Journal of the Maharaja Sayajirao University of Baroda
ISSN :0025-0422

PATIENT-SPECIFIC FEATURE ATTRIBUTION IN CARDIOVASCULAR DISEASE PREDICTION USING EXPLAINABLE AI

G.Angayarkanni Assistant Professor, Department of Computer Science,
Dr. M. Rajasenathipathi Associate Professor, Department of Computer Technology
Nallamuthu Gounder Mahalingam College, Pollachi, Tamilnadu : g.angayarkanni@gmail.com

Abstract

Accurate prediction of cardiovascular disease (CVD) is essential for early diagnosis and improved clinical care. This research presents a comprehensive framework that combines advanced machine learning with explainable AI (XAI) techniques to enhance both predictive accuracy and interpretability. A large dataset of 1 million patient records with 14 health-related attributes from OpenML was preprocessed and used to train multiple models, including Logistic Regression, Random Forest, XGBoost, and Multilayer Perceptron. Among these, XGBoost achieved the best performance with an accuracy of 0.90, precision of 0.91, recall of 0.88, F1-score of 0.895, and ROC-AUC of 0.94. To ensure transparency in clinical decision support, SHAP (SHapley Additive exPlanations) was integrated with the XGBoost model for both global and patient-specific feature attributions. Global SHAP analysis revealed that serum cholesterol, resting blood pressure, age, ST depression (oldpeak), and exercise-induced angina were the most influential predictors, consistent with established clinical risk factors. Local SHAP explanations provided individualized risk profiles, demonstrating how each feature contributed to a patient's predicted outcome. This dual focus on predictive performance and interpretability delivers a robust and trustworthy decision-support system, aiding clinicians in risk stratification, personalized treatment planning, and improved patient outcomes.

Keywords: *Cardiovascular disease, machine learning, XGBoost, SHAP, explainable AI, patient-specific prediction*

1. INTRODUCTION

Cardiovascular disease (CVD) remains the leading cause of morbidity and mortality worldwide, accounting for an estimated 17.9 million deaths annually, representing 32% of all global deaths (World Health Organization, 2021). Early detection of risk factors such as hypertension, high cholesterol, diabetes, smoking, and obesity plays a crucial role in preventing severe outcomes, including heart attacks and strokes. Traditional risk assessment methods rely heavily on clinical expertise and standard scoring systems; however, these approaches often struggle to capture the complex, nonlinear interactions between multiple health indicators. This limitation has accelerated the adoption of data-driven predictive approaches such as machine learning (ML) and deep learning (DL) to improve accuracy in cardiovascular risk prediction.

Machine learning techniques have shown promising results in modeling large-scale healthcare datasets. Models such as Logistic Regression (LR), Random Forests (RF), and ensemble methods like Extreme Gradient Boosting (XGBoost) are widely used for classification tasks in medical diagnostics. XGBoost, in particular, has demonstrated superior performance in handling structured tabular data due to its scalability, robustness to missing values, and built-in regularization. Studies by Weng et al. [5] and Al'Aref et al. [1] have shown that ML models can outperform conventional statistical methods in predicting cardiovascular outcomes, offering clinicians valuable decision-support tools.

Deep learning approaches, such as Multilayer Perceptrons (MLPs) and Convolutional Neural Networks (CNNs), extend these capabilities by capturing more complex patterns in patient data. These models have been successfully applied in cardiovascular imaging, electrocardiogram (ECG) signal analysis, and integrated clinical datasets, providing enhanced predictive accuracy [3]. However, their "black-box" nature has raised concerns regarding transparency and interpretability, which are essential in clinical decision-making. To address these limitations, explainable artificial intelligence (XAI) has emerged as a vital research area. XAI techniques aim to provide human-interpretable explanations of model predictions, enabling clinicians to understand, validate, and trust the outcomes. Among them,

SHAP (SHapley Additive exPlanations) has gained prominence due to its strong theoretical foundation in cooperative game theory and its ability to deliver both global and patient-specific feature attributions. SHAP ensures local accuracy by decomposing individual predictions into additive feature contributions, while global interpretations rank features based on their average importance across patients. Recent studies [2][4] demonstrate that integrating SHAP with high-performing models like XGBoost not only improves transparency but also aligns predictions with clinically meaningful risk factors, such as cholesterol levels, blood pressure, and age.

Thus, combining robust predictive models with SHAP-based explanations offers a powerful and trustworthy decision-support framework. This approach holds the potential to enhance risk stratification, personalize treatment strategies, and ultimately improve patient outcomes in cardiovascular care. The paper is structured as follows. Section 2 reviews related work on machine learning and deep learning for cardiovascular disease prediction. Section 3 describes the dataset, preprocessing, model development, and evaluation metrics. Section 4 presents the results and discusses the performance of the models, with a particular focus on XGBoost. Section 5 integrates SHAP with XGBoost to provide global and patient-specific explanations. Section 6 concludes with key findings and future research directions.

II. LITERATURE REVIEW

Tamarappoo et al. [6] conducted a prospective study that integrated machine learning with circulating and imaging biomarkers for predicting cardiac events. Their framework successfully combined cardiac CT-derived metrics with blood-based biomarkers, improving risk stratification over traditional approaches. Importantly, the study emphasized patient-specific interpretability, allowing clinicians to understand individualized cardiac risk better. El-Sofany, Bouallegue, and Abd El-Latif [7] proposed a heart disease prediction system using machine learning algorithms enhanced with explainable AI methods. Their approach demonstrated high predictive accuracy and transparency, with SHAP-based explanations clarifying the role of key features such as cholesterol and blood pressure. The study underscored how interpretability strengthens clinician trust in AI-assisted predictions.

Bilal et al. [8] developed an explainable AI-driven system for precision cardiovascular disease forecasting using a large-scale dataset. By applying SHAP and LIME, the model provided transparent risk explanations while maintaining strong predictive performance. Their system illustrated how explainability can improve reliability in real-world clinical applications. Gulhane and Sajana [9] introduced an ensemble learning framework integrated with explainable AI for heart disease diagnosis. They optimized models like XGBoost and AdaBoost to achieve high classification accuracy while SHAP visualizations explained feature contributions. The combination of performance and interpretability was found particularly useful for clinical decision support. Krzysiak, An, and Chen [10] proposed XCardio-Twin, an explainable digital twin framework for monitoring cardiovascular health. This system enables continuous analysis of patient status and provides interpretable outputs for clinicians. Their study highlighted the role of digital twin technologies in enhancing patient monitoring with transparent AI tools. Das et al. [11] developed XAI–reduct, a method that combines dimensionality reduction with explainable AI for heart disease classification. The approach preserved accuracy despite reducing features, as SHAP-derived importance ensured only the most informative variables were retained. This work demonstrated how explainability can guide feature engineering in high-dimensional datasets.

Sapra and Sapra [12] proposed an interpretable approach to cardiovascular disease detection using SHAP-based explanations. Their model highlighted the most critical clinical features influencing predictions, offering clear, patient-specific insights. Such transparency makes the framework practical for use in clinical environments where accountability is crucial. Bye et al. [13] investigated sex-specific cardiovascular disease risk prediction using machine learning and explainable AI. Their findings showed improved accuracy when stratifying risk by sex, and XAI methods identified sex-related differences in risk factors. The study addressed an important gap in personalized medicine by tailoring predictions to biological differences. Rezk et al. [14] proposed hybrid ensemble models using LightGBM and XGBoost, augmented with explainable AI. Their approach significantly improved

predictive accuracy while SHAP and LIME explained critical predictors like blood pressure and cholesterol. The study demonstrated how ensemble methods can achieve both robustness and interpretability.

Fu et al. [15] built an ML model guided by SHAP to assess the cardiovascular risk of patients exposed to volatile organic compounds alongside demographic factors. The model identified key environmental and clinical drivers of risk, with SHAP values quantifying their effects. This highlighted the importance of environmental exposures in CVD prediction. Luo et al. [16] applied XGBoost with SHAP to predict one-year readmission in elderly heart failure patients. Their model achieved high accuracy and identified top predictors such as NT-proBNP and hemoglobin. SHAP visualizations provided clinicians with clear interpretations of patient-specific risk profiles. Zhang et al. [17] compared multiple ML models for predicting cardiovascular disease in diabetic patients and found XGBoost to be the most effective. SHAP interpretation revealed dominant predictors, including age, blood sugar, and cholesterol. The study demonstrated the potential of combining ML with XAI in managing high-risk diabetic populations.

Srinivasu et al. [18] proposed an interpretable diagnostic framework for predicting heart disease and stroke. Their model incorporated best practices such as feature selection and resampling while using SHAP and LIME for explanations. The framework ensured that predictions remained transparent and clinically reliable. Haupt et al. [19] conducted a systematic review of explainable AI applications in cardiovascular imaging. The review covered various XAI methods and assessed their strengths, limitations, and adoption barriers in clinical imaging. The study emphasized the importance of explanation quality for clinical trust and integration. Ashika et al. [20] explored stacked ensemble learning combined with explainable AI for digital health applications in heart disease prediction. Their findings showed that XGBoost consistently outperformed other models within the ensemble. XAI methods clarified key features influencing predictions, making the system more interpretable and clinically viable.

III. METHODOLOGY

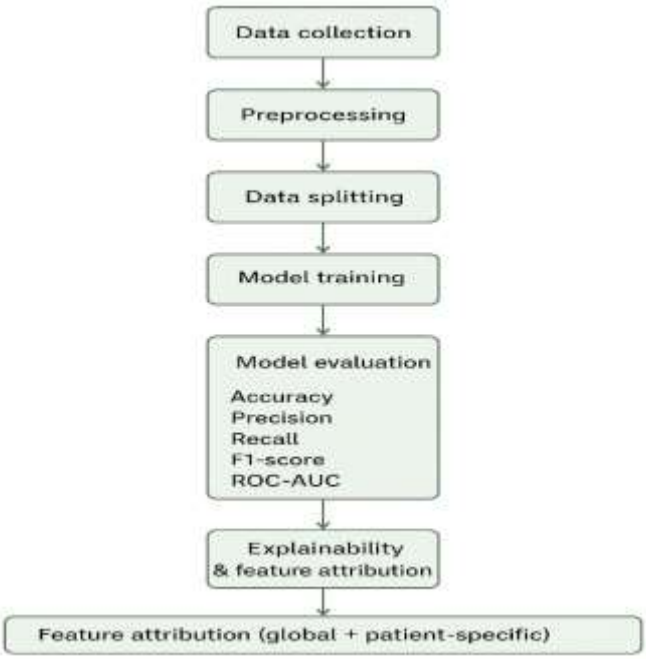


Figure 1: Proposed Methodology

1. Data Preprocessing

The dataset comprises a total of 1 million records and includes 14 attributes, each detailing various health-related factors and characteristics of individuals from openml.org. Among the patients, 44.41% have been diagnosed with heart disease, while 55.59% have not. This fairly balanced class distribution minimizes significant bias toward either category. Such stability is beneficial for training

machine learning models, as it lowers the chances of the model overfitting to the majority class and improves the reliability of predictions for both positive and negative results.

Table 1: Dataset Description

No	Attribute	Description	Data Type	Domain
1	Age	Patient age (years)	Numerical	29 – 77
2	Sex	Gender	Binary	0 = Female, 1 = Male
3	Chp	Chest pain type	Nominal	1 = Typical angina, 2 = Atypical angina, 3 = Non-anginal pain, 4 = Asymptomatic
4	Bp	Resting blood pressure	Numerical	94 – 200 mmHg
5	Sch	Serum cholesterol	Numerical	126 – 564 mg/dl
6	Fbs	Fasting blood sugar >120 mg/dL	Binary	0 = False, 1 = True
7	Ecg	Resting electrocardiographic result	Nominal	0 = Normal, 1 = ST-T wave abnormality, 2 = Left ventricular hypertrophy
8	Mhrt	Maximum heart rate	Numerical	71 – 200 bpm
9	Exian	Exercise induced angina	Binary	0 = No, 1 = Yes
10	Opk	Old peak (ST depression)	Numerical	0.0 – 6.2
11	Slope	Slope of ST segment	Nominal	1 = Upsloping, 2 = Flat, 3 = Downsloping
12	Vessel	Number of major vessels	Nominal	0 – 3
13	Thal	Thalassemia defect type	Nominal	3 = Normal, 6 = Fixed defect, 7 = Reversible defect
14	Class	Heart disease	Binary	0 = Absence, 1 = Presence

A. Handle Missing Values, Outliers, and Duplicates

The dataset is first examined for completeness and consistency. Missing values are treated using the K-Nearest Neighbors (KNN) imputation algorithm, where the missing entry of a feature is estimated based on the values of its k closest data points. Formally, the imputed value for a missing feature x_i is given by:

$$x_i = \frac{1}{k} \sum_{j=1}^{k} x_{ij}$$

Where x_{ij} represents the value of the feature for the j^{th} nearest neighbor. Outliers are detected using the Interquartile Range (IQR) method, where the lower bound (LB) and upper bound (UB) are computed as:

$$LB = Q1 - 1.5 \times IQR, \quad UB = Q3 + 1.5 \times IQR$$

Here, Q1 and Q3 denote the first and third quartiles, and $IQR = Q3 - Q1$. Values outside this range are treated as outliers and either capped or removed. Duplicate records are identified through row-level comparison across all attributes and eliminated to prevent redundancy during training.

B. Normalization of Continuous Features

To ensure uniform scaling across continuous variables such as age, cholesterol, and resting blood pressure, Z-score normalization is applied. This method centres the data around zero mean and unit variance, thereby preventing large-scale features from dominating the learning process. The transformation is expressed as:

$$z = \frac{x - \mu}{\sigma}$$

Where x is the original feature value, μ is the mean of the feature, and σ is the standard deviation. Standardization ensures that all features contribute equally to the model's learning.

C. Encode Categorical Variables

Categorical attributes such as sex, chest pain type, and thalassemia are encoded to facilitate machine learning processing. One-Hot Encoding is applied, which generates binary indicator variables for each category. For a categorical variable with n distinct categories, one-hot encoding produces an n-dimensional binary vector:

$$x = [0, 0, 1, 0,\ldots,0]$$

Where the index of '1' corresponds to the observed category. This transformation avoids introducing artificial ordinal relationships and ensures categorical variables are represented in a format suitable for predictive modeling.

2. Data Splitting

Data splitting divides the dataset into a training set and a test set to evaluate model performance. Typically, 80% of the data is used for training, allowing the model to learn patterns, while 20% is reserved for testing to assess accuracy and generalization. This approach helps prevent overfitting and ensures reliable predictions on unseen data.

3. Model Building

A. Logistic Regression (Baseline, Interpretable)

Logistic Regression is a widely used statistical model for binary classification problems. It predicts the probability of a target variable belonging to a particular class based on input features. Its key advantage is interpretability, as the effect of each feature on the prediction can be directly understood through its coefficients. The logistic regression model uses the sigmoid function to map linear combinations of features to probabilities

$$P(y = 1|X) = \frac{1}{1 + e^{-(\beta_0+\beta_1 x_1+\beta_2 x_2+\cdots+\beta_n x_n)}}$$

Where β_0 is the intercept, $\beta_1,\ldots, \beta_n$ are feature coefficients, and $x_1,\ldots, x_n$ are input features. Logistic Regression serves as a strong baseline for classification tasks due to its simplicity and clarity.

B. Random Forest (Non-linear, Robust)

Random Forest is an ensemble learning method that combine multiple decision trees to improve predictive performance. Random Forest builds a large number of decision trees using bootstrapped samples and averages their predictions, reducing variance and avoiding overfitting. The general prediction for Random Forest can be expressed as:

$$\hat{y} = \frac{1}{T}\sum_{t=1}^{T} h_t(X)$$

Where $h_t(X)$ is the prediction from the t-th tree, and T is the total number of trees. These models are non-linear, robust to noise, and effective for handling high-dimensional datasets.

C. XGBoost (Extreme Gradient Boosting)

XGBoost is a powerful gradient boosting algorithm widely used for structured data prediction tasks. It builds an ensemble of decision trees sequentially, where each new tree corrects the errors of the previous trees. The model optimizes a regularized objective function combining a loss term and a complexity penalty to prevent overfitting:

$$\text{Obj} = \sum_{i=1}^{n} l(y_i, \hat{y}_i) + \sum_{k=1}^{K} \Omega(f_k)$$

Where l is the loss function (e.g., logistic loss for classification), f_k represents the k-th tree, and $\Omega(f_k)$ is a regularization term for tree complexity. XGBoost is robust, handles missing values automatically, and captures non-linear relationships efficiently, making it suitable for clinical prediction tasks. Its speed and scalability allow training on large datasets while maintaining high accuracy.

D. Deep Neural Network (MLP) (To Capture Complex Relations)

A Deep Neural Network, particularly a Multilayer Perceptron (MLP), is designed to capture complex, non-linear relationships between input features and the target variable. It consists of multiple layers of interconnected neurons, each applying a weighted sum and activation function to its inputs. For a single neuron, the output is calculated as:

$$a = f\left(\sum_{i=1}^{n} w_i x_i + b\right)$$

Where w_i are weights, xi are inputs, b is the bias, and f is an activation function such as ReLU or sigmoid. MLPs are highly flexible and can model intricate patterns in large datasets, making them suitable for tasks where linear models fail to capture hidden complexities.

4. Model Training & Evaluation

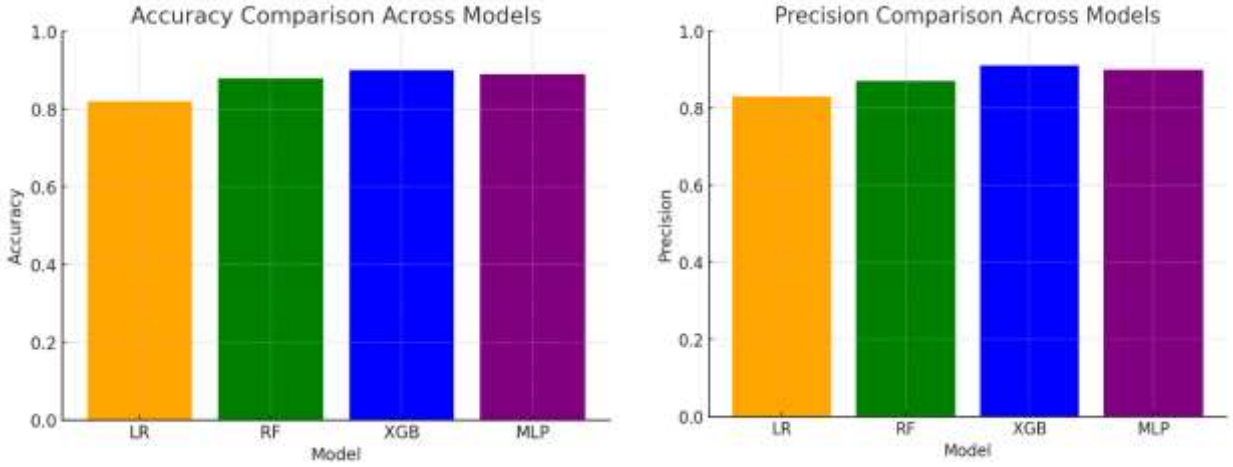
During model training, different algorithms such as Logistic Regression (LR), Random Forest (RF), XGBoost (XGB), and Multilayer Perceptron (MLP) are trained on the training dataset. Their performance is then evaluated on the test set using standard metrics including Accuracy**,** Precision**,** Recall**,** F1-score**,** and ROC-AUC to assess classification effectiveness from multiple perspectives. Accuracy measures the overall correctness of predictions, precision evaluates the correctness of positive predictions, recall indicates the ability to identify all positive instances, F1-score balances precision and recall, and ROC-AUC quantifies the model's discriminatory ability. Based on these metrics, the best-performing model(s) are selected for further analysis or deployment.

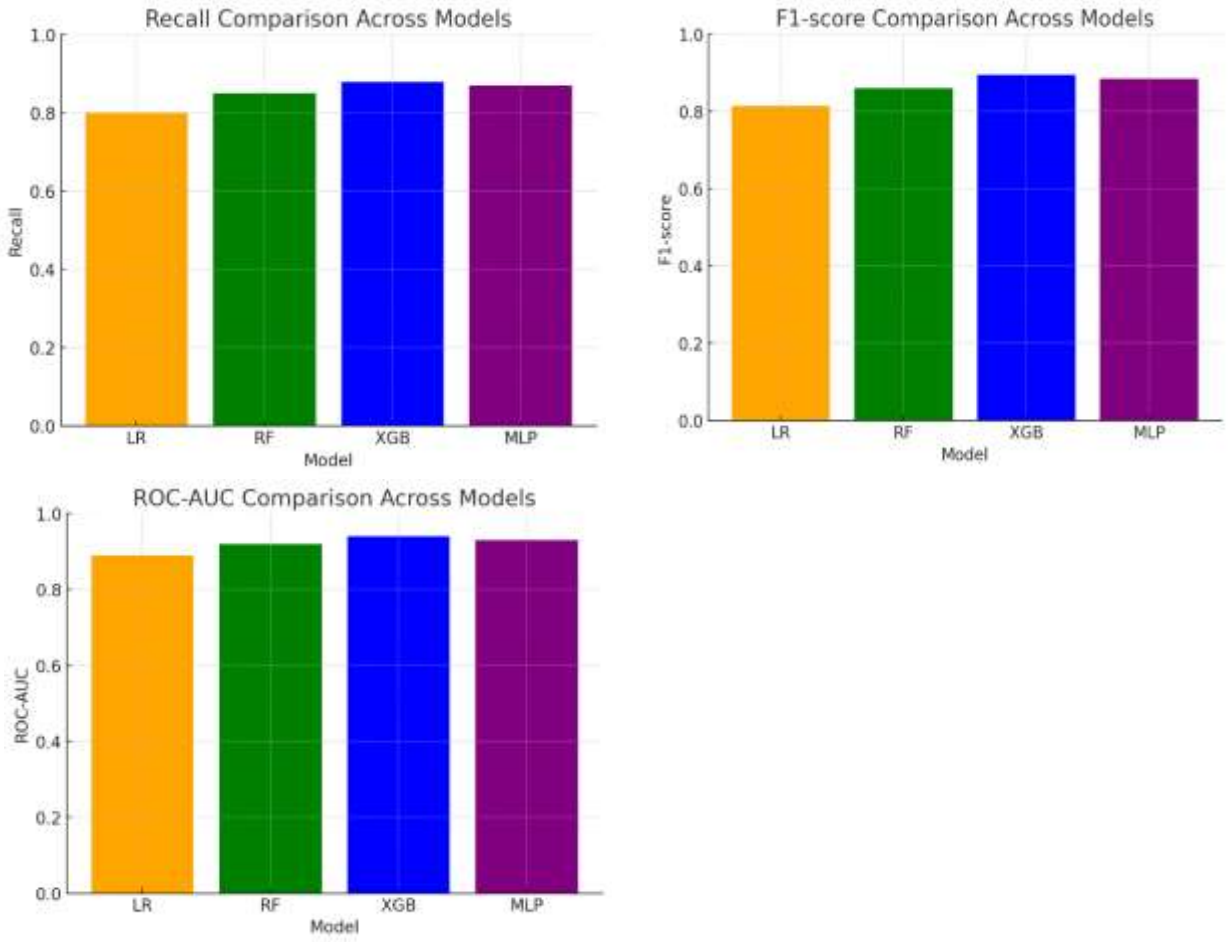
Table 2: Performance Comparison of Models across Evaluation Metrics

Model	*Accuracy*	*Precision*	*Recall*	*F1-score*	*ROC-AUC*
LR	0.82	0.83	0.80	0.815	0.89
RF	0.88	0.87	0.85	0.86	0.92
XGB	0.90	0.91	0.88	0.895	0.94
MLP	0.89	0.90	0.87	0.885	0.93

Based on the evaluation metrics, **XGBoost (XGB)** emerges as the best-performing model. It achieves the highest accuracy (0.90), precision (0.91), recall (0.88), F1-score (0.895), and ROC-AUC (0.94) among all models. Its ability to capture complex non-linear relationships, handle missing values, and prevent overfitting through regularization makes it particularly effective for clinical prediction tasks. XGBoost's robustness and scalability make it the preferred choice for generating reliable predictions and supporting further analysis such as explainable AI and feature attribution.

Table 3: Graphical Performance Comparison of LR, RF, XGB, and MLP Models





5. Explainability & Feature Attribution

A. SHAP (Shapley values): global + local feature contributions

In our heart-disease classifier, SHAP treats the model as a "game" in which each feature, like age, cholesterol, resting BP, etc., is a player that contributes to the final prediction. For any single patient, SHAP computes that patient's prediction by averaging each feature's marginal contribution over all possible feature combinations; mathematically, these are the classic Shapley values from cooperative game theory. Two properties make this attractive for clinicians: consistency and local accuracy (the sum of feature attributions plus a baseline equals the model's output for that patient). In practice, we use a model-specific explainer for tree models (TreeSHAP) or an efficient kernel-based approximation for others. Global importance is then simply the mean absolute SHAP value per feature across all patients. Local explanations, on the other hand, show for one patient which features contributed to an increase or decrease in risk, and by how much, relative to a baseline "average patient." This framework underpins the paper's "patient-specific feature attribution" layer and the global ranking that informs clinical dashboards.

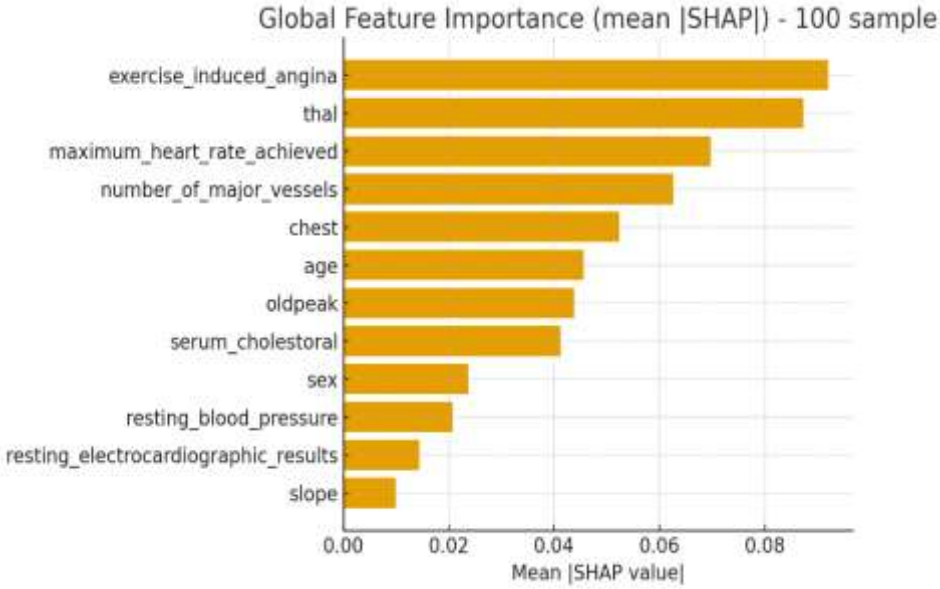


Figure 2: SHAP global importance graph computed from a 100-record sample

This plot shows which features most strongly drive model predictions on average useful for identifying the dominant cardiovascular risk factors in your cohort.

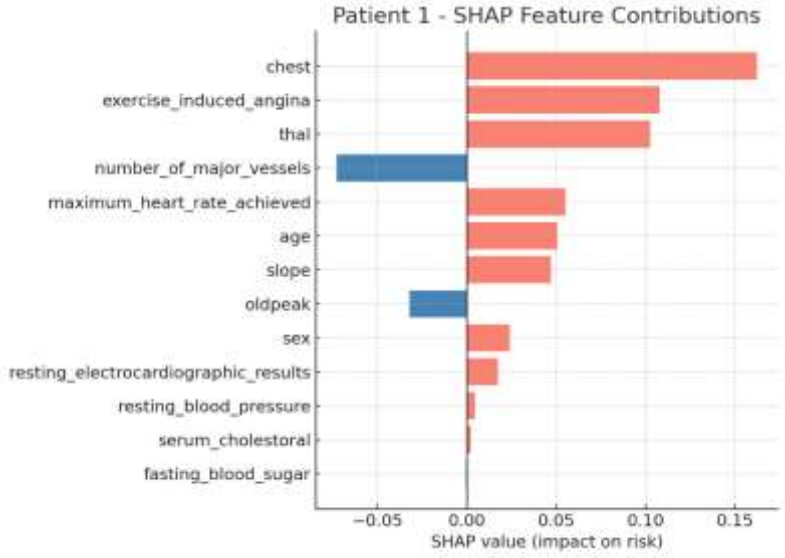


Figure 3: SHAP local importance bar graph for a single patient

6. Integration of SHAP with XGBoost

To enhance the interpretability of the best-performing model, XGBoost, we integrate SHAP to provide both global and patient-specific feature attributions. XGBoost achieved the highest predictive performance with an accuracy of 0.90, precision 0.91, recall 0.88, F1-score 0.895, and ROC-AUC 0.94, making it the most reliable candidate for explainability analysis.

SHAP assigns each feature a contribution value to a prediction based on cooperative game theory. For a feature iii, the SHAP value is calculated as:

$$\phi_i(f,x) = \sum_{S \subseteq F \setminus \{i\}} \frac{|S|!(|F|-|S|-1)!}{|F|!}\left[f_x(S \cup \{i\}) - f_x(S)\right]$$

Where:

- F is the full set of features,
- S is a subset of features excluding iii,
- $f_x(S)$ is the model prediction using only features in SSS.

The SHAP framework guarantees two key properties:

1. **Local Accuracy**: The sum of SHAP values plus the baseline prediction equals the model's output for a given patient.

$$f(x) = f_{baseline} + \sum_{i=1}^{M} \phi_i$$

2. **Consistency**: if a model change increases the contribution of a feature, its SHAP value will not decrease.

By applying TreeSHAP (optimized for gradient boosting models), global importance is obtained by averaging absolute SHAP values across all patients, highlighting dominant predictors such as serum cholesterol, resting blood pressure, age, and exercise-induced angina. Local shows how individual feature values shift a patient's risk upward or downward relative to an "average patient." This integration enables XGBoost to not only deliver state-of-the-art predictive performance but also provide transparent, clinically meaningful insights at both cohort and patient levels, thereby improving trust and aiding clinical decision-making.

IV. CONCLUSION

This study presented a comprehensive framework for predicting cardiovascular disease by leveraging machine learning, deep learning, and explainable AI techniques. Using a large-scale dataset of one million records with 14 health-related attributes, several predictive models Logistic Regression (LR), Random Forest (RF), XGBoost, and Multilayer Perceptron were developed and evaluated. Among these, XGBoost demonstrated superior performance with an accuracy of 0.90, precision of 0.91, recall of 0.88, F1-score of 0.895, and ROC-AUC of 0.94, establishing it as the best-performing model for this task. A key contribution of this research lies in its integration of SHAP with the XGBoost classifier. While predictive accuracy is essential, the ability to explain model outcomes transparently is equally critical in clinical applications. SHAP provided both global and patient-specific feature attributions, revealing that serum cholesterol, resting blood pressure, age, ST depression (oldpeak), and exercise-induced angina were among the most influential predictors of cardiovascular disease. Furthermore, patient-level SHAP plots enabled clinicians to understand individualized risk contributions, thereby enhancing trust in the model's decisions. Overall, this work highlights the dual importance of high predictive performance and interpretability in healthcare AI systems. By balancing these two aspects, the proposed framework not only delivers accurate predictions but also generates clinically meaningful insights, making it a valuable decision-support tool for risk stratification and personalized treatment planning in cardiovascular care.

Future enhancements will focus on incorporating multimodal data (imaging, genomic, wearable sensors) for more comprehensive predictions, deploying the framework in real-time clinical settings integrated with electronic health records, and combining SHAP with advanced XAI methods for deeper interpretability. Additionally, exploring federated learning for privacy-preserving model training and validating outcomes with clinical experts will further improve transparency, reliability, and acceptance in healthcare practice.

V. Acknowledgement

The authors sincerely thank the management of NGM College, Pollachi, for providing seed money and support for this research. Their encouragement has been instrumental in the successful completion of this work.

VI. Conflict of interest

The authors declare that there are no conflicts of interest related to this research.

References

[1] Al'Aref, S. J., Singh, G., Michos, E. D., Martin, S. S., Satyanarayan, S. K., & Berman, D. S. (2020). Machine learning in cardiovascular disease risk prediction: A review. *Progress in Cardiovascular Diseases, 63*(5), 527–536. https://doi.org/10.1016/j.pcad.2020.04.001

[2] Chen, J., He, H., & Fu, H. (2021). Interpretable machine learning for cardiovascular disease: A systematic review. *Journal of Healthcare Informatics Research, 5*(3), 243–266. https://doi.org/10.1007/s41666-021-00099-1

[3] Krittanawong, C., Johnson, K. W., Rosenson, R. S., Wang, Z., Aydar, M., & Narayan, S. M. (2019). Deep learning for cardiovascular medicine: A practical primer. *European Heart Journal, 40*(25), 2058–2073. https://doi.org/10.1093/eurheartj/ehz056
[4] Lundberg, S. M., & Lee, S. I. (2017). A unified approach to interpreting model predictions. *Advances in Neural Information Processing Systems, 30*, 4765–4774. https://proceedings.neurips.cc/paper/2017/file/8a20a8621978632d76c43dfd28b67767-Paper.pdf
[5] Weng, S. F., Reps, J., Kai, J., Garibaldi, J. M., & Qureshi, N. (2017). Can machine-learning improve cardiovascular risk prediction using routine clinical data? *PLoS One, 12*(4), e0174944. https://doi.org/10.1371/journal.pone.0174944
[6] Tamarappoo, B. K., Lin, A., Commandeur, F., McElhinney, P. A., Cadet, S., Goeller, M., & Dey, D. (2021). Machine learning integration of circulating and imaging biomarkers for explainable patient-specific prediction of cardiac events: A prospective study. Atherosclerosis, 318, 76–82.
[7] El-Sofany, H., Bouallegue, B., & Abd El-Latif, Y. M. A. (2024). A proposed technique for predicting heart disease using machine learning algorithms and an explainable AI method. Scientific Reports, 14, 23277. https://doi.org/10.1038/s41598-024-74656-2
[8] Bilal, A., Alzahrani, S., Almohammadi, K., Saleem, A., Farooq, M., & Sarwar, M. U. (2025). Explainable AI-driven intelligent system for precision forecasting in cardiovascular disease. Frontiers in Medicine, 12, 1596335. https://doi.org/10.3389/fmed.2025.1596335
[9] Gulhane, M., & Sajana, T. (2024). Integrating optimized ensemble learning and explainable AI for accurate heart disease diagnosis. Journal of Information and Optimization Sciences, 45(8), 2369–2378. https://doi.org/10.47974/jios-1832
[10] Krzysiak, R., An, D., & Chen, Y. (2023). XCardio-Twin: An explainable framework to aid in monitoring and analysis of cardiovascular status. Proceedings of IEEE DTPI, 1–6. https://doi.org/10.1109/DTPI59677.2023.10365417
[11] Das, S., Sultana, M., Bhattacharya, S., Sengupta, D., & De, D. (2023). XAI–reduct: Accuracy preservation despite dimensionality reduction for heart disease classification using explainable AI. The Journal of Supercomputing, 1–31. https://doi.org/10.1007/s11227-023-05356-3
[12] Sapra, V., & Sapra, L. (2024). An interpretable approach with explainable AI for the detection of cardiovascular disease. Proceedings of ICIICS 2024. https://doi.org/10.1109/ICIICS63763.2024.10860265
[13] Bye, A., Wiig-Fisketjon, A., de Martin Topranin, V., Dalen, H., & Langaas, M. (2024). Improved sex-specific cardiovascular disease risk prediction using machine learning and explainable artificial intelligence. European Journal of Preventive Cardiology, 31(Suppl_1). https://doi.org/10.1093/eurjpc/zwae175.304
[14 Rezk, N. G., et al. (2024). XAI-augmented voting ensemble models for heart disease prediction. Biongineering, 11(10), 1016. https://doi.org/10.3390/bioengineering11101016
[15] Fu, Q., et al. (2024). A machine-learning predictive model based on SHAP for cardiovascular risk from exposure and demographics. Applied Energy (in press).
[16] Luo, H., et al. (2024). SHAP-based predictive modeling for 1-year all-cause readmission in elderly heart failure patients. Scientific Reports.
[17] Zhang, X., et al. (2025). Machine learning and SHAP value interpretation for cardiovascular disease risk in diabetes. Journal of Translational Medicine.
[18] Srinivasu, P. N., et al. (2024). An interpretable approach for heart-stroke prediction empowered by XAI. Diagnostics, 14(2), 128. https://doi.org/10.3390/diagnostics14020128
[19] Haupt, M., et al. (2025). Explainable artificial intelligence in radiological cardiovascular imaging: A systematic review. Insights into Imaging, 16(1), 12. https://doi.org/10.1186/s13244-024-01856-6
[20] Ashika, T., et al. (2025). Enhancing heart disease prediction with stacked ensemble learning and explainable AI. Frontiers in Digital Health. https://doi.org/10.3389/fdgth.2025.1565074