disease prediction using strength scores with significant predictors. BMC Med. Inform. Decis. Mak. 2021, 21, 194.
5. Gour, S.; Panwar, P.; Dwivedi, D.; Mali, C. A machine learning approach for heart attack prediction. Intell. Sustain. Syst. 2022, 2555(1), 741-747.
6. Doppala, B.P.; Bhattacharyya, D.; Janarthanan, M.; Baik, N. A reliable machine intelligence model for accurate identification of cardiovascular diseases using ensemble techniques. J. Healthc. Eng. 2022, 2022, 1-13.
7. Aljanabi, M.; Qutqut, M.H.; Hijjawi, M. Machine learning classification techniques for heart disease prediction: a review. Int. J. Eng. Technol. 2018, 7(4), 5373-5379.
8. Saravanan, S.; Tirumurugan, P. Performance analysis of glioma brain tumor segmentation using ridgelet transform and CANFES methodology. J. Med. Imaging Health Inform. 2020, 10(11), 2642-2648.
9. Swathy, M.; Saruladha, K. A comparative study of classification and prediction of cardiovascular diseases (CVD) using machine learning and deep learning techniques. ICT Express 2022, 8(1), 109-116.
10. Sharma, A.M.; Gupta, A.; Kumar, P.K.; Rajan, J.; Saba, L.; Nobutaka, I.; Laird, J.R.; Nicolades, A.; Suri, J.S. A review on carotid ultrasound atherosclerotic tissue characterization and stroke risk stratification in machine learning framework. Curr. Atheroscler. Rep. 2015, 17, 1-3.
11. Tao, R.; Zhang, S.; Huang, X.; Tao, M.; Ma, J.; Ma, S. Magnetocardiography-based ischemic heart disease detection and localization using machine learning methods. IEEE Trans. Biomed. Eng. 2019, 66(6), 1658-1667.
12. Kavitha, M.; Gnaneswar, G.; Dinesh, R.; Sai, Y.R.; Suraj, R.S. Heart disease prediction using hybrid machine learning model. In Proceedings of the 6th International Conference on Inventive Computation Technologies (ICICT), Coimbatore, India, 20 January 2021; IEEE: Piscataway, NJ, USA, 2021; pp. 1329-1333.
13. Ouyang, S. Research of heart disease prediction based on machine learning. In Proceedings of the Fifth International Conference on Advanced Electronic Materials, Computers and Software Engineering (AEMCSE), 2022; pp. 315-319.
14. Ahmad, E.; Tiwari, A.; Kumar, A. Cardiovascular diseases (CVDs) detection using machine learning algorithms. Int. J. Res. Appl. Sci. Eng. Technol. 2020, 8(6), 2341-2346.
15. Haq, A.U. A hybrid intelligence system framework for prediction of heart disease using ML. Mobile Inf. Syst. 2021, 2021, 3860146.
16. Alom, Z.; Early-stage detection of heart failure using machine learning techniques. In Proceedings of the International Conference on Big Data, IoT, and Machine Learning (BDIoTML), 2021; pp. 75-88.
17. Nadeem, W.; Goh, H.G.; Khan, M.A.; Hussain, M.; Mushtaq, M.F.; Ponnusamy, V. Fusion-based machine learning architecture for heart disease prediction. Comput. Mater. Continua 2021, 67(2), 2481-2496.
18. Rubini, P.E.; Subasini, C.A.; Katharine, A.V.; Kumaresan, V.; Kumar, S.G.; Nithya, T.M. A cardiovascular disease prediction using machine-learning algorithms. Ann. Romanian Soc. Cell Biol. 2021, 904-912.
19. Absar, N.; Das, E.K.; Shoma, S.N.; Khandaker, M.U.; Miraz, M.H.; Faruque, M.R.; Tamam, N. The efficacy of machine-learning-supported smart system for heart disease prediction. Healthcare 2022, 10(6), 1137.
20. Yeh, C.H.; Muo, C.H.; Sung, F.C.; Yen, P.S. Risk of ischemic heart disease associated with primary dysmenorrhea: a population-based retrospective cohort study. J. Pers. Med. 2022, 12(10), 1610.
21. Jothikumar, R.; Sivakumar, N.; Ramesh, P.S. Heart disease prediction system using ANN, RBF, and CBR. Int. J. Pure Appl. Math. 2017, 117(21), 199-217.
22. Gjoreski, M.; Gradišek, A.; Gams, M.; Simjanoska, M.; Peterlin, A.; Poglajen, G. Chronic heart failure detection from heart sounds using a stack of machine-learning classifiers. In Proceedings of the 13th International Conference on Intelligent Environments (IE), Seoul, South Korea, 21-24 August 2017; IEEE: Piscataway, NJ, USA, 2017; pp. 14-19.
23. Ravindhar, N.; Hariharan, R.S. Intelligent diagnosis of cardiac disease prediction using machine learning. Int. J. Innov. Technol. Explor. Eng. 2019, 9(11), 1417-1421. Khan, M. F., Iftikhar, A., Anwar, H., & Ramay, S. A. (2024). Brain Tumor Segmentation and Classification using Optimized Deep Learning. Journal of Computing & Biomedical Informatics, 7(01), 632-640.

24. Iftikhar, A., Elmagzoub, M. A., Shah, A. M., Al Salem, H. A., ul Hassan, M., Alqahtani, J., & Shaikh, A. (2023). Efficient Energy and Delay Reduction Model for Wireless Sensor Networks. Comput. Syst. Sci. Eng., 46(1), 1153-1168.
25. Sajjad, R., Khan, M. F., Nawaz, A., Ali, M. T., & Adil, M. (2022). Systematic analysis of ovarian cancer empowered with machine and deep learning: a taxonomy and future challenges. Journal of Computing & Biomedical Informatics, 3(02), 64-87.
26. Weng, S.F.; Reps, J.; Kai, J.; Garibaldi, J.M.; Qureshi, N. Can machine learning improve cardiovascular risk prediction using routine clinical data? PLoS One 2017, 12(4), e0174944.
27. Reddy, S.K.; Meghana, P.; Reddy, N.V.; Rao, B.A. Prediction on cardiovascular disease using decision tree and naïve Bayes classifiers. J. Phys. 2022, 2161, 1-8.