Heart Disease Detection And Classification Using Ensemble Technique

Priyanshi Agrawal[1], Priyanshi Rajawat[2], Akshara Sharma[3], Yasha Istwal[4]
[1-2-3-4]*Computer Science & Engineering (AI&DS)G.L. Bajaj Institute of Technology and Management Greater Noida, India*

***Abstract*—cardiovascular diseases (CVD) account for a substantial proportion of deaths worldwide, underlining the necessity of accurate diagnostic methods for early intervention. This study offers a ML framework for CVD prediction, using an openly accessible data set comprising clinical and demographic data. The process involves refining the dataset, identifying important predictors, and benchmarking several classification techniques such as tree-based approaches like Random Forest and advanced boosting techniques such as Gradient Boosting (GB) and XGBoost (XGB). A soft-voting ensemble model was implemented to enhance prediction accuracy, with performance evaluated using metrics like measures of correctness (accuracy), relevance (precision), completeness (recall), harmonic mean (F1 score), and discriminative ability (ROC-AUC). The results demonstrate the potential of ensemble techniques in improving diagnostic efficiency, paving the way for integration into clinical decision-making systems.**

I. INTRODUCTION

Cardiovascular diseases stand as a predominant global cause of mortality, contributing substantially to worldwide death statistics. According to WHO, CVDs contribute to approximately 17.9 million deaths annually, underscoring the critical importance of early detection in mitigating the impact of heart disease (WHO, 2021) [1]. Traditional diagnostic methods often lack the precision required for early and accurate detection, leading researchers to explore advanced machine-learning techniques to improve diagnostic accuracy (Smith et al., 2018) [2].

Machine learning, particularly in healthcare, has demonstrated promising potential in recent years. It enables the analysis of large datasets to generate predictions that aid in early disease detection (Johnson & Patel, 2019) [3]. Ensemble learning, a methodological approach that integrates multiple machine learning models to enhance predictive robustness, has gained recognition as an effective strategy for boosting model accuracy and minimizing overfitting risks (Dietterich, 2000) [4]. Numerous studies have demonstrated that ensemble methods, including tree-based approaches (Random Forest) and advanced boosting technique (Gradient Boosting), consistently surpass the performance of individual classifiers in classification and predictive tasks (Zhou, 2012) [5].

This study concentrates on leveraging ensemble machine-learning methodologies to inflate the relevance (precision) & effectiveness of CVD detection and classification. The objective of employing classifiers including tree-based approaches (RF), advanced boosting technique (Gradient Boosting), and Voting Classifiers is to assess whether ensemble methods can provide superior accuracy compared to single-model approaches. The study uses publicly available heart disease datasets, and pre-processing steps including feature selection and data cleaning are performed to ensure high-quality input data (Smith et al., 2018) [2].

The outcomes of this research are anticipated to enrich the expanding corpus of evidence advocating the usage of ML in health sector, with a specific emphasis on early detection and classification of CVD. Ensemble models could conceivably augment the reliability and accuracy of diagnostic systems, which could aid in reducing the impact of heart diseases globally (Zhou, 2012) [5].

II. LITERATURE REVIEW

The field of heart disease prognosis has seen notable strides through the adoption of ensemble machine-learning techniques, with several studies contributing to improved accuracy, robustness, and interpretability. Mondal et al. came up with a dual-stage stacking ensemble model integrating tree-based approaches (Random Forest), advanced boosting technique

Extreme Gradient Boosting (XGBoost), and Decision Tree (DT) classifiers as base learners, with XGBoost acting as the meta-classifier. The study demonstrated the benefits of a two-stage architecture, achieving 96% accuracy, a completeness i.e. recalls of 0.98 and discriminative ability i.e. ROC-AUC score of 0.96 on the primary dataset, while validation on a secondary dataset yielded an accuracy of 96.88%. Despite its robustness, the dual-stage stacking approach increases computational complexity, a potential limitation in resource-constrained environments. [7]

Chaudhari et al. drew attention to the importance of feature selection and explored multiple machines learning models, highlighting Random Forest's efficacy due to its capacity for dealing with complex datasets. The study applied a voting-based feature selection technique and attained the measures of correctness (accuracy) of 90.63% and a sensitivity of 95.12%, showcasing the critical aspect of feature selection for enhancing interpretability and performance. [8]

Asif et al. validated the importance of hyperparameter optimization, demonstrating that the Extra Tree Classifier (ETC), when tuned using RandomizedSearchCV and GridSearchCV, achieved measures of correctness i.e. accuracy of 98.15% and harmonic mean i.e. F1-score of 98.10%, outperforming other models such as XGBoost and CatBoost. [9]

Reddy et al. introduced a hybrid stacking ensemble approach combining gradient boosting models (LightGBM, XGBoost) and neural networks (MLP) with Logistic Regression as the meta-classifier. The proposed model produces measures of correctness i.e. accuracy of 95.8% and a harmonic mean i.e. F1-score of 96.2%, outperforming standalone classifiers. However, integrating neural networks contributed to increased computational demands, emphasizing the balance between precision and resource utilization. [10]

El-Sofany et al. focused on model interpretability by integrating SHAP with XGBoost and applied SMOTE to handle class imbalance. They achieved the measures of correctness (accuracy) of 97.57%, a relevance (precision) of 97.91%, and discriminative ability (ROC-AUC) of 98%. Their study validated the inclusion of SHAP for feature importance analysis and SMOTE for balancing datasets, both of which are utilized in the current research. However, the study's limited dataset size may have impacted the external validity of the results. [11]

Mohan et al. put forward a hybrid approach, HRFLM that combined Random Forest and Linear Models to acquire both non-linear and linear relationships in the data. This method brought about an accuracy of 88.7% and a sensitivity of 92.8%, although its performance was limited by the lack of advanced feature selection or ensemble stacking. [12]

Velusamy and Ramasamy proposed a heterogeneous ensemble model combining RF, SVM, and k-NN, a popular machine learning algorithm (k-Nearest Neighbors) which achieved a perfect accuracy of 100% on a dataset with 303 records. Their work highlighted the effectiveness of combining diverse classifiers to leverage their unique strengths. The use of feature selection and voting mechanisms, as well as the application of SMOTE for class imbalance, has influenced the current study's design, particularly in the context of feature importance and ensemble learning techniques. [13]

Mienye et al. emphasized the importance of both accuracy and explainability by utilizing Bayesian optimization to fine-tune an XGBoost model and integrate SHAP for feature interpretation. They achieved the accuracy (measure of correctness) of 98.4% on the Cleveland dataset, illustrating the power of optimization and explainability in heart disease prediction. Their approach to optimizing hyperparameters and integrating SHAP inspired the current study's focus on both model performance and interpretability. Nonetheless, the study's heavy reliance on structured data and computationally expensive SHAP analysis may limit its scalability to larger datasets. [14]

Nissa et al. presented a technical comparison of boosting ensemble techniques, with AdaBoost emerging as the most effective algorithm due to its superior generalization capabilities. AdaBoost achieved an accuracy of 95%, excelling in specificity and resilience to overfitting. Their study provided valuable insights into boosting methods, but it did not explore other ensemble architectures like stacking or heterogeneous classifiers, leaving room for further exploration in the current study. [15]

These studies collectively highlight the advancements in heart disease prediction through the coalescence of diverse ensemble methods, feature selection approach, hyperparameter optimization, and explainable AI.

III. METHODOLOGY

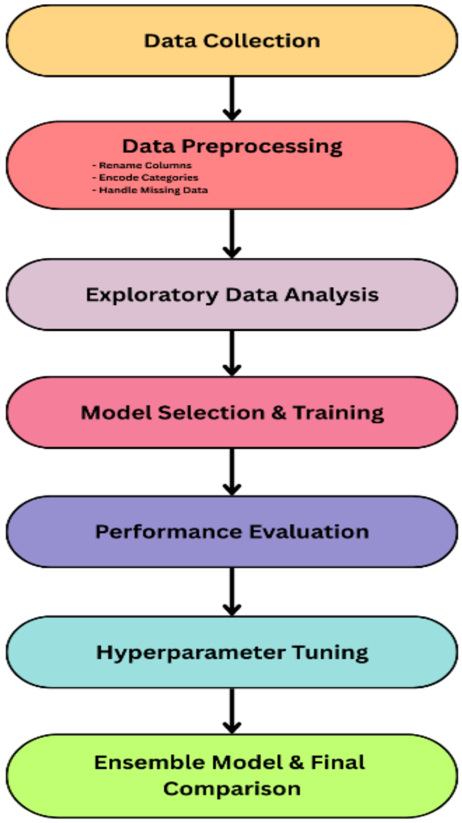

Fig I — Flowchart representing the methodology

Dataset Description:

Dataset utilized in this study is a curated collection derived from multiple heart disease datasets, including the Cleveland and Hungary subsets of the UCI Heart Disease Repository. It consists of structured clinical data encompassing 11 input features and 1 binary target variable, denoting the existence (1) or non-existence (0) of CVD.

Features include both categorical and continuous variables, capturing essential patient metrics such as:

Demographics: age, sex

Medical Measures: resting blood pressure(rbp), cholesterol(chl), max heart rate achieved(mhra), st depression(stdep)

Symptoms and Indicators: chest pain type(cpt), fasting blood sugar(fbs), rest ecg, exercise induced angina(eia), st slope(stsl)

The dataset comprises approximately 900 records, ensuring a balanced representation of both healthy and heart-affected individuals. Each observation is a unique patient case, and the dataset exhibits a moderately balanced class distribution, as revealed through exploratory visualization.

The diversity and medical relevance of the features make the dataset highly suitable for supervised classification tasks aimed at predicting cardiovascular risk. Prior to training, this raw dataset underwent comprehensive preprocessing to convert categorical labels, eliminate invalid values, and prepare the data for machine learning model input.

Target Variable: The target is a binary classification label indicating heart disease status: 1 signifies the existence of heart disease, while 0 indicates its non-existence. This outcome variable is used to train supervised models for predicting cardiovascular risk.

Data Preprocessing:

The dataset used in this study comprises clinical features, including age, gender, cholesterol levels, resting BP, maximum heart rate achieved, and various categorical indicators like cpt, resting ecg results, and stsl. The original dataset underwent several preprocessing steps:

Column Renaming: All feature names were renamed for readability and standardization.

Categorical Encoding: Nominal categorical features such as cpt, rest ecg, and stsl were mapped to descriptive string labels and subsequently encoded using one-hot encoding to preserve the non-ordinal relationships.

Binary Transformation: Features such as sex and exercise_induced_angina were converted into binary indicators.

Missing and Invalid Value Handling: Entries with undefined or invalid slope values were removed, and the dataset was checked for null values.

Feature Distribution Check: Exploratory Data Analysis (EDA) was conducted using visualization techniques to assess class balance and distribution of features across target classes.

Exploratory Data Analysis (EDA):

Target Variable Distribution

The dataset exhibits balanced distribution of the binary target var. Out of the total instances, 629 patients were diagnosed with heart disease (positive class), and 561 patients were not (negative class). This balance is advantageous for supervised learning models, as it minimizes the risk of biased training outcomes.

Demographic Feature Distribution

The demographic attributes, including age and sex, were analyzed to understand the dataset's population characteristics. The gender distribution showed a higher proportion of male patients compared to females, which aligns with epidemiological trends in cardiovascular disease prevalence.

Feature Relationships and Correlation Analysis

To explore interdependencies among features and their influence on the target variable, several visual and statistical techniques were applied:

Correlation Heatmap: A Pearson correlation matrix was plotted to highlight linear relationships between numerical features. Notably, ST Depression exhibited a strong positive correlation with heart disease presence, while Max Heart Rate Achieved showed a moderate negative correlation.

Distribution Plots: Univariate distributions of numerical features such as Cholesterol, Resting BP, and Max Heart Rate were visualized using histograms. These plots helped identify skewed variables and outliers, particularly in cholesterol values, which showe a right-skewed distribution.

Categorical Comparisons: Count plots for variables like CPT, Resting ECG, and STSL were generated across target classes. These visualizations revealed that certain categories, such as 'asymptomatic' chest pain and 'flat' ST slope, were more prevalent among heart disease patients.

Class Balance and Feature Influence

An essential part of EDA involved ensuring that the target classes were balanced and no feature introduced significant bias. The visualization results indicated that several clinical features exhibited discriminative potential for predicting cardiovascular outcomes, justifying their inclusion in the subsequent modeling phase.

Model Selection and Training

A broad spectrum of supervised ML algo was selected to ensure a complete comparative analysis of classification performance. Chosen models included both linear and non-linear classifiers, tree-based ensembles, svm, and neural network. Specifically, the models implemented were:

Logistic Regression – a baseline linear classifier

Decision Tree Classifier – a non-linear model based on recursive partitioning

Random Forest Classifier – a bagging ensemble of decision trees

Gradient Boosting Classifier – a boosting ensemble focusing on residual minimization

XGBoost Classifier – an optimized implementation of gradient boosting with regularization

Extra Trees Classifier – a randomized ensemble variant of decision trees

SVM – a kernel-based margin classifier

SGD Classifier – a linear model optimized using stochastic gradient descent

Multilayer Perceptron (MLP) – a feedforward neural network trained via backpropagation

All models were implemented using Scikit-learn and XGBoost libraries. Before training, dataset was split into 80% training and 20% testing parts using stratified sampling to preserve the class distribution across both sets.

Evaluation Metrics:

To estimate the efficacy of developed ML models, the following assessment metrics were employed. These metrics give a thorough evaluation of the classifiers, with a particular focus on healthcare-related prediction tasks:

Acc:

$$\text{Acc} = \frac{\text{True Positives} + \text{True Negatives}}{\text{Total Samples}}$$

Prec and Rec:

$$\text{Prec} = \frac{\text{True Positives}}{\text{True Positives} + \text{False Positives}}$$

$$\text{Rec} = \frac{\text{True Positives}}{\text{True Positives} + \text{False Negatives}}$$

F1 Score:

$$\text{F1S} = 2 \times \frac{\text{Precision} \times \text{Recall}}{\text{Precision} + \text{Recall}}$$

ROC-AUC:

$$\text{TPR} = \frac{\text{True Positives}}{\text{True Positives} + \text{False Negatives}}$$

$$\text{FPR} = \frac{\text{False Positives}}{\text{False Positives} + \text{True Negatives}}$$

Hyperparameter Tuning:

To enhance model performance, hyperparameter tuning was conducted using GridSearchCV with 5-fold cross-validation. This approach systematically explored predefined parameter grids for the top-performing models: XGB, RF, and Extra Trees. The tuning focused on key parameters, including the number of estimators, max depth, learning rate, and sampling ratios.

This tuning significantly improved model generalization, with the XGBoost classifier achieving an F1S of 0.953 and an ROC-AUC of 0.974 -the highest among all models.

Final Ensemble and Comparison:

To enhance predictive robustness, a soft voting ensemble was created by combining the probabilistic outputs of the top three tuned models—XGBoost, Random Forest, and Extra Trees. This approach leverages model strengths, reduces variance, and stabilizes performance.

After tuning and ensembling, all models were evaluated using metrics like Accuracy, Precision, Recall, F1-Score, and ROC-AUC. The tables below summarize their performance.

S.No.	Model	Acc	Prec	Rec	F1S	ROC AUC
1	XGBoost	0.950	0.938	0.968	0.953	0.967
2	Extra Trees	0.950	0.960	0.944	0.952	0.982
3	Random Forest	0.950	0.960	0.944	0.952	0.975
4	Voting Ensemble	0.945	0.945	0.952	0.949	0.978
5	Gradient Boosting	0.920	0.902	0.952	0.927	0.961
6	Logistic Regression	0.861	0.855	0.889	0.872	0.940
7	MLP Classifier	0.853	0.876	0.841	0.858	0.923
8	Decision Tree	0.845	0.874	0.825	0.849	0.846
9	SVM	0.718	0.761	0.683	0.720	0.796
10	SGD Classifier	0.588	0.967	0.230	0.372	0.860

Table I — Performance of tuned models and the final ensemble classifier

IV. RESULT

Tuned XGBoost –

The tuned XGBoost model demonstrated exceptional performance with confusion matrix showing a strong true positive rate, while the ROC curve highlights its excellent discriminative ability.

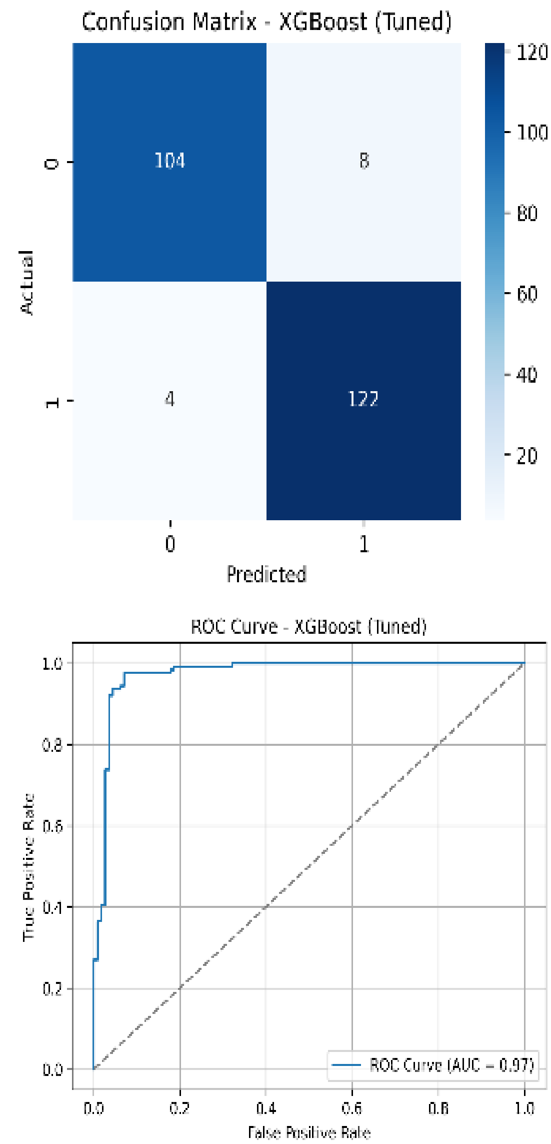

Fig II — Confusion matrix and ROC-AUC curve of Tuned XGBoost

The model correctly classified 104 negative and 122 positive cases, with only 12 total misclassifications, reflecting high sensitivity and specificity. The ROC curve shows excellent separability with an AUC of 0.97, indicating strong model performance in distinguishing between the classes.

Tuned Random Forest –

The confusion matrix confirms strong classification performance and ROC curve indicates reliable sensitivity and specificity.

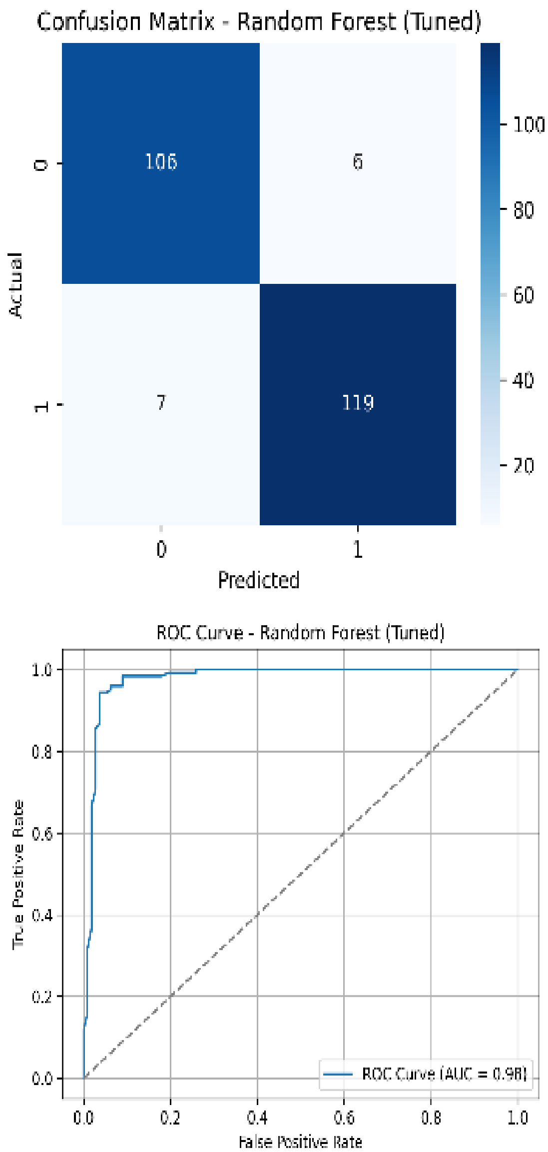

Fig III — Confusion matrix and ROC-AUC curve of Tuned Random Forest

The tuned Random Forest achieved 106 true negatives and 119 true positives, indicating reliable performance with only 13 total misclassifications. An AUC of 0.98 reflects high sensitivity and specificity, confirming strong overall performance of the model.

TUNED EXTRA TREES –

The Extra Trees model showed strong precision and recall. Its confusion matrix and ROC curve confirm reliable classification performance.

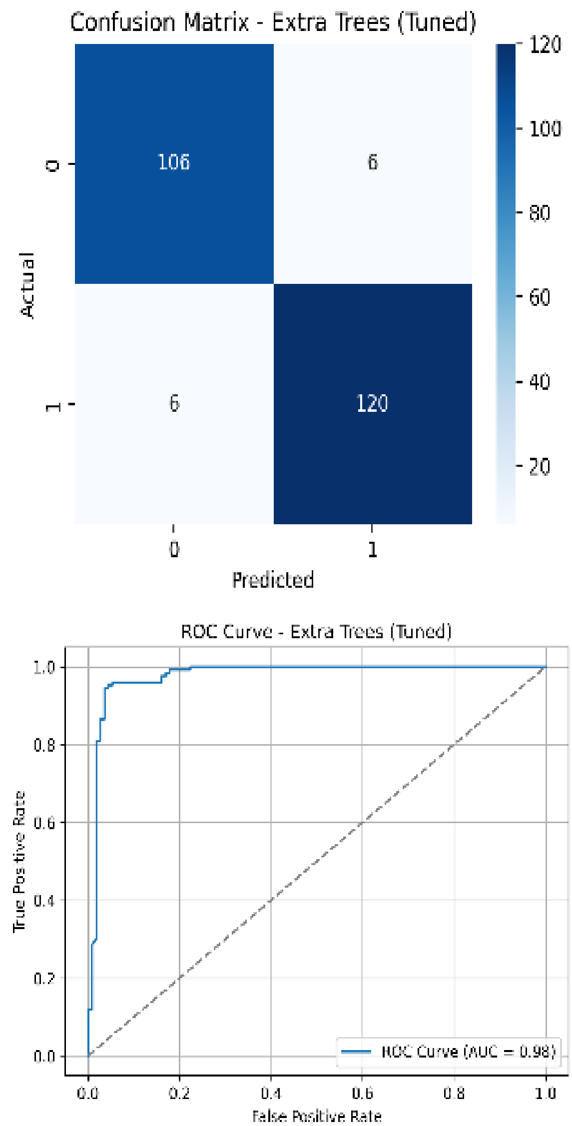

Fig IV — Confusion matrix and ROC-AUC curve of Tuned Extra Trees

The model accurately predicted 106 negatives and 120 positives, with just 6 errors in each class—showcasing balanced classification capability. With an AUC of 0.98, the Extra Trees model displays excellent discriminative power, maintaining low false positives across thresholds.

TUNED SOFT VOTING –

The ensemble model outperformed individual learners with improved generalization. Both confusion matrix and ROC curve reflect its robust predictive capability.

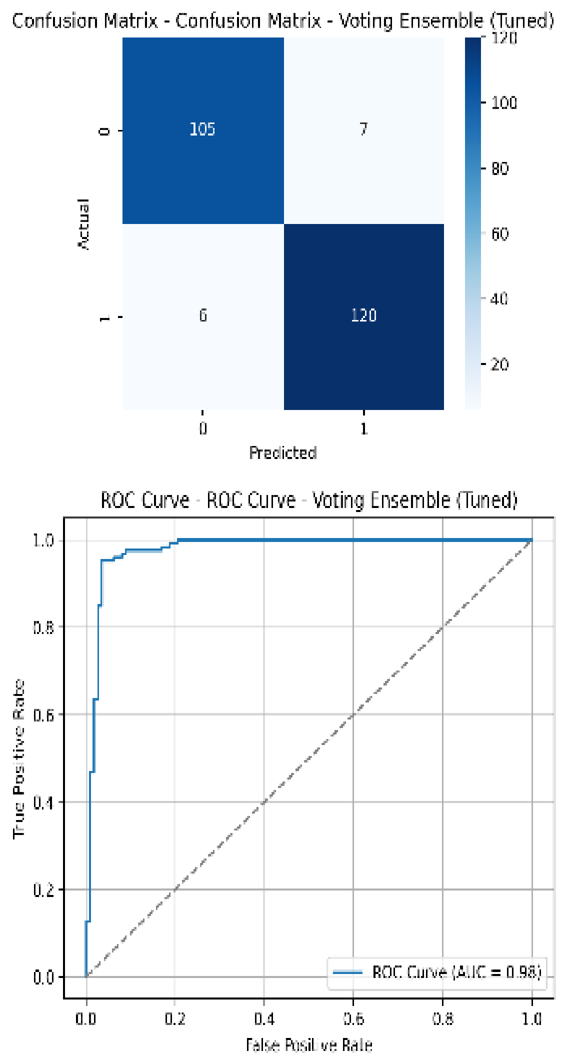

Fig V — Confusion matrix and ROC-AUC curve of Tuned Soft Voting

The ensemble model correctly classified 105 negatives and 120 positives, with a total of 13 misclassified samples—demonstrating robust generalization. The soft voting classifier attained an AUC of 0.98, reinforcing its capability to effectively aggregate strengths of base learners.

PERFORMANCE COMPARISON OF VARIOUS PAPERS – The table below presents key evaluation metrics—accuracy, precision, recall, and ROC AUC—from recent studies. Our tuned soft voting ensemble model outperformed prior methods on most metrics, highlighting its strong predictive power and robustness. This comparison emphasizes the benefits of combining hyperparameter tuning with ensemble techniques for heart disease risk assessment.

Ref.	Year	Publication	Method	Result
[15]	2023	IEEE	Soft Voting	Acc - 91.85 % & AUC - 0.9344
[16]	2022	MDPI	XGB boost	Acc - 94.7%, Rec - 96.1%, Prec – 93.4%, F1S – 94.6% and AUC - 98.0%
[17]	2023	MDPI	Soft Voting	Acc – 95%
[18]	2022	Journal of Healthcare Engineering	Soft Voting classifier	Acc - 93.39 %, Prec - 99%, Rec - 88%, F1S - 90%
[20]	2022	IEEE	Hard Voting & Soft Voting Classifier	Acc (HVS) - 83.2% & Acc (SVC) - 82.5%
[21]	2022	IEEE	Soft Voting & Hard Voting	Acc – .94.2%
		Our Proposed Paper	Soft Voting using XGB, RF & Extra tree	Acc – 95%, Prec – 94.5%, Rec – 95.2%, F1S – 95%, ROC-AUC – 98%

Table II — Performance Comparison of Various Papers

V. CONCLUSION

This study demonstrates the efficacy of supervised ML algo for prediction of CVD using a structured dataset bounding demographic, clinical, and exercise-induced features. A robust methodology was applied, beginning with data preprocessing—including feature renaming, label encoding, and invalid entry removal—followed by exploratory data analysis to reveal trends and correlations within the data.

A comprehensive set of models was implemented, including LR, DT, RF, GB, XGBoost, Extra Trees, SVC, SGD Classifier, and Multilayer Perceptron. Each model was evaluated using stratified cross-validation and assessed through multiple metrics: Acc, Prec, Rec, F1S, and ROC-AUC.

Hyperparameter tuning was performed on the top-performing models using GridSearchCV, resulting in significant improvements in classification performance. Among these, the tuned XGBoost classifier achieved the highest F1S of 0.953 and ROC-AUC of 0.974, confirming its strong generalization capability.

To further enhance stability and predictive robustness, a soft voting ensemble was constructed using XGBoost, Random Forest, and Extra Trees classifiers. This ensemble method maintained competitive performance attaining AUC of 0.98 while reducing model variance, making it suitable for real-world deployment.

Feature importance analysis highlighted key predictors such as ST Depression, Max Heart Rate Achieved, and Cholesterol, which align with clinical knowledge in cardiovascular diagnostics. These findings reinforce the potential of ensemble machine learning models to support early diagnosis and decision-making in healthcare systems.

This research lays the groundwork for deploying intelligent, data-driven systems in clinical environments to aid cardiologists in risk stratification and improve cardiovascular health outcomes.

VI. FUTURE SCOPE

Machine learning (ML) offers transformative potential in cardiovascular disease (CVD) prediction and classification. Despite advancements in deep learning for medical image analysis and ECG interpretation, challenges such as data privacy regulations, limited dataset diversity, and model robustness hinder progress. Enhancing transparency through Explainable Artificial Intelligence (XAI) is crucial for fostering clinical adoption by clarifying predictive features. The integration of ML with IoT devices, such as wearable health monitors, enables real-time cardiovascular monitoring, with Generative Adversarial Networks (GANs) addressing data scarcity and federated learning supporting decentralized, privacy-compliant training.

Big data frameworks are essential for processing high-resolution clinical data and enabling real-time insights. Tackling algorithmic biases and ensuring fairness is necessary for equitable healthcare. Ensemble strategies like stacking and blending enhance predictive accuracy by leveraging diverse models. Developing user-friendly tools and scalable applications for CVD management is key to applying these innovations clinically. Collaborative efforts among clinicians, data scientists, and technical experts will ensure these solutions are practical and globally accessible, ultimately improving patient outcomes and transforming CVD diagnostics.

REFERENCES

[1] WHO, “CVD” United Nations.

[2] N. Chandrasekhar, "Improving the accuracy of heart disease predictions using ml & optimization techniques," Processes, vol. 11, no. 4, Apr. 2023.

[3] S. Secinaro, D. Calandra, A. Secinaro, "AI in healthcare: A structured review of the literature," BMC Medical Informatics and Decision Making, vol. 21, no. 1, Dec. 2021.

[4] S. Džeroski, P. Panov, and B. Ženko, "ML: Ensemble methods," in Encyclopedia of Complexity and Systems Science, R. Meyers, Ed. New York, NY, USA: Springer, 2009.

[5] Z.-H. Zhou, Ensemble Methods. Boca Raton, FL, USA: Chapman and Hall/CRC, 2012.

[6] Subhash Mondal, R. Maity, Y. Omo, S. Ghosh, and A. Nag, "A two-stage stacked machine learning model for predicting heart disease risk," IEEE Access, vol. 12, pp. 7255–7270, 2024.

[7] C. S. Gautam, A. A. Waoo, S. Chaudhari, M. Chandra, and S. Gautam, "A comparative analysis of ensemble-based machine learning models for heart disease prediction accuracy improvement."

[8] D. Asif, M. Bibi, M. S. Arif, and A. Mukheimer, "Using ensemble learning and hyperparameter optimization for heart disease prediction," Algorithms, vol. 16, no. 6, June 2023.

[9] N. N. Reddy, L. Nipun, U. Baba, N. Rishindra, and T. Shilpa, "Application of ensemble and hybrid machine learning methods for enhancing

heart disease prediction," Int. J. Electr. Comput. Eng., vol. 14, no. 5, pp. 5744–5754, October 2024.
[10] H. F. El-Sofany, "Using machine learning with various classification methods for heart disease prediction," IEEE Access, vol. 12, pp. 106146–106160, 2024.
[11] S. Mohan, C. Thirumalai, and G. Srivastava, "Prediction of heart disease using hybrid techniques in machine learning," IEEE Access, vol. 7, pp. 81542–81554, 2019.
[12] D. Velusamy and K. Ramasamy, "Diagnosing coronary artery disease using an ensemble of diverse classifiers with optimized features," Comput. Methods Programs Biomed., vol. 198, January 2021.
[13] I. D. Mienye and N. Jere, "Heart disease prediction enhanced with optimized ensemble methods and interpretable AI," Information (Switzerland), vol. 15, no. 7, July 2024.
[14] N. Nissa, S. Jamwal, and M. Neshat, "Boosting ensemble-based techniques for a comparative heart disease prediction framework," Computation, vol. 12, no. 1, January 2024.
[15] K. Sen and B. Verma, "Heart disease prediction leveraging a soft voting ensemble of gradient boosting, random forest, and Gaussian Naive Bayes," in Proc. 2023 4th Int. Conf. Emerging Technology (INCET), Belgaum, India, 2023, pp. 1–7.
[16] S. Zhang, Y. Yuan, Z. Yao, X. Wang, and Z. Lei, "Enhancing coronary artery disease prediction through the XGBoost algorithm and advanced feature processing," Electronics (Switzerland), vol. 11, no. 3, February 2022.
[17] N. Chandrasekhar and S. Peddakrishna, "Enhancing heart disease prediction accuracy through machine learning techniques and optimization," Processes, vol. 11, no. 4, p. 1210, Apr. 2023
[18] B. P. Doppala, D. Bhattacharyya, M. Janarthanan, and N. Baik, "A reliable machine intelligence model for accurate identification of cardiovascular diseases using ensemble techniques," J. Healthc. Eng., vol. 2022, Article ID 2585235, 13 pages, Mar. 2022
[19] M. Özcan and S. Peker, "A classification and regression tree algorithm for heart disease modelling and prediction," Healthcare Analytics, vol. 3, p. 100130, Nov. 2023
[20] P. Vivek, D. Deepika, S. Pritpal, B. Rahul, and S. Saket, "Coronary heart disease prediction using voting classifier ensemble learning," in Proc. 3rd Int. Conf. Advances in Computing, Communication Control and Networking (ICAC3N), 2021, pp. 181–185
[21] S. Cyriac, R. Sivakumar, N. Raju, and Y. W. Kim, "Heart disease prediction using ensemble voting methods in machine learning," in Proc. 13th Int. Conf. Information and Communication Technology Convergence (ICTC), Jeju Island, South Korea, Oct. 2022