RESEARCH PAPER

An IoT-Enabled Explainable Stacked Ensemble Framework with SMOTE-ENN for Robust Heart Disease Prediction

Ritu Singh[1*], Shaheen Naz[2], Himanshu Payal[3], Harendra Singh[4]

[1*] Research Scholar, Department of Electrical Electronics and Communication Engineering, School of Engineering and Science, Sharda University, Greater Noida, India; ritu.singh@sharda.ac.in
[2] Assistant Professor, Department of Electrical Electronics and Communication Engineering, Centre of Cyber Security and Cryptology, Sharda University, Greater Noida, India; shaheen.naz@sharda.ac.in
[3]Assistant Professor, Department of Mechanical Engineering, School of Engineering and Science, Sharda University, Greater Noida, India; himanshu.payal@sharda.ac.in
[4] Assistant Professor, Department of Electronics & Communication Engineering, G.L. Bajaj Institute of Technology and Management, Greater Noida, Uttar Pradesh 201306, India; harendramvn2010@gmail.com

***Corresponding Author:** Ritu Singh
Research Scholar, Department of Electrical Electronics and Communication Engineering, School of Engineering and Science, Sharda University, Greater Noida, India; ritu.singh@sharda.ac.in

Abstract: Cardiovascular disease continues to be a primary cause of death globally, requiring precise and clinically interpretable predictive models. This paper presents an innovative IoT-enabled hybrid framework for predicting heart disease, incorporating advanced data balancing, feature optimization, ensemble learning, and explainable artificial intelligence (XAI). The proposed model utilizes SMOTE-ENN to rectify class imbalance, succeeded by a two-tier feature selection process that integrates Mutual Information and SHAP-based importance ranking. A heterogeneous ensemble comprising XGBoost, LightGBM, Random Forest, CatBoost, and TabNet classifiers is constructed, with their probabilistic outputs integrated through a logistic regression meta-learner. Comprehensive experiments performed on the UCI Heart Disease dataset and real time database collected from hospital demonstrate that the proposed method attains exceptional performance, with accuracy at 98.56%, sensitivity at 97.5%, specificity at 97%, AUC at 0.99, and Matthews Correlation Coefficient at 0.94 with 5-fold cross validation. Moreover, SHAP analysis offers clear global and local explanations, thereby improving clinical trust and interpretability. The findings validate that the suggested framework is both resilient and appropriate for practical healthcare decision-support systems.

Keywords: Heart Disease, SMOTE-ENN, CatBoost & TabNet, Machine Learning, Tree-Importance, Explainable AI, SHAP, and Stacked Ensemble learning.

How to cite this article: Singh R, Naz S, Payal H, Singh H, An IoT-Enabled Explainable Stacked Ensemble Framework with SMOTE-ENN for Robust Heart Disease Prediction. Int J Drug Deliv Technol. 2026;16(4s): 974-987; DOI: 10.25258/ijddt.16.4s.115

1. Introduction

The heart powers veins and arteries, which supplies blood and oxygen to every part of the body. Any malfunction could result in fatalities or major health hazards. In the present context inside the United States, an individual dies from cardiovascular disease (CVD) every 34 seconds, and globally, a life is lost around every 1.5 seconds. The leading cause of death and chronic illness worldwide is still CVD, which include conditions like chronic heart disease, stroke, heart failure, rheumatic heart disease, and related disorders [1]. Unhealthy habits like not having balanced diet, lack of exercise, smoking, and drinking too much alcohol are the main causes of cardiovascular illnesses. Factors like as stress, pollution, genetics, age, and lack of access to healthcare increase the risk, which is already elevated by hypertension, diabetes, obesity, and cholesterol. The inflammation from infections (COVID-19 & Autoimmune disorders) are also key factors in the rising number of CVD cases around the world. Strong immune reactions damage blood vessel linings, causing clots and plaque that promote atherosclerosis. A weak heart muscle and poor pumping power might result from myocarditis [2]. Covid-19's high-clotting condition can clog coronary veins, and severe infections can cause cytokine storms that damage cardiac tissues. In autoimmune disorders, the immune system targets the heart and arteries, producing chronic inflammation and damage. Months after infection, COVID-related inflammation and arrhythmias increase heart disease risk. In 2022, approximately 19.8 million individuals succumbed to cardiovascular diseases worldwide, representing nearly 32 percent of total mortality [3]. Prior evaluations indicated that the annual global mortality due to CVD is roughly 17.9 million. According to the World Health Organization (WHO), cardiovascular illnesses cause around 23.6 million deaths, with coronary conditions accounting for over 82%. Disturbingly, almost 80% of these fatalities take place in low and middle-income nations, frequently during people's peak earning years, because of delayed or insufficient medical treatment. The CVD influences both

*Author for Correspondence: ritu.singh@sharda.ac.in genders; however, studies indicate that males are at a considerably greater risk of developing this condition [4]. As a disorder, that significantly diminishes life expectancy globally, early and precise identification has emerged as a crucial priority in healthcare. Considering its position as a primary contributor to global mortality, prompt prediction can significantly save lives and alleviate the substantial economic burden, potentially resulting in billions in healthcare savings. Rapid advancements in medical technology present opportunities to enhance diagnostic accuracy via advanced analytical methods. Now a days advanced diagnostic tools such as electrocardiograms and CT scans are crucial for identifying coronary heart disease; however, they frequently pose financial barriers and accessibility challenges for numerous individuals [5]. Even though the application of artificial intelligence and sophisticated machine learning techniques in biomedical applications continues to show strong promise for enhancing early diagnosis and facilitating more efficient clinical decision-making and robust validation frameworks for trustworthy heart disease prediction models are still not fully investigated [6].

An artificial intelligence (AI) and machine learning (ML) are changing healthcare by making it easier to anticipate diseases, design treatments, and manage patients, but depending on ML makes people worry about privacy, security, and prejudice. ML greatly improves the accuracy and speed of finding heart disease by looking at big datasets to guess when the disease will start, find people who are at high risk, and help doctors decide how to treat patients [7]. Metrics like accuracy, recall, precision, and ROC-AUC are used to evaluate how well a model works. These metrics measure how well the model classifies data, how well it can identify cases, how reliable its positive predictions are, and how well it predicts overall. Due to large medical datasets and diverse risk factors, heart disease prediction is difficult. Early risk detection requires accurate decision support. Computer-aided diagnosis (CAD) and AI-enabled deep learning and machine learning models are improving cardiovascular disease detection [8]. A variety of machine learning techniques have been employed to discern significant patterns and produce predictions within medical datasets. These techniques include Logistic Regression, Random Forests, SVMs, KNN, Gradient Boosting, Neural Networks, and more sophisticated models such as XGBoost, MARS, and Bayesian GLMs, but deep learning classifiers have been more popular in cardiovascular disease research lately since they are more accurate and can be used on a larger scale. Starting in 2022, there has been a growing focus on the necessity of continuous and real-time cardiac monitoring for early diagnosis and proactive measures. This study introduces an enhanced heart disease management system that attains superior performance through the integration of wearable sensors, Internet of Things (IoT) technologies, and sophisticated AI–ML models, facilitating continuous, real-time cardiac monitoring and early disease identification. Biomedical sensors collect important physiological data, such as heart rate, ECG, blood pressure, oxygen saturation, and physical activity ete and then send this data securely across IoT networks to cloud or edge platforms for quick analysis. Machine learning and artificial intelligence tools look at these signals to find problems, figure out how likely someone is to have heart problems, and help doctors make decisions quickly. The suggested integrated system improves the accuracy of diagnoses as compare to available AI-ML models, makes it easier to follow patients from a distance and lowers the cost of healthcare.

Predicting cardiac disease early and accurately remains challenging due to the complexity, dynamic nature, and variety of physiological data, despite the fact that numerous dependable methods have been developed for this purpose. With XGBoost, LightGBM, Random Forest, CatBoost, and TabNet serving as base learners and Logistic Regression acting as the meta-learner to enhance predictive accuracy and stability, the proposed internet of things cardiac estimation framework is a merged smart system that employs a hybrid-stacking ensemble. By creating minority-class samples and removing noisy instances, SMOTE-ENN addresses data imbalance and noise. Here are the key aspects of the research proposal:

- In order to make cardiac risk prediction more accurate and robust, this study presents a hybrid stacking ensemble framework that combines learning models based on bagging and gradient boosting.
- To efficiently handle class imbalance and remove ambiguous and noisy samples from clinical datasets, a data balancing strategy based on SMOTE-ENN is used.
- A tree-based feature selection method is utilized to decrease dimensionality, augment computational efficiency, and further improve the ability to generalize of the predictive model.
- The proposed framework includes SHAP-based explainable artificial intelligence, which gives clear and clinically useful information about model predictions to help doctors make informed decisions.
- The accuracy, precision,recall, F-1 score and ROC value of the proposed hybrid stacking model is 98.56 %, 97.5 %, 97 %, 97.5 %, and 99 % when evaluated at 5-fold validation.

The subsequent sections of this paper are organized in the following manner: Section 2 examined the literature pertaining to heart disease. Section 3, presents the proposed IOT–enabled Explainable stacked ensemble heart disease prediction model. Section 4, presents the results and discussion of proposed model. Ultimately, the paper is summarized in Section 5.

2. Related work

There has been a recent uptick in research into using machine-learning algorithms in conjunction with Internet of Things (IoT) technologies to enable remote patient monitoring and continuous, real-time prediction of cardiac disease. Current studies utilize classifiers like Random Forest, Gradient Boosting, and stacking ensembles to enhance diagnostic accuracy; however, issues concerning data imbalance, noise, scalability, and interpretability persist, prompting the need for additional optimization. This section provides a summary of the most current research on frameworks for predicting cardiac disease using artificial intelligence and machine learning. El-

Sofany *et. al.* [9] proposed a machine learning-based methodology for the early prediction of heart disease, utilizing various feature selection techniques and assessing ten ML classification algorithms on a composite dataset that integrates the Cleveland Heart Disease database with a proprietary dataset gathered from Egyptian hospitals, comprising 503 samples with 13 clinical features. The Synthetic Minority Oversampling Technique (SMOTE) was used to fix the class imbalance, and three approaches for selecting features; ANOVA, Chi-square, and Mutual Information were utilized to construct feature subsets (SF-1, SF-2, and SF-3). The XGBoost was the best model tested, with the SF-2 subset. It had an accuracy of 97.57%, a sensitivity of 96.61%, a specificity of 90.48%, a precision of 95.00%, an F1-score of 92.68%, and an AUC of 0.98. Further author uses Explainable AI through SHAP to improve interpretability and help clinicians comprehend prediction elements, and a mobile app for real-time prediction helps remote users. The findings show that XGBoost has strong predictive power and could support early diagnosis and decision-making in low-resource healthcare settings. Ferdowsi *et. al.* [10] developed a responsible AI framework for the early prediction of cardiovascular disease (CVD), emphasizing accuracy and patient privacy while promoting transparency in clinical decision-making. The researchers used a publicly available IEEE Data port dataset with 918 records and 11 clinical parameters. They used data anonymization and Differential Privacy using Laplace noise to protect sensitive information like age and gender. A Random-Forest-based wrapper technique found important predictors as Old peak, Maximum Heart Rate, Cholesterol, Chest Pain Type, and Exercise-Induced Angina after pre-processing and fixing class imbalance with Gaussian noise up-sampling. We looked at a number of machine learning algorithms. The classic Random Forest model had the greatest AUC of 0.974, while the privacy-preserving Logistic Regression model had the best secure performance (AUC 0.848). Hazra *et. al.* [11], critically evaluates the significance of machine learning and data mining approaches in the diagnosis and prediction of cardiac disease. It brings together results from a lot of previous research to see how well common algorithms like Decision Trees, Naïve Bayes, Support Vector Machines, Artificial Neural Networks, fuzzy logic systems, and ensemble and hybrid models operate. The study shows that hybrid methods are usually more accurate and reliable than single classifiers. The importance of data pre-processing, feature selection and cross-validation in improving prediction performance is stressed. The author suggests that well-designed data-driven decision support systems can significantly enhance the early diagnosis of cardiovascular disorders and facilitate more reliable clinical decision-making. Chaichuum *et. al.* [12] examines subclinical cardiac adverse effects post-COVID-19 vaccination through the application of two-dimensional speckle tracking echocardiography (2D-STE). A 121 vaccinated individuals were assessed, comprising symptomatic subjects who reported cardiac discomfort and asymptomatic controls. Despite conventional cardiac biomarkers, electrocardiographic findings, and left ventricular ejection fraction being within normal limits, myocardial strain analysis indicated notable decreases in global longitudinal strain (GLS) and global circumferential strain (GCS) in symptomatic individuals. The findings suggest the presence of subtle myocardial dysfunction that standard diagnostic methods fail to detect. The research indicates that 2D-STE serves as a sensitive and non-invasive method for the early identification of cardiac changes associated with vaccination. The authors conclude that GLS and GCS serve as significant prognostic indicators for detecting early myocardial alterations post-COVID-19 vaccination and may enhance clinical monitoring. Chang *et. al.* [13] suggested a study that delves into the topic of cardiac illness prediction. The patient health parameters like age, gender, blood pressure, and cholesterol were used to construct a framework based on machine learning. Data pre-processing, feature encoding, and analysis may be done securely with the help of Python packages like Scikit-learn, NumPy, Pandas, and Matplotlib. After comparing various algorithms, the Random Forest classifier proved to be the most effective, with a training accuracy of around 83%. The findings highlight the value of data-driven analytics and ensemble learning in enhancing diagnostic precision and bolstering trustworthy clinical decision-making systems for the prediction of cardiac illness. Jawalkar *et. al.* [14] introduces a machine learning methodology for the early prediction of heart disease, utilizing a decision tree-based random forest classifier augmented by stochastic gradient boosting. The framework includes comprehensive data pre-processing, encompassing normalization, exploratory data analysis, visualization, and the selection of pertinent clinical features. The optimized DTRF classifier accurately differentiates between the presence and absence of heart disease by effectively modeling nonlinear patterns and minimizing overfitting. Experimental validation using a real-world public dataset demonstrates robust predictive performance, attaining an overall accuracy of 96%, accompanied by high precision, recall, and F1-score. The results indicate that the integration of ensemble learning with loss optimization provides an effective, scalable, and clinically significant decision-support mechanism for precise cardiovascular risk evaluation and early diagnosis. Chinmayi [15] suggested a study to investigate how artificial intelligence and machine learning are changing heart disease prediction, diagnosis, and management in cardiovascular and pediatric care. Through automated decision assistance, data-driven models enable early risk assessment, eliminate diagnostic errors, and reduce clinical workload. The paper discusses AI in cardiac imaging, surgical planning, postoperative monitoring, and prognosis. It also analyzes the rising worldwide burden of cardiovascular diseases and addresses data limits, ethical difficulties, model interpretability, and clinical practice integration, highlighting AI's value in modern healthcare. Alwakid *et. al.* [16] proposed a model for detection of cardiovascular disease with advanced feature selection methods like Chi-square, Information Gain, Forward Selection, and Backward Elimination to get a small, useful collection of features from many different clinical data. The researchers suggested a number of machine learning classifiers, such as Decision Tree, Random Forest, Logistic Regression, and XGBoost. XGBoost has the best performance, with accuracy, precision, and F1-score rates of up to 99%. Using PCA to reduce dimensionality makes things even more efficient while keeping diagnostic reliability. The paradigm incorporates ethical AI concepts, including transparency, interpretability, and bias prevention, to guarantee fairness and reliability in clinical decision support systems. Teja and Rayalu [17] proposed an optimized heart disease prediction framework based on advanced machine learning approaches. The proposed model robustness and generalizability are restricted in existing studies due to inadequate feature evaluation, limited classifiers, or incomplete performance metrics. The results show that ensemble-based methods, especially XGBoost, Random Forest, and Bagged Trees, considerably beat conventional classifiers, with ROC-AUC values reaching 95% and accuracy reaching 93%. Ensemble models are suitable for practical heart disease prediction systems, as their stability and reliability are further confirmed by cross-validation results. Using scalable Apache Hadoop, Rao *et. al*. [18] presents AttGRU-HMSI, a state-of-the-art hybrid deep learning framework for accurate prediction of cardiac disease from huge and imbalanced healthcare datasets. This method optimizes an attention-based GRU model with the help of a bio-inspired HMSI algorithm after integrating IKC-based outlier removal, SMOTE for class balancing, and RFE for optimal feature selection. With experimental results showing a 95.58% F1-score, 95.42% accuracy, 92.51% precision, and 98.86% recall, the suggested model proves to be more effective than traditional classifiers, proving its reliability and suitability for clinical decision support. Aslam *et. al.* [19] presents a machine-learning framework for the classification of valvular heart disease utilizing clinical cardiovascular risk factors derived from a Pakistani dataset. A two-stage methodology integrating wrapper-based feature selection with various classifiers, such as SVM, RF, XGBoost, ANN, and ensemble models, is assessed. Experimental findings indicate that the SVM+PCA model attains optimal performance, exhibiting 96.97% accuracy and a 98.61% ROC-AUC, thereby validating that dimensionality reduction and ensemble learning substantially enhance VHD prediction and facilitate dependable clinical decision-making. Khan *et. al.* [20] put forward a classification framework based on ensemble stacking for predicting cardiovascular and diabetes diseases using clinical datasets. The method uses Support Vector Machine and Random Forest as meta-classifiers to improve predictive performance by combining KNN, Naïve Bayes, Linear Discriminant Analysis, and Decision Tree. Using stratified data splitting and standard metrics for experimental evaluation shows that the stacking method always works better than individual classifiers. The proposed model has an accuracy of 88.71% for predicting cardiovascular disease and 97.35% for predicting diabetes. The method demonstrates enhanced precision, recall, and F-measure, indicating that ensemble stacking of multiple classifiers improves robustness and diagnostic reliability. This presents stacking-based learning as an efficient decision-support approach for timely identification of cardiovascular diseases and diabetes. Almulihi *et. al.* [21] presents an enhanced ensemble-stacking framework for the early prediction of heart disease, integrating two hybrid deep learning architectures, CNN–LSTM and CNN–GRU, with a Support Vector Machine (SVM) functioning as a meta-learner. The method uses Recursive Feature Elimination (RFE) to find the most important clinical features, which cuts down on redundancy and makes the model more generalizable. During training, stratified data splitting and advanced hyperparameter optimization techniques are used to make sure that learning is strong. The proposed framework is tested on two well-known heart disease benchmark datasets. It consistently beats both traditional machine learning classifiers and individual hybrid deep learning models. The stacking ensemble achieves a maximum accuracy of 97.17% on the Cleveland dataset and 78.81% on a large-scale heart disease dataset, with significant improvements in precision, recall, and F1-score. These results show that ensemble stacking works well for combining different deep learning models. Singh *et. al.* [22] presents a comprehensive machine learning framework aimed at the early prediction of congestive heart failure (CHF) utilizing the Cardiovascular Health Study (CHS) dataset. The proposed method tackles data quality issues by integrating C4.5 decision tree-based feature selection with KNN-based imputation for missing data, thereby addressing the challenges posed by noisy, inconsistent, and incomplete clinical records. A Deep Neural Network (DNN) is utilized for classification, and its performance is evaluated against traditional machine learning models such as Logistic Regression, Support Vector Machine, Random Forest, KNN, Naïve Bayes, and Decision Tree. Experimental results demonstrate that the DNN surpasses the comparative models, attaining an accuracy of 95.30%, an F1-score of 97.03%, a sensitivity of 96.49%, and a precision of 97.58%. The results indicate that the combination of effective pre-processing methods and deep learning significantly enhances the accuracy of CHF predictions while minimizing reliance on extensive clinical evaluations. Ilyas *et. al.* [23] proposed a study that compares machine-learning methods for early heart disease prediction using a structured clinical dataset. Rapid Miner analysed 1025 Kaggle heart disease instances with 14 attributes like age, gender, blood pressure, and cholesterol. Supervised learning algorithms, such as Decision Tree, Random Forest, K-Nearest Neighbors, Support Vector Machine, Logistic Regression, Linear Regression, Naïve Bayes, and an ensemble approach, were evaluated for accuracy, precision, recall, and F1-score. The experimental results indicate that the Naïve Bayes classifier outperformed other models, achieving 90% accuracy and high predictive reliability. Machine learning–based decision-support systems can improve clinical decision-making, early diagnosis, and cardiovascular risk, according to the study. Advanced models, explainable AI, and IoT-enabled personalized monitoring are future priorities. In conclusion, the literature review shows that early diagnosis, continuous monitoring, and support for medical decisions have all been greatly enhanced by machine learning-based heart disease detection systems, especially when combined with IoT-enabled sensing and communication frameworks. Data heterogeneity, class imbalance, real-time scalability, and limited interpretability of predictive models are some of the remaining challenges, despite the promising accuracy and automation as reported in existing studies. To fill these gaps and create scalable, explainable AI, advanced feature engineering, ensemble learning, and clinically deployable heart disease prediction systems, strong IoT-assisted architectures are necessary.

3. Proposed IOT–Enabled Explainable Stacked Ensemble Model

In this research, an IoT-enabled intelligent heart disease prediction model is proposed to support accurate and reliable clinical decision-making by determining whether heart disease is present or absent in a patient. Figure 1 shows the full architecture of the proposed system, which combines cloud computing, advanced data pre-processing (Feature scaling & SMOTE-ENN technique), advanced ensemble machine learning techniques with the collection of physiological signals, and Explainable AI (SHAP) approach. The framework starts with wearable and diagnostic biomedical sensors that collect multiple types of physiological and clinical data from the body. A signal conditioning and data acquisition unit processes these signals to make sure the data is correct. Then, an Internet of Things (IoT) interface module sends them to the cloud platform. At the same time, the heart disease dataset stored in the cloud is systematically pre-processed to get rid of noise, deal with missing values, and normalize features. The SMOTE-ENN technique is used to fix the class imbalance that is often seen in medical datasets. It does this by improving the representation of the minority class and getting rid of noisy and misclassified samples. Then, a tree-based feature importance method is used to find the most important clinical attributes. This reduces the number of dimensions and makes the model work better and faster. Then, the cleaned-up dataset is split into two parts: a training set and a testing set, with a 70:30 percent split ratio. Five strong tree-based learning models; XGBoost, LightGBM, Random Forest, CatBoost, and TabNet classifiers are used as base learners for the classification task. To make predictions even better, a stacking ensemble method is used, with Logistic Regression as the meta-learner to best combine the predictions of the base classifiers. This ensemble strategy successfully diminishes model bias and variance while encompassing varied decision boundaries. We train and test the proposed ensemble model with both dataset-derived and IoT-generated testing data to see how well it generalizes. Finally, the trained model gives predicted outcomes that show whether heart disease is present or not and this is backed up by a thorough performance evaluation that uses standard metrics like accuracy, precision, recall, F1-score, and ROC–AUC. Additionally, SHAP (SHapley Additive exPlanations) is also used to explain the proposed model by measuring how much each clinical feature contributes to each prediction. This makes the model more transparent, builds trust in the model, and makes decisions more reliable.

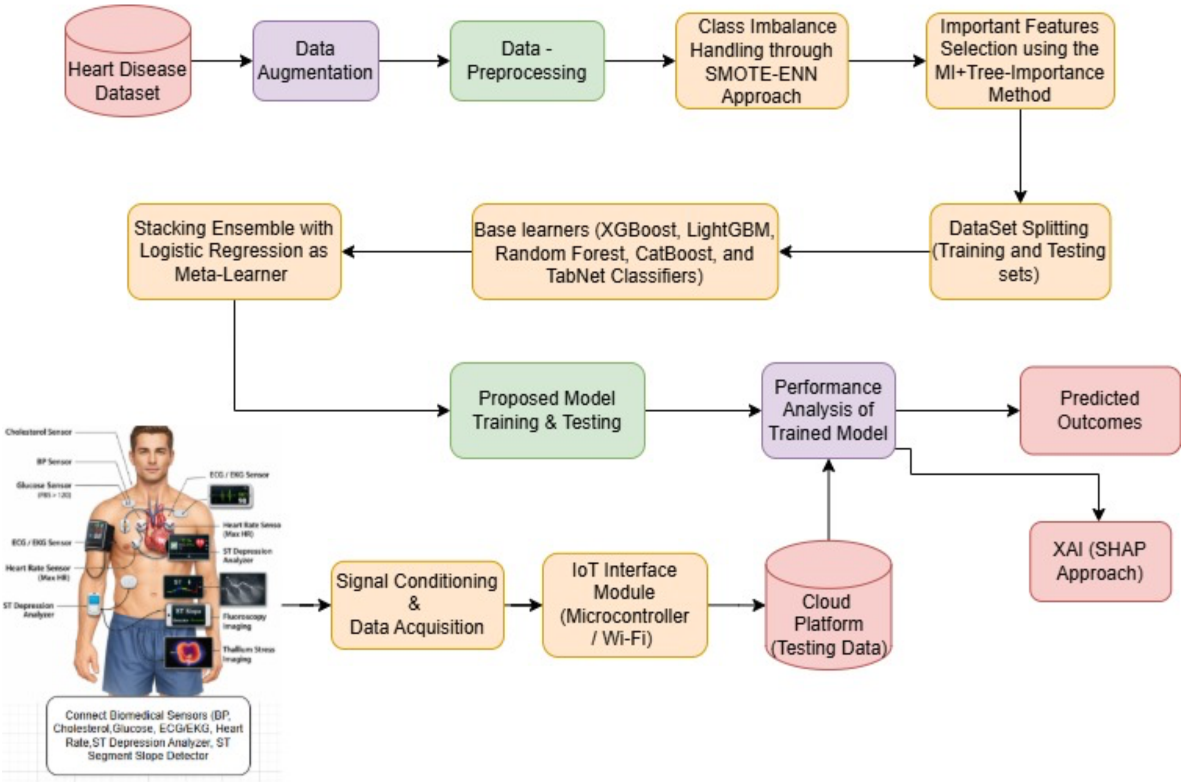


Figure 1: Proposed IOT enabled Explainable Stacked Ensemble Model

3.1 Dataset Collection

For predicting heart disease, the proposed framework uses a hybrid data acquisition method that combines learning from offline datasets with real-time physiological monitoring to make the model more robust and clinically relevant. The training data come from a Kaggle dataset that is open to the public and contains structured and clinically relevant patient records for assessing cardiovascular risk. These records include important demographic, physiological, and diagnostic information such as age, sex, type of chest pain, resting blood pressure (BP), serum cholesterol, fasting blood sugar (FBS > 120 mg/dL), electrocardiographic (EKG) findings, maximum heart rate achieved (Max HR), exercise-induced angina, ST depression, and ST-segment slope, and other related clinical parameters. Each record has a clear label that shows whether or not heart disease is present. This makes supervised machine learning and model training more reliable. At the same time, wearable and diagnostic biomedical sensors are constantly gathering real-time physiological data from the human body. A signal conditioning and data acquisition unit improves these signals to make sure they are accurate and free of noise. Then, an Internet of Things (IoT) interface module sends them to a cloud platform. This allows for continuous monitoring, centralized data management, and easy integration with the trained prediction model for real-time heart disease assessment. The histograms of all responsible attributes for the prediction of heart disease are shown below in figure 2.

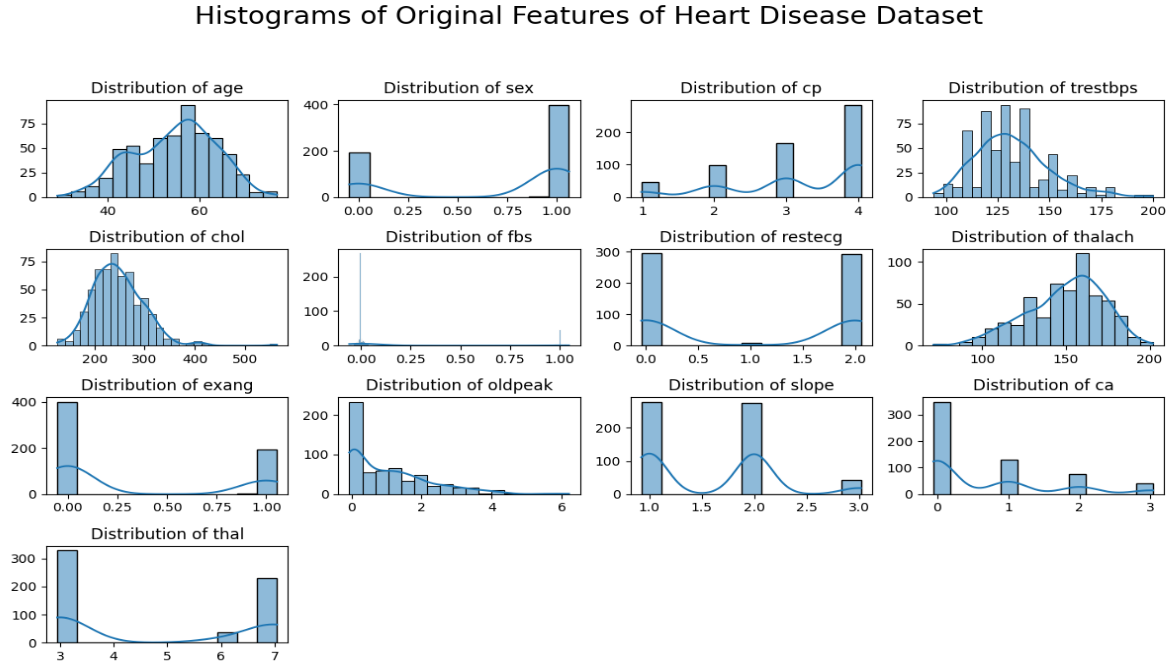


Figure 2: Histograms of original heart disease features

Figure 2 shows histograms of the original heart disease dataset features, which show how key demographic, physiological, and clinical variables are spread out. Continuous attributes exhibit diverse distribution patterns, whereas categorical features manifest as distinct value clusters. The figure shows that the data is not uniform, is skewed, and is not balanced, which shows that pre-processing is necessary before training the model.

3.2 Feature Scaling

In machine learning, feature scaling is an important step that makes sure that all of the input features contribute equally to training the model. Age, blood pressure, and cholesterol are just a few examples of variables that can be on different scales in real-world datasets. If these variables are not scaled, they can affect how learning algorithms work [24]. This problem is especially important for models that use distance or gradients, such as Neural Networks, K-Nearest Neighbors, Support Vector Machines, and Logistic Regression. Proper feature scaling makes models more accurate, speeds up convergence, makes numbers more stable, and stops features with big values from taking over the learning process. Min-Max normalization, Z-score standardization, and robust scaling are all common methods. This study uses feature standardization because Logistic Regression as the final meta-classifier, with a stacked ensemble of Extreme Gradient Boosting (XGBoost), Light Gradient Boosting Machine (LGBM), and Random Forest classifiers.

3.3 SMOTE and ENN model

To fix the class imbalance and improve the quality of the UCI cardiac disease dataset, a hybrid resampling method that combines the Synthetic Minority Over-sampling Technique (SMOTE) and Edited Nearest Neighbors (ENN) is used. The SMOTE method creates fake examples for the minority class to balance out the class distribution. The ENN method, on the other hand, uses nearest-neighbor estimation to get rid of misclassified and noisy examples from both the majority and minority classes. This two-step process not only fixes the class imbalance, but it also reduces noise and class overlap, making the training data cleaner [25].

For each minority sample (x_i) in the heart disease dataset, its k-nearest neighbors from the minority class are first identified. SMOTE then generates synthetic samples using linear interpolation, as expressed in Eq. (1), thereby increasing the density of minority instances within the feature space without simply duplicating existing data points.

$$x_{new} = x_i + \delta(\hat{x}_i - x_i) \qquad \text{(Eq. 1)}$$

Where,

δ = Random scaling factor; lie between 0 &1

x_i = Original minority-class sample

$\hat{x}_i$ = Randomly chooses k-nearest minority neighbors

In Edited Nearest Neighbors approach, sample computed with its k-nearest neighbors (k = 3). If the majority class among neighbors disagrees with x's label, remove x. Remove x when,

$$\arg max_c \sum_{j \in N_{k(x)}} 1[y_j = c] \neq y_x \qquad \text{(Eq. 2)}$$

Where,

x = Used dataset

y_x = True class label associated with sample x

$N_{k(x)}$ = The set of the k-nearest neighbors of sample x using Euclidean distance

y_j = class label of the jth neighbor within the neighborhood $N_{k(x)}$

c = Possible class label

$\sum_{j \in N_{k(x)}} 1[y_j = c]$ = The count of neighbors belonging to class c.

$\arg max_c$ = An operator that returns the class label c with the maximum count among the neighbors.

The flow chart of SMOTE and ENN approach are given in figure 3.

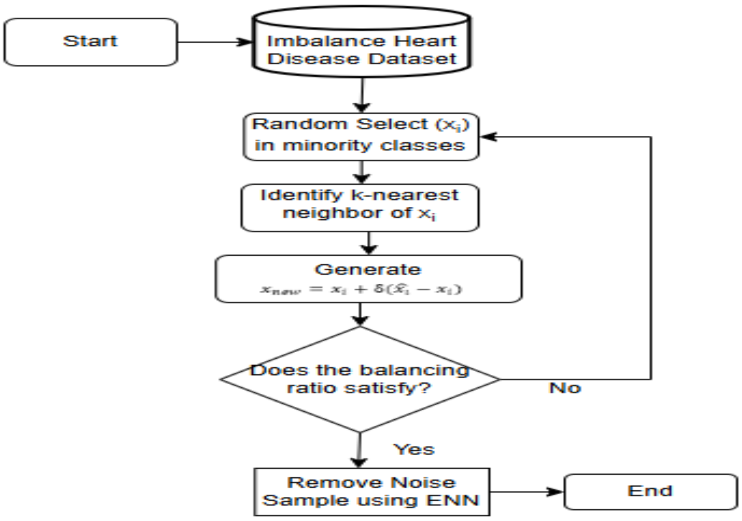


Figure 3: Flow chart of SMOTE and ENN approach for data Imbalance

Figure 4 illustrates the class and feature distribution plots after applying the SMOTE–ENN approach. The figure shows how the heart disease dataset is distributed by feature after the SMOTE-ENN technique is used, with the target class (heart disease present and absent) separated. The overlaid density plots show better balance between the two classes for each demographic, physiological, and clinical attribute. This means that class imbalance has been effectively reduced. A few characteristics, like the type of chest pain, the highest heart rate, exercise-induced angina, ST depression, the number of vessels seen by fluoroscopy, and the results of the thallium test, show a clearer separation between classes, which shows how well they can tell the difference. The figure shows that SMOTE-ENN improves class representation while keeping meaningful feature distributions, which makes the dataset better for training machine learning models.

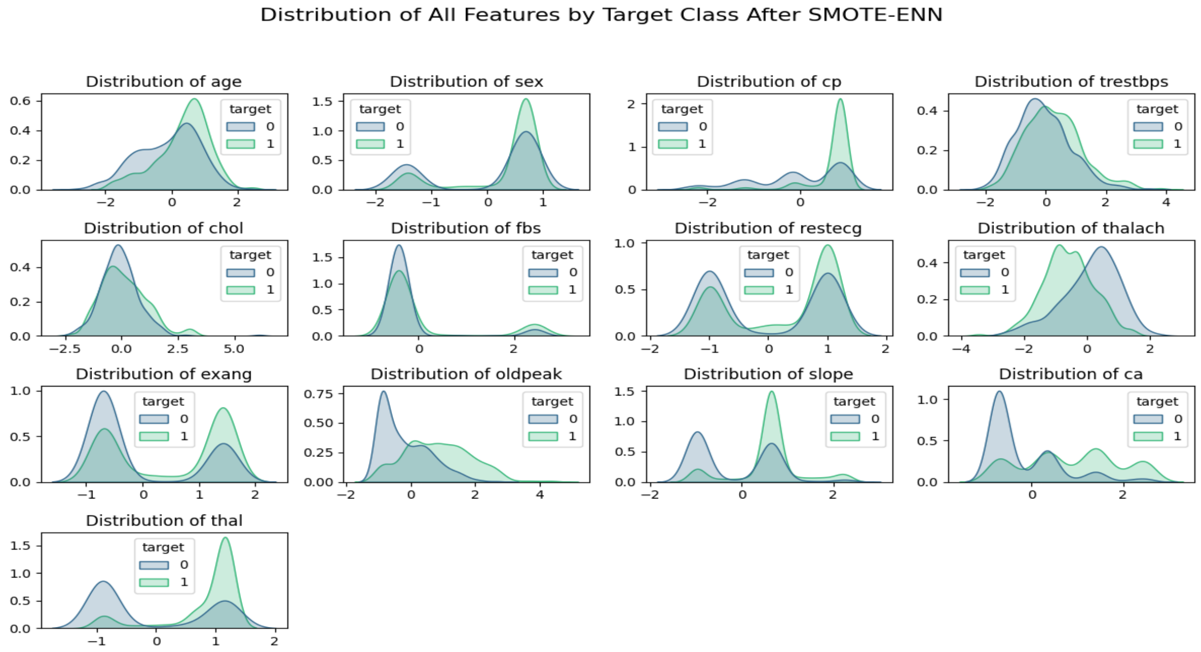


Figure 4: Class and feature distribution plots after applying the SMOTE–ENN approach

3.4 Tree Importance based Feature Selection

Feature selection is a crucial pre-processing phase in machine learning that retains the most informative variables while discarding irrelevant or redundant features. Tree-importance-based feature selection is prevalent because of its interpretability, scalability, and efficacy with high-dimensional and nonlinear data [26]. Tree-inspired algorithms automatically figure out how important each feature is during training by looking at how much each one helps make decisions in order to decrease the ensemble of trees prediction error or impurity of nodes. In classification problems, tree-based theories determine split quality using impurity measures like the Gini Index or Entropy, while regression trees focus on variance reduction for their splits.

For a given feature f_j, its importance shines through as the cumulative impurity reduction it brings to all splits and trees in the ensemble as given in Eq.3 and normalizing the importance value yields one [27].

$$\text{Importance } (f_j) = \sum_{t \in T} \sum_{n \in N_{j,t}} \Delta I_{n,t} \qquad \text{(Eq.3)}$$

Where,

T = Set of trees in ensemble

$N_{j,t}$ = Set of nodes in tree t where feature f_j is used for splitting

$\Delta I_{n,t}$= Decrement in impurity achieved at node n

Figure 5 illustrates the distribution of selected features across target classes using the three-feature-importance approach.

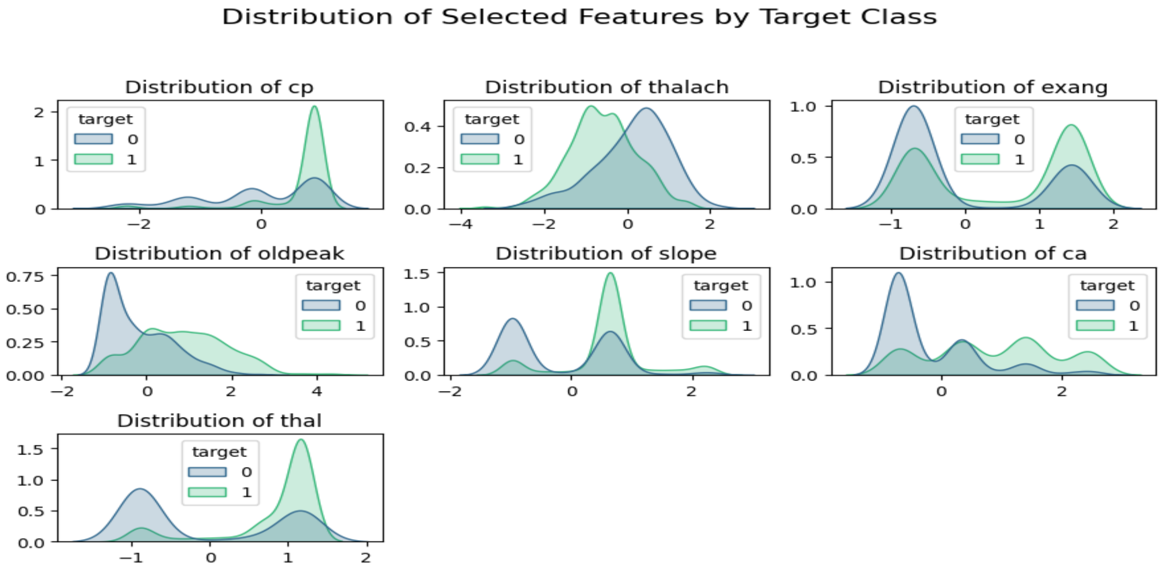


Figure 5: Distribution of selected features by target class through three-importance approach

The suggested Tree-importance feature selection method says that the most important factors for the cardiovascular disease estimation framework are the type of chest pain, the highest heart rate, exercise-induced angina, ST depression, the slope of the peak exercise ST segment, the number of major vessels (fluoroscopy), and the results of the thallium test. The chest pain, especially typical angina, indicates myocardial ischemia resulting from diminished coronary blood flow. A low maximum heart rate and angina that happens during exercise are signs that the heart is not working properly when it is under stress. A depressed ST segment or a flat or down sloping ST segment are signs of myocardial ischemia and a higher risk of heart problems. Fluoroscopy showing more affected vessels and abnormal thallium test results are signs of severe and widespread coronary artery disease.

3.5 Stacking Ensemble Approach

Stacking ensemble learning improves prediction accuracy by integrating diverse base models via a meta-learner that effectively combines their outputs, in contrast to bagging or boosting, which rely on predetermined aggregation methods. The collected datasets for heart disease prediction are inherently complex, exhibiting characteristics such as heterogeneity, nonlinearity, class imbalance, and complex interactions among clinical risk factors [28]. A robust stacking ensemble framework incorporating XGB Classifier, LGBM Classifier, Random Forest, CatBoost, and TabNet is employed to effectively address these challenges. The Random Forest model offers consistent and dependable predictions by reducing variance and noise via bagging techniques. XGBoost improves predictive accuracy through iterative correction of residual errors via gradient boosting, effectively capturing complex nonlinear relationships among features. LightGBM enhances computational efficiency and scalability by utilizing histogram-based learning and leaf-wise tree expansion, facilitating rapid convergence and robust performance on high-dimensional medical data [29]. CatBoost makes the ensemble stronger by efficiently handling structured and categorical data and reducing overfitting. TabNet also adds a deep learning angle by using attention mechanisms to focus on the most useful features at each decision point. To produce out-of-fold predictions, each base learner is cross-validated on the same dataset during training. These predictions summarize the learned patterns and are input features for the next stage. The framework prevents information leakage and improves generalization by transferring prediction-level information instead of raw features. In the stacking framework, three models function as base learners, and their out-of-fold prediction outputs serve as inputs to a Logistic Regression meta-learner. This meta-learner systematically determines the optimal contribution of each base model, facilitating a more efficient integration of their strengths compared to traditional averaging or majority voting techniques. The proposed stacking ensemble model as shown in Figure 6, enhances accuracy, resilience against noisy and correlated features, and the modeling of nonlinear patterns, providing a dependable and clinically effective framework for early heart disease risk prediction.

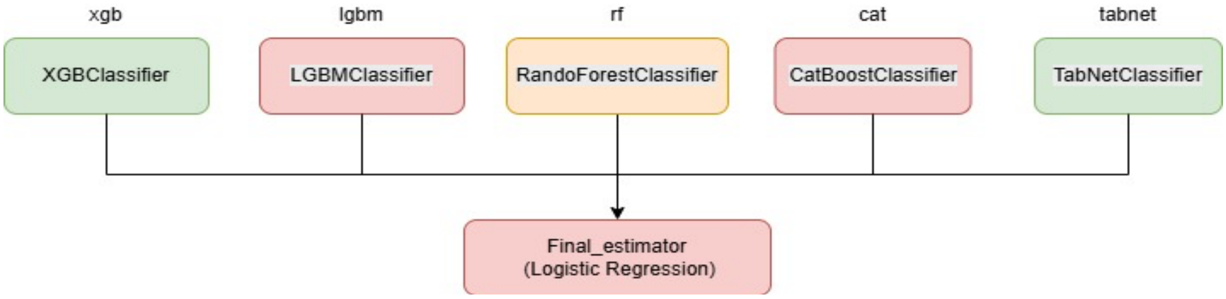


Figure 6: Suggested Stacking Ensemble model for heart disease prediction.

3.5 Explainable AI using SHAP

Explainable Artificial Intelligence (XAI) employs SHapley Additive exPlanations (SHAP) to elucidate intricate machine learning models by measuring the impact of each input feature on a model's prediction. SHAP is based on cooperative game theory, which sees each feature as a player who helps shape the result [30]. For a certain prediction f(x), SHAP explanation is shown in Eq.4 as an additive feature attribution mode.

$$f(x) = \emptyset_0 + \sum_{j=1}^{n} \emptyset_j \qquad \text{(Eq.4)}$$

Where

$\emptyset_0$ = Represents the model's baseline output (expected prediction)

n = Total number of features

$\emptyset_j$ = Shapley value corresponding to the j^{th} feature

$$= \sum_{S \subseteq N\{j\}} \frac{|S|!(n-|S|-1)!}{n!} [f(S \cup \{j\}) - f(S)]$$

S = Subset of all features excluding feature j

N = Full feature set

SHAP offers both local interpretability in heart disease prediction by elucidating individual patient-level predictions and global interpretability by ranking clinical features based on their overall significance. As a result, SHAP increases clinical trust, transparency, and the acceptability and dependability of AI-driven healthcare decision support systems [31]. In heart disease prediction, diagnostic utilization offers the exciting opportunity for high accuracy alongside clear interpretability. Explainable AI using SHAP wonderfully addresses this need by clearly illustrating how each clinical feature contributes to the model's prediction. The SHAP plot for heart disease prediction model is shown in figure 7. For a specific patient record, SHAP calculates the marginal contribution of each feature (e.g., age, cholesterol level, type of chest pain, resting blood pressure, or maximum heart rate) to the alteration of the predicted risk of heart disease. Adding up all the SHAP values and a baseline prediction gives you the exact model output, which makes it easy to understand both locally and globally. Ca, tha, oldpeak, and maximum heart rate (Max HR) are the most influential features.

Figure 7: SHAP plot for heart disease dataset

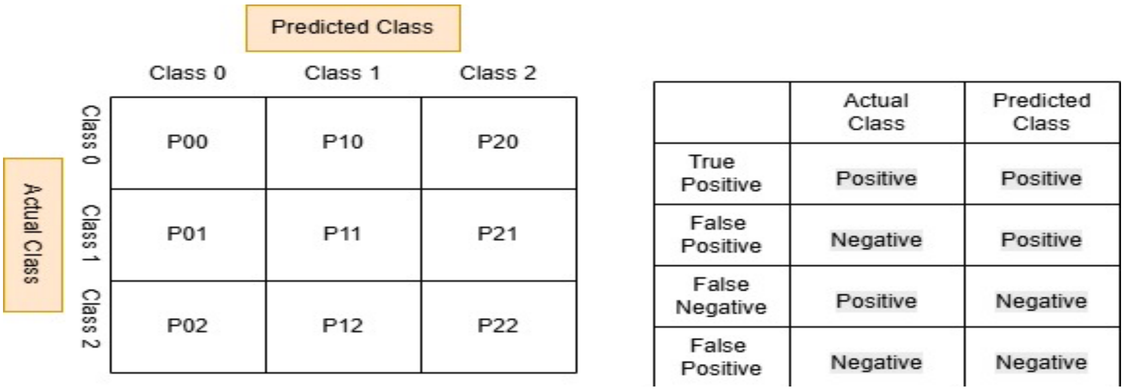
The SHAP plot in Figure 7, which is based on the heart disease dataset, shows the model's most important clinical features and how their values affect the predicted risk of heart disease right now. As indicated by the red points on the positive SHAP side, the predicted risk is significantly increased. The number of main vessels (ca) is the most important risk factor, and greater values greatly increase the risk of heart disease. Abnormal thalassemia conditions and increased ST depression (oldpeak) significantly influence predictions of illness existence, indicating underlying myocardial ischemia. On the other hand, a greater maximal heart rate (thalach) usually lowers the anticipated risk, which means that the heart is in better shape.

4. Results and Discussion

The proposed IoT-enabled heart disease prediction framework, utilizing a stacking ensemble architecture, is deployed within the Google Colab environment using Python 3.8 and five sophisticated learning algorithms. The experimental assessment is performed on a system featuring an 8th-generation Intel Core i7 processor running at 2.5 GHz and 4 GB of RAM. The model gets verified utilizing a heart disease dataset consisting of 1,504 samples, obtained through class-wise data augmentation to ensure balanced representation and a dependable evaluation of performance.

4.1 Confusion Matrix

Model performance is meticulously evaluated through accuracy, precision, sensitivity (recall), and F1-score, all of which are computed from the predicted confusion matrix. These evaluation metrics collectively offer a thorough and dependable assessment of the proposed model's predictive performance. The general confusion matrix for three classes is given below in figure 8.


Predicted Class

Actual Class	Class 0	Class 1	Class 2
Class 0	P00	P10	P20
Class 1	P01	P11	P21
Class 2	P02	P12	P22

	Actual Class	Predicted Class
True Positive	Positive	Positive
False Positive	Negative	Positive
False Negative	Positive	Negative
False Positive	Negative	Negative

Figure 8: General Confusion Matrix for Multi-Class

Equations (5)–(8) illustrate the mathematical formulations employed to calculate the accuracy, precision, sensitivity (recall), and F1-score of the machine-learning model, respectively.

$$\text{Precission} = \frac{\text{True Positive}}{(\text{True Positive} + \text{False Positive})} = \frac{P_{00}}{P_{00}+P_{01}+P_{02}} \qquad \text{(Eq.5)}$$

$$\text{Sensitivity} = \frac{\text{True Positive}}{(\text{True Positive} + \text{False Positive})} = \frac{P_{00}}{P_{00}+P_{10}+P_{20}} \qquad \text{(Eq.6)}$$

$$\text{Accuracy} = \frac{\text{True Positive} + \text{True Negative}}{(\text{Positive} + \text{Negative})}$$

$$= \frac{P_{00}+P_{11}+P_{22}}{P_{00}+P_{01}+P_{02}+P_{10}+P_{11}+P_{12}+P_{20}+P_{21}+P_{22}} \qquad \text{(Eq.7)}$$

$$\text{F}-1\ \text{Score} = \frac{2*\text{Precision}*\text{Recall}}{(\text{Precision}+\text{Recall})} \qquad \text{(Eq.8)}$$

Table 1 presents a summary of the performance metrics obtained from the previously mentioned formulations, evaluated through the confusion matrices of the proposed stacked ensemble framework alongside the respective individual machine learning models.

Table 1: Comparative evaluation of different model's performance using 5-fold cross-validation

ML Approach	**Accuracy**	**Precision**	**Recall**	**F1-Score**
SVM Classifier	89 %	92 %	86 %	88 %
Decision Tree Classifier	93 %	93 %	92 %	92 %
Random Forest Classifier	93 %	94 %	90 %	92 %
CatBoost Classifier	95 %	96 %	93 %	94 %
TabNet Classifier	88 %	90 %	85 %	87 %
Proposed Stacked Ensemble Model (without SMOTE-ENN)	**94.2 %**	**92%**	**88.5%**	**92.2%**
Proposed Stacked Ensemble Model (with SMOTE-ENN)	**98 %**	**97.5 %**	**97 %**	**97.5 %**

Figure 9, illustrates that the assessment of the proposed stacked ensemble absent SMOTE-ENN leads to a notable decline in performance, with accuracy reducing to approximately 94.2%, precision to 92%, recall to 88%, and F1-score to nearly 92.2%.

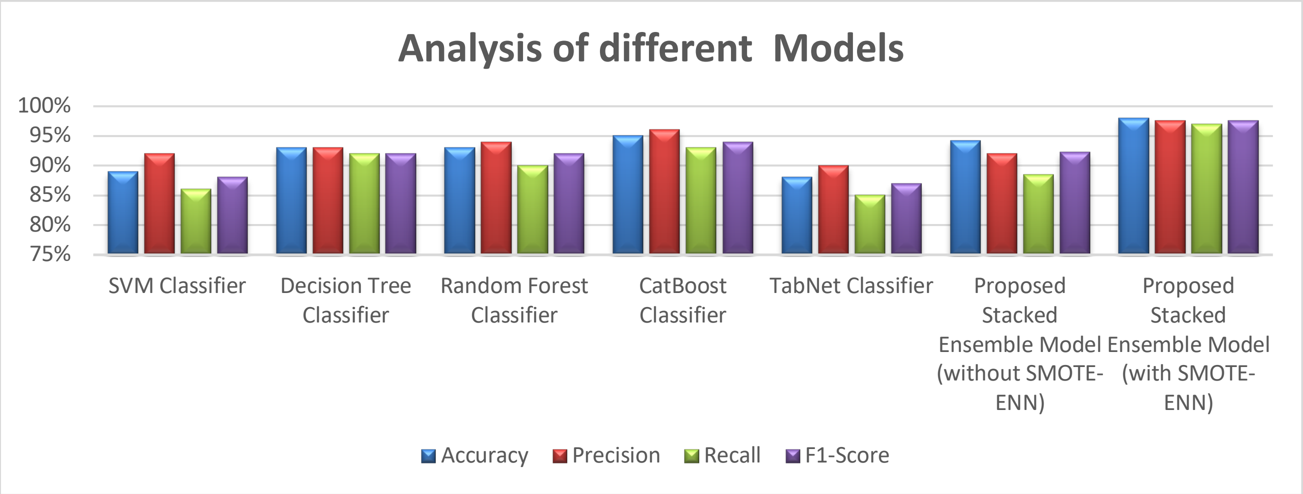


Figure 9: Comparative evaluation of different models' performance.

4.2 ROC and AUC Plots

The Receiver Operating Characteristic (ROC) curve shows the balance between the True Positive Rate (sensitivity) and the False Positive Rate at different classification thresholds to differentiate multi-class problems. The ROC curve shows how well a classifier balances sensitivity and specificity, regardless of class imbalance, by looking at this relationship. The area under the curve (AUC) takes all of this information and turns it into one number that shows how likely it is that the model gives a positive instance a higher score than a negative one. An AUC of 0.5 means random performance, while values close to 1.0 show great discriminative power. Figure 10, demonstrate the comparative discriminative efficacy of the individual machine learning models employed in the heart disease prediction framework by illustrating the trade-off between the true positive rate and false positive rate across different classification thresholds. The AUC functions as a quantitative measure of each model's capacity to effectively distinguish between patients with and without heart disease. Figures 7 depict the ROC curve of different individual machine learning models utilized in proposed IoT-enabled stacked ensemble model for heart disease prediction.

According to figure 10, support vector machine (SVM) model exhibits robust predictive performance with an AUC of 0.9340, signifying dependable sensitivity and specificity. The Decision Tree model attains an AUC of

0.9200, indicating relatively inferior discrimination owing to its constrained generalization capacity. The Random Forest model demonstrates a notable enhancement with an AUC of 0.9890, underscoring the efficacy of ensemble learning in improving robustness and minimizing misclassification.

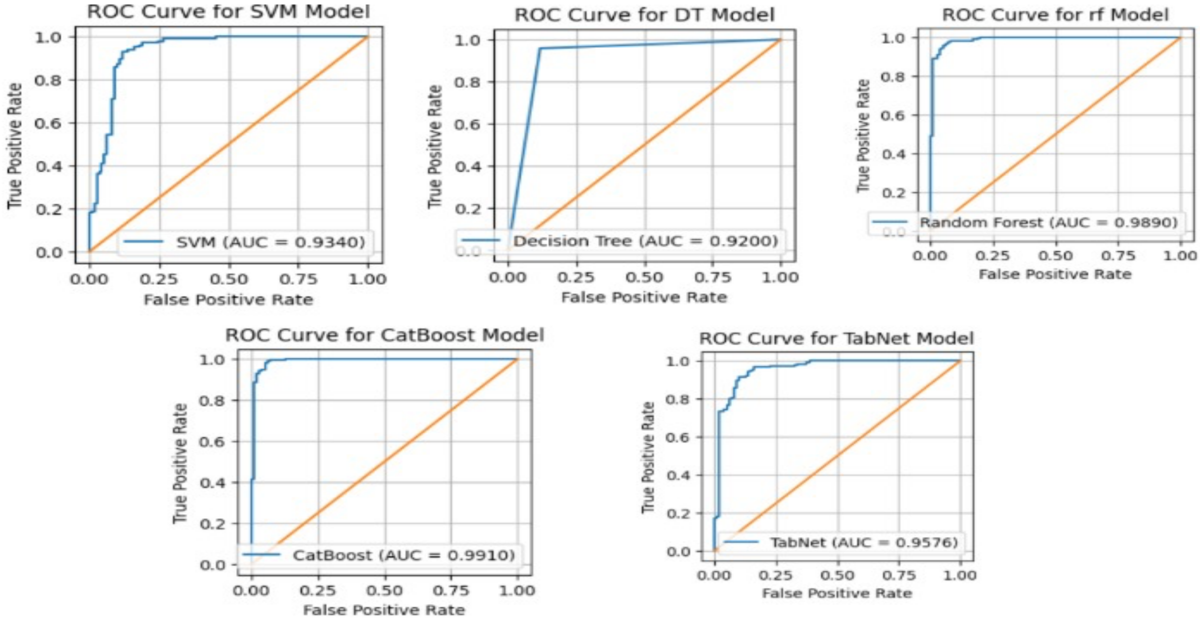


Figure 10: ROC Curve and AUC value of different ML models

The CatBoost model achieves the highest AUC of 0.9910, indicating outstanding classification efficacy and superior management of intricate feature interactions in clinical data. The TabNet model achieves an AUC of 0.9576, demonstrating robust predictive capability and providing the added benefit of inherent interpretability via attention-based feature selection.

The ROC curves collectively demonstrate that advanced ensemble and deep tabular models surpass traditional classifiers in predicting heart disease, thereby validating their incorporation as base learners in the proposed stacked ensemble framework to enhance accuracy, reliability, and generalization. Figure 11, shows the ROC & AUC curve for the proposed stacked ensemble heart disease prediction model.

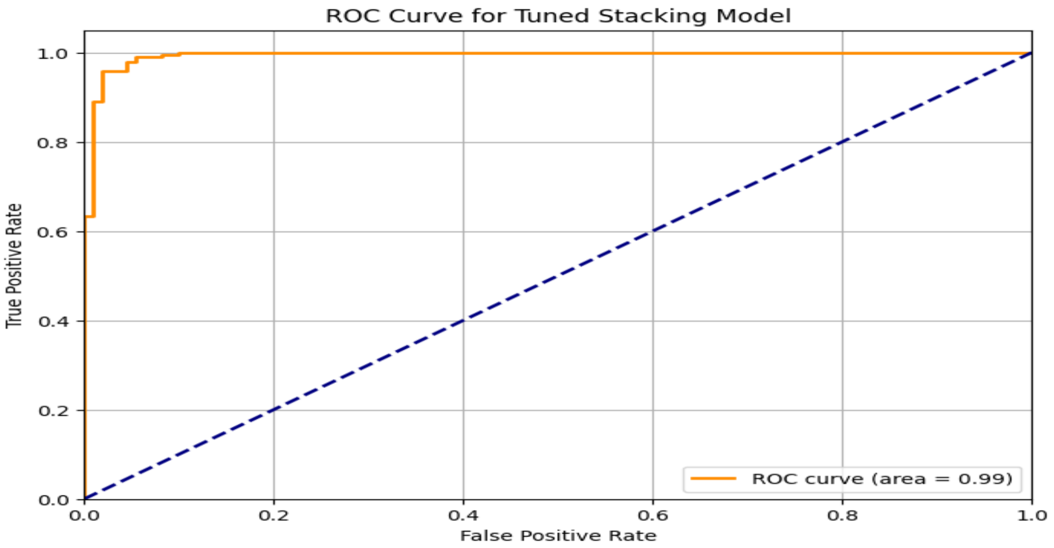


Figure 11: ROC –AUC Value of proposed stacked Ensemble model

The proposed model achieves the highest AUC of 0.99, indicating outstanding classification efficacy and superior management of intricate feature interactions in clinical data.

Table 2: Comparative study report of Proposed Stacked Ensemble Model with previous related work

Sr. No	**Authors**	**Model accuracy**	**Classification Approach**
1.	H. El-Sofany [9]	97.57 %	Binary Class

2.	V. Chang et. al. [13]	83 %	Binary Class
3.	Jawalkar et al [14]	96 %	Binary Class
4.	Teja & Rayalu [17]	93 %	Binary Class
5.	Aslam et. al [19]	96.17 %	Binary Class
6.	Khan et al. [20]	97.35 %	Binary Class
7.	Almulihi et al. [21]	90 %	Binary Class
8.	**Proposed Stacked Model**	**98.56 %**	**Binary Class**

The table 2 illustrates a distinct performance disparity between current binary classification methods and the proposed model. Previous studies indicate accuracies between 83% and 97.57%, whereas the proposed stacked ensemble attains the highest accuracy of 98.56%, surpassing all comparative methods. This enhancement illustrates the advanced proficiency of the proposed framework in identifying intricate clinical patterns and providing reliable heart disease predictions.

5. Conclusion

This work introduces a cohesive IoT-enabled hybrid framework for predicting heart disease, integrating class imbalance reduction, intelligent feature selection, ensemble learning, and explainable AI. The framework demonstrates reliable performance on both benchmark and real-time datasets through the utilization of SMOTE-ENN, heterogeneous stacking, and SHAP-driven interpretability, indicating its significant applicability to IoT-based clinical decision-support systems. Future research will concentrate on federated learning for privacy, multimodal data integration, and real-time implementation utilizing advanced machine and deep learning methodologies.

References

[1] Manisha, S., R. K. Anuradha, and K. Kodanda Reddy. “Cardiovascular Disease: A Comprehensive Review of Epidemiology, Risk Factors, and Advances in Management.” *World Journal of Biology Pharmacy and Health Sciences*, vol. 21, no. 01, 2025, pp. 495–500.

[2] A. Denegri, V. Dall’Ospedale, M. Covani, M. Pruc, L. Szarpak, and G. Niccoli, “Cardiovascular complications of COVID-19 disease: A narrative review,” *Diseases*, vol. 13, no. 8, p. 252, 2025, doi: 10.3390/diseases13080252.

[3] *World Heart Federation,* World Heart Report / World Heart Observatory: Reports and Country Profiles on Cardiovascular Disease Burden and Policy*, 2023–2025.*

[4] J. P. Lopez-Lopez, M. Perez-Mayorga, M. Rebolledo-Del Toro, J. Otero, D. Martinez-Bello, S. Rangarajan, A. A. Garcia-Peña, S. Yusuf, G. O'Donovan, and P. Lopez-Jaramillo, “Sex differences in cardiovascular disease risk factor prevalence, morbidity, and mortality in Colombia: Findings from the PURE Study,” *Global Heart*, vol. 19, no. 1, pp. 1-10, 2024.

[5] V. Shorewala, “Early detection of coronary heart disease using ensemble techniques,” *Informatics in Medicine Unlocked*, vol. 26, p. 100655, 2021.

[6] M. A. Jabbar, B. L. Deekshatulu, and P. Chandra, “Prediction of heart disease using machine learning algorithms,” *International Journal of Computer Science and Engineering*, vol. 7, no. 1, pp. 1–5, 2024.

[7] J. V. Zhao *et al.*, “Beyond one-size-fits-all: addressing sex differences in cardiovascular disease prevention and treatment,” *BMC Medicine*, vol. 23, no. 1, 2025, doi: 10.1186/s12916-025-04373-8.

[8] G. M. Rao, D. Ramesh, V. Sharma, A. Sinha, M. M. Hassan, and A. H. Gandomi, “AttGRU-HMSI: enhancing heart disease diagnosis using hybrid deep learning approach,” *Scientific Reports*, vol. 14, no. 7833, pp. 1–19, Apr. 2024.

[9] H. El-Sofany, B. Bouallegue, and Y. M. Abd El-Latif, “A proposed technique for predicting heart disease using machine learning algorithms and an explainable AI method,” *Scientific Reports*, vol. 14, no. 23277, 2024. doi: 10.1038/s41598-024-74656-2.

[10] M. Ferdowsi, M. M. Hasan, and W. Habib, “Responsible AI for cardiovascular disease detection: Towards a privacy-preserving and interpretable model,” *Computer Methods and Programs in Biomedicine*, vol. 254, p. 108289, 2024.

[11] A. Hazra, S. K. Mandal, A. Gupta, A. Mukherjee, and A. Mukherjee, “Heart disease diagnosis and prediction using machine learning and data mining techniques: A review,” *Advances in Computational Sciences and Technology*, vol. 10, no. 7, pp. 2137–2159, 2017.

[12] S. Chaichuum, C.-L. Tseng, S.-C. Chang, C.-H. Chan, C.-Y. Hsu, E. Chiang, M. Daimon, S.-J. Chiang, and H.-H. Chen, “Assessment of cardiac adverse events following COVID-19 vaccination by speckle tracking echocardiography,” *Scientific Reports*, vol. 14, Art. no. 10849, 2024, doi: 10.1038/s41598-024-61641-y.

[13] V. Chang, V. R. Bhavani, A. Q. Xu, and M. A. Hossain, “An artificial intelligence model for heart disease detection using machine learning algorithms,” *Healthcare Analytics*, vol. 2, Art. no. 100016, 2022, doi: 10.1016/j.health.2022.100016.

[14] A. P. Jawalkar, P. Swetcha, N. Manasvi, P. Sreekala, S. Aishwarya, P. K. D Bhavani, and P. Anjani, “Early prediction of heart disease with data analysis using supervised learning with stochastic gradient boosting,” *Journal of Engineering and Applied Science*, vol. 70, Art. no. 122, 2023, doi: 10.1186/s44147-023-00280-y.

[15] N. R. Chinmayi, “Heart disease prediction using artificial intelligence and machine learning,” *International Journal of Preclinical & Clinical Research*, vol. 4, no. 2, pp. 32–37, Aug. 2023, doi: 10.51131/IJPCCR/v4i2.23.11.

[16] G. Alwakid, F. Ul Haq, N. Tariq, M. Humayun, M. Shaheen, and M. Alsadun, "Optimized machine learning framework for cardiovascular disease diagnosis: A novel ethical perspective," *BMC Cardiovascular Disorders*, vol. 25, no. 1, Art. no. 123, 2025, doi: 10.1186/s12872-025-04550-w.

[17] M. Darshan Teja and G. Mokesh Rayalu, "Optimizing heart disease diagnosis with advanced machine learning models: A comparison of predictive performance," *BMC Cardiovascular Disorders*, vol. 25, no. 212, pp. 1–12, 2025, doi: 10.1186/s12872-025-04627-6.

[18] G. Madhukar Rao, Dharavath Ramesh, Vandana Sharma, Anurag Sinha, Md. Mehedi Hassan, and Amir H. Gandomi, "AttGRU-HMSI: Enhancing heart disease diagnosis using hybrid deep learning approach," *Scientific Reports*, vol. 14, art. no. 7833, pp. 1–19, 2024, doi: 10.1038/s41598-024-56931-4.

[19] M. U. Aslam, S. Xu, S. Hussain, M. Waqas, and N. L. Abiodun, "Machine learning-based classification of valvular heart disease using cardiovascular risk factors," *Scientific Reports*, vol. 14, art. no. 24396, pp. 1–18, 2024, doi: 10.1038/s41598-024-67973-z.

[20] A. Khan, A. Khan, M. M. Khan, K. Farid, M. M. Alam, and M. B. M. Su'ud, "Cardiovascular and diabetes diseases classification using ensemble stacking classifiers with SVM as a meta classifier," *Diagnostics*, vol. 12, no. 11, art. no. 2595, pp. 1–15, Oct. 2022, doi: 10.3390/diagnostics12112595.

[21] A. Almulihi, H. Saleh, A. M. Hussien, S. Mostafa, S. El-Sappagh, K. Alnowaiser, A. A. Ali, and M. R. Hassan, "Ensemble learning based on hybrid deep learning model for heart disease early prediction," *Diagnostics*, vol. 12, no. 12, art. no. 3215, pp. 1–17, Dec. 2022, doi: 10.3390/diagnostics12123215.

[22] M. S. Singh, K. Thongam, P. Choudhary, and P. K. Bhagat, "An integrated machine learning approach for congestive heart failure prediction," *Diagnostics*, vol. 14, no. 7, art. no. 736, pp. 1–21, Mar. 2024, doi: 10.3390/diagnostics14070736.

[23] A. R. Ilyas, S. Javaid, and I. L. Kharisma, "Heart Disease Prediction Using Machine Learning," *Engineering Proceedings*, vol. 107, no. 124, pp. 1–9, Oct. 2025, doi: 10.3390/engproc2025107124.

[24] C. Wongoutong, "The impact of neglecting feature scaling in k-means clustering," *PLoS ONE*, vol. 19, no. 12, Art. e0310839, Dec. 2024, doi: 10.1371/journal.pone.0310839.

[25] G. Husain, "SMOTE vs. SMOTE-ENN: A study on the performance of resampling algorithms for addressing class imbalance in regression models," *Algorithms*, vol. 18, no. 1, Art. 37, 2025.

[26] G. Louppe, L. Wehenkel, A. Sutera, and P. Geurts, "Understanding variable importances in forests of randomized trees," in *Advances in Neural Information Processing Systems*, vol. 26, Lake Tahoe, NV, USA, 2013, pp. 431–439.

[27] H. Liu, Y. Wang, and Z. Li, "A comprehensive review of feature selection methods in machine learning," *IEEE Access*, vol. 11, pp. 118245–118268, 2023.

[28] J. Zhang, Y. Liu, and R. Jin, "A stacking ensemble learning framework for improving classification performance," *IEEE Access*, vol. 8, pp. 193109–193120, 2020.

[29] M. Ganaie, M. Hu, M. Tanveer, and P. Suganthan, "Ensemble deep learning: A review," *IEEE Access*, vol. 9, pp. 61240–61256, 2021.

[30] E. Hashmi, S. Y. Yayilgan, M. M. Yamin, S. Ali and M. Abomhara, "Advancing Fake News Detection: Hybrid Deep Learning With FastText and Explainable AI," *IEEE Access*, vol. 12, pp. 44462–44480, 2024, doi: 10.1109/ACCESS.2024.3381038.

[31] K. Alam, K. Kifayat, G. A. Sampedro, V. Karovič and T. Naeem, "SXAD: Shapely eXplainable AI-Based Anomaly Detection Using Log Data," in *IEEE Access*, *vol. 12, pp. 95659-95672, 2024, doi: 10.1109/ACCESS.2024.3425472.*