

Forecasting Decade-Long Heart Disease Risk Through a Tailored Random Forest Framework: An Analysis Based on Framingham Dataset

Sharmeen Jahan Seema^{1*}, Prithviraj Chowdhury¹, Anik Paul¹

Abstract

Cardiovascular disease remains the leading cause of mortality worldwide, highlighting the need for accurate and early risk prediction to support effective clinical intervention. This study presents a customized Random Forest classifier for ten-year coronary heart disease prediction using the Framingham Heart Study dataset. To address the class imbalance inherent in the data, class-weighted learning is employed while deliberately avoiding synthetic data generation to maintain the original data distribution. The proposed model achieved an overall accuracy, precision, recall, and F1-score of 81%, with a ROC-AUC of 72%, indicating moderate discriminative ability. SHAP-based explainability analysis identified age and systolic blood pressure as the most influential predictors, followed by sex, daily cigarette consumption, and a history of hypertension, aligning with established cardiovascular risk factors. However, class-wise evaluation revealed limited sensitivity for the positive class, with a recall of 34% and a precision of 37%, reflecting the challenges posed by the severe class imbalance in the Framingham dataset. In contrast, the model demonstrated strong performance in identifying negative cases, achieving 88% precision and 90% recall. By integrating predictive modelling with clinical interpretability, this study demonstrates the applicability of interpretable machine learning for healthcare settings where transparency and clinical relevance are essential.

Keywords: Heart disease prediction, Random Forest, SHAP explainability, cardiovascular risk factors, machine learning

1. Introduction

Worldwide, cardiovascular disease remains the primary cause of mortality. In this category, Coronary Artery Disease, creates significant challenges for both individuals and the wider healthcare sector (Di Cesare et al. 2024). Since major

*Corresponding Author email: sharmeen_seema@bup.edu.bd

¹Department of Information and Communication Technology, Bangladesh University of Professionals, Dhaka, Bangladesh

DOI: <https://doi.org/10.64494/JFST/v4i1/ss/2026/05/93-108>

triggers like nicotine use and hypertension are manageable, the focus must shift to early detection. Accurate prediction models allow us to step in earlier, which is key to improving both clinical results and public health policies (Vu et al. 2024).

Despite growing awareness of cardiovascular risk factors, many individuals are diagnosed only after experiencing a major event such as a heart attack or stroke. The daily practice of using basic risk charts and graphs combined with the clinician's own clinical judgment and years of clinical experience is considered a common practice, but there are occasions when using such methods can overlook patients who possess several moderate-risk factors (Panattoni et al. 2025). Since universal advanced screening is not financially sustainable, we need a different approach. The priority is developing cost-effective models that can use standard clinical data to identify and prioritise the individuals who need care most (Hossain et al. 2024).

In recent years, the Framingham dataset has been used to analyze predictions for heart disease through several Machine Learning techniques (Mir et al. 2024). This dataset remains one of the most widely used benchmark datasets for cardiovascular risk prediction due to its longitudinal design and clinically relevant risk factors. A vast majority of studies have shown good accuracy in predicting heart disease using Machine Learning techniques; however, there are several practical limitations to this due to a very unbalanced training set being created. For instance, although the outcome variable for ten-year events is accurate for all subjects, it creates a heavily biased model because there will be many more cases in the majority class compared to the minority class (Mir et al. 2024). A lot of these approaches tend to focus excessively on predicting a model's overall accuracy. The inability to validate models accurately in clinical settings leads to an overstated assessment of their usefulness clinically. Therefore, for a model to be suitable for clinical use, it must not only be accurate but also provide transparency to the end-user and align with current medical knowledge (El-Sofany et al. 2024).

Random Forest Classifiers are an ideal choice for this job because of their ability to model complex, non-linear relationships and because they tend to be robust to noise (Yang et al. 2020). Yet, typical implementations of Random Forests do not allow for consideration of imbalance across classes, nor will they typically provide interpretations that can be used by clinicians to guide treatment decisions (O'Brien and Ishwaran 2019). This motivates the development of a customized Random Forest model that is designed for the Framingham dataset and that supports predictive performance while providing clinically meaningful interpretation.

This study addresses these needs by investigating heart disease prediction on the Framingham dataset using a customized Random Forest framework. The main contributions are:

- a. A customized Random Forest classifier is developed for ten-year heart disease prediction on the Framingham dataset, with class-weighted learning used to handle the strong imbalance between positive and negative cases.
- b. Clinically interpretable feature analysis is conducted to identify and explain the most influential cardiovascular risk factors in the model, providing actionable insight for early intervention and risk management.
- c. A wide range of performance metrics and validation techniques are used to assess the suggested model, and class-wise results are analysed.

2. Literature Review

The fields of medical science and machine learning have begun to converge in recent years, to the benefit of improved prediction and detection of heart disease. Multiple impactful medical studies now show that data-driven models can allow for earlier screening and beneficial clinical decisions.

2.1 Explainable AI and Advanced Models

HXAI-ML possesses a low-complexity, streamlined pipeline for early detection of CVD, demonstrating clinically realistic drivers through random oversampling and Tomek Links, tree-classifiers (Extra Trees), SHAP, LIME, and permutation feature importance with accuracy/precision/F1 of 98.78% on Cleveland and 99.24% on Framingham. However, there were several drawbacks, such as the exclusion of multimodal feature-fusion, Deep Learning (DL) techniques weren't explored, and transformer-based approaches were not considered, which are effective in handling sequential data; furthermore, the use of a singular 80/20 split, devoid of k-fold or external validation, undermines optimistic generalization (Talukder et al. 2025).

2.2 Transformer-Based Approaches

Speaking of Transformer-based approaches, on the UCI heart-disease dataset, the study by Talaat and Aly (2025) proposes a "Cardio Transformer" with the combination of dense layers with multi-head attention. The model was tested on a previously held-out test set, and reported an accuracy of 99% and AUC=1.00 with a confusion matrix of 159 TN, 146 TP, 0 FP, and 3 FN. The model outperformed baseline traditional ML models in terms of accuracy, precision/recall, and F1, where the mechanism is said to capture complex feature

interactions while maintaining high precision/recall. However, there is limited generalizability beyond the dataset used, and there is a possibility of performance fluctuations with different data quality or feature availability, which indicates the necessity of further testing in real-world conditions.

2.3 Neural Networks and Traditional Classifiers

Another comparison of SVM, Random Forest, and a small neural network on the UCI heart-disease dataset reports that the SVM achieved essentially 85% accuracy and slightly better ROC outcomes, including AUC, accuracy, F1, precision, and recall. However, the Results section introduced a dataset of "12 thousand" records, while the confusion matrices are instead based on 9,200 test cases with varying denominators. Moreover, the abstract and tables report mean classification accuracy values of 0.85 for SVM and 0.84 for both RF and NN. Consequently, the validation approach is unclear, and the interpretability as well as the generalizability of the findings are compromised (Setiyadi et al. 2025). An Enhanced RNN with stacked GRUs and SMOTE attained 81.43% accuracy on Framingham without oversampling, 98.78% with SMOTE, and 100% raw/99.18% with SMOTE on Cleveland. Yet, training plateaued at approximately epoch 860 before signs of overfitting emerged, suggesting the learning optimum had been crossed (Krishnan et al. 2023). It is somewhat difficult to assess the generalizability, given that the study was limited to only 705 Framingham cases, relied on a somewhat unclear 70/30 split, and continued training even after early stopping was declared. The emphasis on accuracy rather than comprehensive diagnostics further restricts clinical applicability. When compared to HXAI-ML, it provided less interpretability with similar headline scores (Krishnan et al. 2023; Talukder et al. 2025).

2.4 Multi-Dataset and Resampling Strategies

A mean-imputation-based multi-dataset strategy combined with bootstrap resampling, best feature selection, and an AdaBoost-with-Random-Forest classifier using 10-fold cross-validation provides a broader context. Reported scores were 99.48% (Heart Statlog), 93.90% (Heart UCI), 96.25% (Stroke), 86.52% (Framingham), and 78.36% (CHD), with the optimised individual baseline being Random Forest (Mir et al. 2024). While this design provides breadth of datasets relative to earlier efforts, the authors themselves highlight the limitations of the number of datasets and lack of external/clinical validation; therefore, real-world generalizability is not validated (Mir et al. 2024). On the other hand, Prabu et al. (2022) conducted a grid/random-search study, which emphasizes minimizing false negatives by pairing mean imputation with kernel

methods (GPR/KRR) and tuning, achieving mean recall/F1/accuracy of 90.2%/82.1%/86.0 across $n=100\text{--}300$ samples and, for $n=300$, recall/F1/accuracy of 92.89%/84.23%/89.19%, yet the mixing of regression learners/losses with classification metrics muddies the validity of the reported scores and experiments are limited to a single tabular dataset.

Similarly, optimization-enhanced machine learning models and conventional supervised classifiers have demonstrated promising performance for cardiovascular disease and stroke prediction, although both studies relied on single public datasets without external clinical validation (Nova et al., 2025; Shipra, 2023).

2.5 Classical Baselines

Despite being restricted to the standard Cleveland and Framingham datasets, the study by Mienye and Jere (2024) demonstrated a robust ensemble approach using Bayesian optimization. The model performed exceptionally well, recording accuracies of 0.984 and 0.959, respectively. Beyond the raw metrics, the inclusion of SHAP analysis was vital: by validating that features like age, sex, and ca were driving the predictions, the authors ensured their high accuracy was grounded in medical reality.

Table 1: Summary of Works that Use Tabular Datasets

Study	Dataset(s) (Tabular)	Method / Model	Main Metrics	Results
Talukder et al. 2025	Cleveland, Framingham	HXAI ML (Extra Trees + XAI)	Acc. 99.24% (Framingham), 98.78% (Cleveland)	No
Talaat and Aly 2025	UCI Heart Disease	Cardio Transformer	Acc. 99%, AUC 1.00	No
Setiyadi et al. 2025	UCI Heart Disease	SVM, RF, NN	Acc. 84–85%	No
Krishnan et al. 2023	Framingham, Cleveland	Enhanced RNN (GRU + SMOTE)	Acc. 81–99%	No
Mir et al. 2024	Framingham + other tabular CVD sets	AdaBoost + RF, resampling	Acc. 86.5% (Framingham)	No
Prabu et al. 2022	Single tabular CVD dataset	GPR/KRR + grid search	Acc. $\approx 86\%$, F1 $\approx 82\%$	No

Study	Dataset(s) (Tabular)	Method / Model	Main Metrics	Results
Patil et al. 2022	Kaggle Heart (tabular)	LR, KNN, RF, DT	Acc. 80–90%	No
Mahmoud et al. 2021	Framingham	LR, DT, KNN, SVM, RF	Acc. 85.05%, F1 91.9%	No

Table 1 shows that most prior heart disease prediction studies using tabular datasets such as Framingham, Cleveland, UCI, and related clinical records have primarily focused on maximizing overall predictive accuracy. Both conventional machine learning approaches (e.g., Random Forest, Support Vector Machine, and Logistic Regression) and more recent ensemble or deep learning architectures (e.g., HXAI-ML, Enhanced RNN, and Cardio Transformer) have reported strong performance, often exceeding 90% accuracy on benchmark datasets. However, many of these studies provide limited analysis of class imbalance and minority-class detection, despite their importance in clinical decision-making. In addition, model evaluation is frequently based on aggregate performance metrics, with relatively little emphasis on class-wise assessment, interpretability, and transparent validation of model behaviour. Consequently, a research gap remains in developing clinically meaningful prediction frameworks that not only achieve competitive performance but also provide interpretable and balanced evaluation of cardiovascular risk prediction models. Addressing this gap, the present study employs a class-weighted Random Forest model combined with SHAP-based explainability and comprehensive class-wise performance analysis. Rather than focusing solely on overall accuracy, the study highlights both the strengths and limitations of the proposed approach, providing a more transparent and clinically relevant assessment of ten-year coronary heart disease risk prediction using the Framingham dataset.

3. Methodology

The Framingham Heart Study dataset was preprocessed by handling missing values and selecting clinically relevant features for analysis. The data is initially divided into stratified training and testing sets using an 80/20 split to preserve the original class distribution. A class-weighted Random Forest classifier is then trained using optimized hyperparameters to address the significant imbalance between positive and negative cases. To evaluate model robustness beyond a single train-test split, 5-fold stratified cross-validation has also been performed. The proposed model is further compared with Logistic Regression, Decision Tree, and AdaBoost classifiers using multiple performance metrics. Finally, SHAP-based explainability analysis is conducted to identify the most influential

predictors of ten-year coronary heart disease risk and to enhance the interpretability of the model.

3.1 Dataset Description

The study uses the publicly available Framingham Heart Study dataset, which contains patient-level clinical and demographic features such as male, age, education, currentSmoker, cigsPerDay, BPMeds, prevalentStroke, prevalentHyp, diabetes, totChol, sysBP, diaBP, BMI, heartrate, and glucose. These features are relevant to cardiovascular risk assessment. The target variable is TenYearCHD, which indicates whether an individual developed coronary heart disease within ten years.

3.2 Data Preprocessing

Further analysis of the dataset has shown that some of the columns have missing values, such as education, cigsPerDay, BPMeds, totChol, BMI, heartrate, and glucose columns; these missing values are handled with necessary imputation. To ensure model compatibility and numerical stability, all feature variables are converted to floating-point representations. A correlation heatmap in Figure 1 is generated to examine multicollinearity among features, and highly correlated values are assessed before model training. Among all the features, the education feature is not selected as it is irrelevant to this study. The dataset is then split into training (80%) and testing (20%) subsets using stratified random sampling to maintain class distribution.

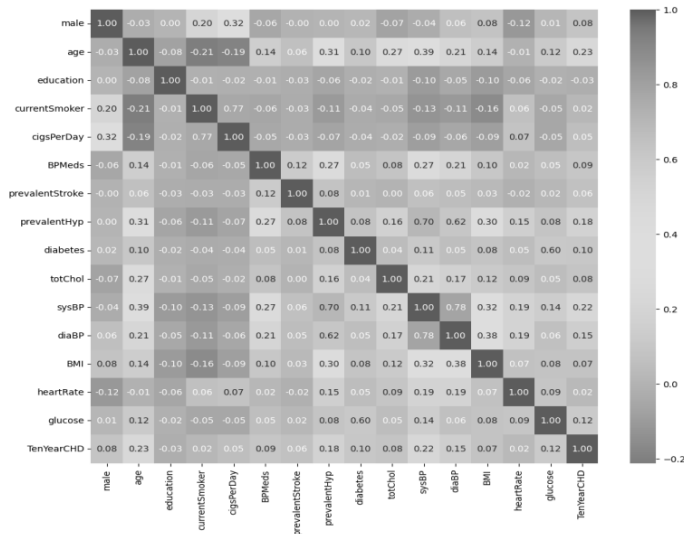


Figure 1: Correlation matrix of clinical and demographic features in the Framingham Heart Study dataset

3.3 Model Development and Training

A Random Forest Classifier is employed due to its robustness against overfitting and ability to handle class imbalance. Hyperparameters are optimised using RandomSearchCV with 5-fold cross-validation followed by further manual tuning. Table 2 shows the best configuration.

Table 2: Tuned Hyperparameters of the Custom Random Forest Model

Hyperparameters	value
Number of estimators	600
Maximum depth	15
Maximum features	sqrt
Criterion	entropy
Minimum samples split	18
Minimum samples per leaf	8
Class weight	balanced

Here, the highly imbalanced class (0:3505 and 1:628) issue is fixed by using class-weighted balanced. SMOTE-ing is avoided as it is a medical dataset, and this technique creates synthetic data. The Custom Random Forest model is trained on the training portion and subsequently validated on the testing part. Predictions are generated for both class labels and class probabilities for result analysis.

3.4 Explainable AI

To enhance model transparency, SHAP (Shapley Additive exPlanations) is used to interpret feature contributions. SHAP values quantify each feature's impact on the model's prediction by fairly distributing its contributions across all input variables. Finally, a summary plot is generated to identify the most influential predictors of the outcome, thereby improving both interpretability and trust in the model.

4. Results

This section provides a comprehensive explanation of the model suggested for heart disease prediction. All experiments have been conducted on a Lenovo IdeaPad Slim 3i laptop with an Intel Core i5-1035G1 CPU, 8GB RAM, and NVIDIA MX330 GPU, running Python 3.11 and Pytorch 2.1. For performance evaluation precision, recall, F1-score, accuracy, and ROC-AUC scores are examined. Comparative experiments with Logistic Regression, Decision Tree, and AdaBoost are also presented to assess the effectiveness of the proposed

approach and shows overall high accuracy does not necessarily mean better performance.

4.1 Overall Classification Performance

Figure 2 shows that the classification model is consistent across core evaluation metrics, achieving 81% precision, recall, F1-score, and overall accuracy. This indicates a well-balanced model with comparable ability to correctly identify positive cases and avoid false predictions. The ROC–AUC score of 72% depicted in Figure 3 suggests a moderate discriminative capability between classes.

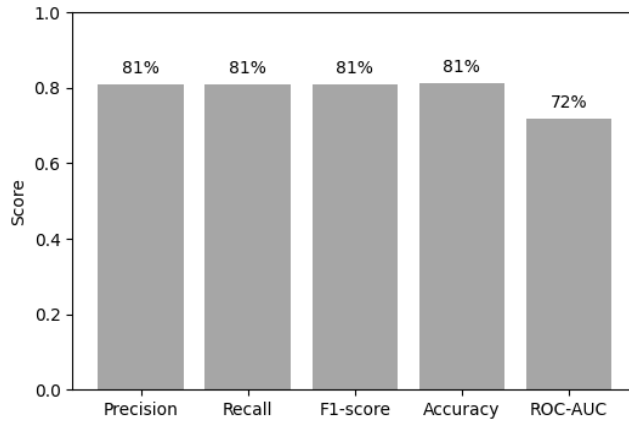


Figure 2: The overall Precision, Recall, F1-score, Accuracy, and AUC of the Proposed model

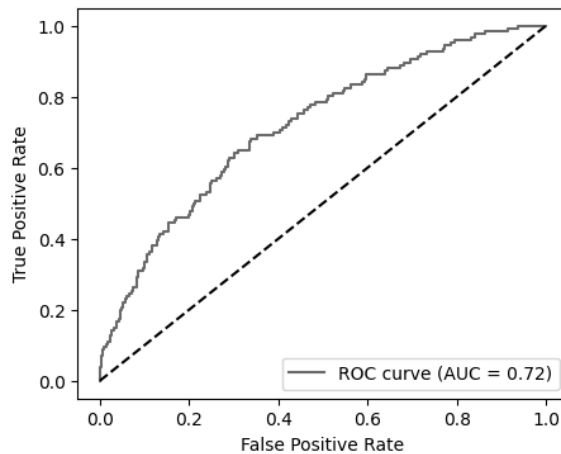


Figure 3: ROC AUC curve of the proposed model

4.2 Class-wise Classification Performance

The class-wise evaluation shown in Figure 4 illustrates a clear disparity in model performance between the two classes. The majority class achieves high precision (88%), recall (90%), and F1-score (89%), indicating strong and consistent predictive capability. In contrast, the minority class shows substantially lower precision (37%), recall (34%), and F1-score (36%), suggesting difficulties in correctly identifying this class. This imbalance highlights that, despite its strong overall performance, the model is biased toward the dominant class, as the sample difference is very high.

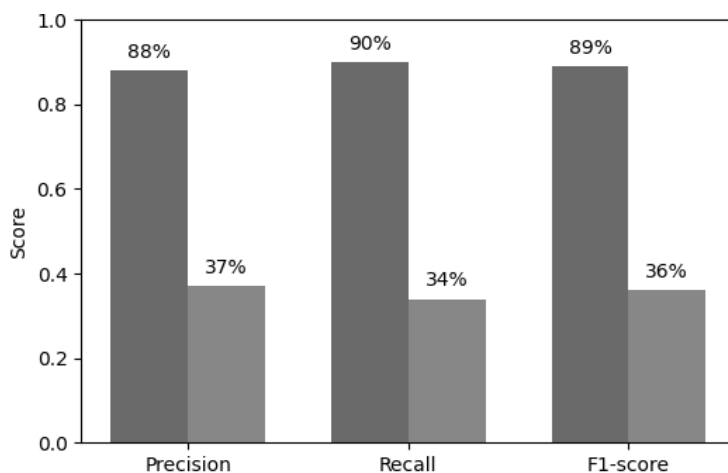


Figure 4: Class-wise Precision, Recall and F1-score of the Proposed model

4.3 Cross-Validation Analysis

Cross-validation results are shown in Table 3. A five-fold stratified cross-validation is done to evaluate the proposed model's generalization capabilities. The mean accuracy of the custom Random Forest model is 78.64% with a low standard deviation of 0.87%, indicating consistent performance across different data partitions. The weighted precision and the weighted recall scores are approximately $79.75\% \pm 0.86\%$ and $78.64\% \pm 0.87\%$, respectively. The model also shows a steady weighted F1-score of $79.12\% \pm 0.54\%$ and a ROC-AUC of $70.19\% \pm 1.22\%$. Although the cross-validation accuracy is slightly lower than the 81.14% obtained from the original 80/20 train-test split, the low standard deviations across all metrics suggest that the model exhibits relatively stable performance. This small difference is expected because in cross-validation, multiple data partitions are used.

Table 3: Five-Fold Cross-Validation Performance of the Custom Random Forest Model

Metric	Mean $\pm$ Standard Deviation
Accuracy	0.7864 $\pm$ 0.0087
Weighted Precision	0.7975 $\pm$ 0.0086
Weighted Recall	0.7864 $\pm$ 0.0087
F1-Weighted	0.7912 $\pm$ 0.0054
ROC-AUC	0.7019 $\pm$ 0.0122

4.4 Confusion Matrix Analysis

The confusion matrix in Figure 5 shows that the model correctly classified 628 negative and 43 positive instances, while misclassifying 73 negatives as positives and 83 positives as negatives. This indicates strong performance in identifying the negative class, but a relatively higher rate of false negatives for the positive class, suggesting reduced sensitivity toward positive cases.

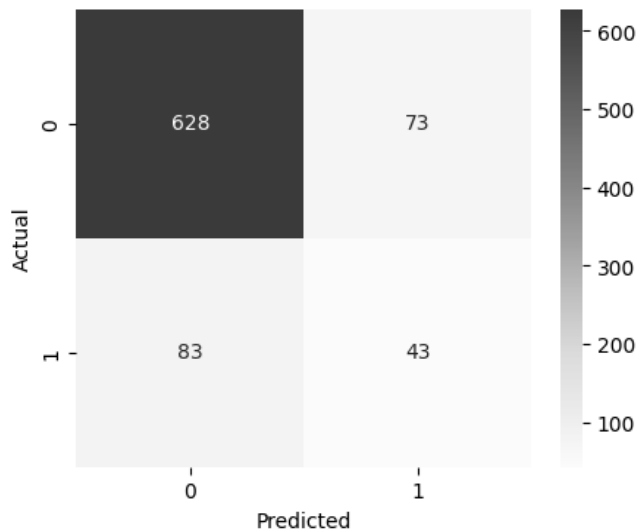


Figure 5: Confusion matrix of the proposed model

4.5 Explainability Analysis Using SHAP

The SHAP-based feature importance analysis in Figure 6 indicates that age is the most influential predictor, followed by systolic blood pressure (sysBP). Sex, smoking intensity (cigsPerDay), and prevalent hypertension also contribute substantially to the model’s predictions, while diastolic blood pressure, total

cholesterol, glucose, heart rate, and BMI show comparatively lower impact. Overall, the results highlight the dominant role of demographic and blood pressure-related factors in driving the model’s decision-making process, which shows similarities with real-world clinical cases.

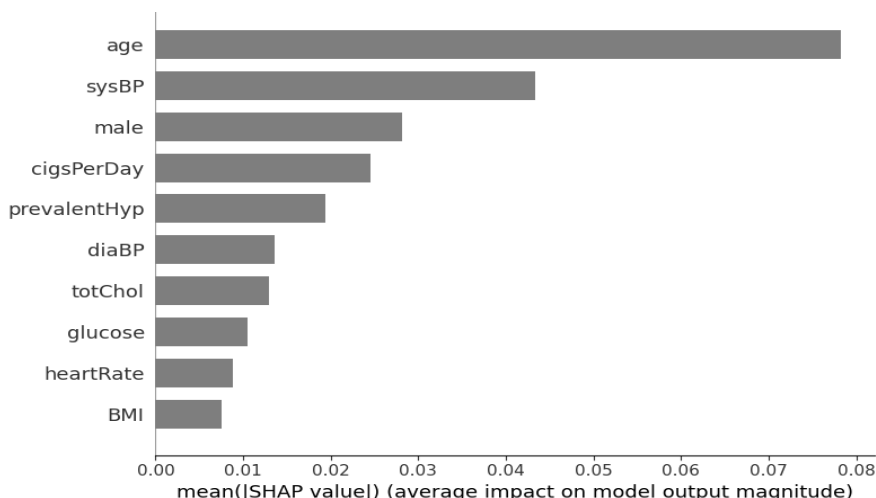


Figure 6: Proposed Model’s Explanation using SHAP analysis for Predicting Stroke

4.6 Comparison with Baseline Machine Learning Models

The proposed class-weighted Random Forest model is evaluated against Logistic Regression, Decision Tree, and AdaBoost classifiers. Table 4 presents the cross-validation performance metrics for these baseline models.

AdaBoost model demonstrates the best overall accuracy of 85.02% as compared to the other models that are tested. But an in-depth class-by-class analysis shows that this measure is hyper-inflated by the harsh data disproportion. The AdaBoost classifier had both a recall of 3% on the positive minority class and an F1-score of 6%. This means that there is almost total failure to detect real cases of heart disease.

On the other hand, the overall accuracies of the Logistic Regression and Decision Tree are 66.85% and 77.16%, respectively. By comparison, the Random Forest model is able to reach a trustworthy 81% overall accuracy. It can strike a moderate balance between class detection and maximizing its raw accuracy. The findings are clear evidence that overall accuracy is a grossly misleading factor when used in medical prediction exercises. A balanced approach is a prerequisite to meaningful clinical applications.

Table 4: Performance of Baseline Models

Model	Accuracy	F1-Score	ROC-AUC
Logistic Regression	66.85%	71.36%	72.34%
Decision Tree	77.16%	76.91%	54.77%
AdaBoost	85.02%	78.86%	70.37%
Proposed Random Forest	81.14%	81.00%	72.00%

5. Conclusion

This study developed and evaluated a class-weighted Random Forest model for ten-year coronary heart disease risk prediction using the Framingham Heart Study dataset. The proposed model achieved 81% accuracy on the hold-out test set and demonstrated stable performance under 5-fold stratified cross-validation ($78.64\% \pm 0.87\%$ accuracy; ROC-AUC: $70.19\% \pm 1.22\%$). SHAP-based analysis identified age, systolic blood pressure, sex, cigarette consumption, and prevalent hypertension as the most influential predictors, supporting the clinical interpretability of the model. Despite an overall satisfactory performance, class-wise evaluation revealed limited detection of positive cases, highlighting the challenges posed by the severe class imbalance. Comparison with Logistic Regression, Decision Tree, and AdaBoost models further demonstrated that overall accuracy alone may provide a misleading assessment of clinical usefulness. Overall, the proposed framework offers an interpretable and transparent approach for cardiovascular risk prediction while emphasizing the importance of class-wise evaluation. Future work should focus on external validation and the development strategies for minority-class detection.

Acknowledgment

The authors thank the editorial board of Journal of FST for the opportunity to publish our work, as well as the anonymous reviewers for their constructive comments and feedback which improved the quality of this manuscript.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Data Availability

The data and code that support the findings of this study are available from the corresponding author upon reasonable request.

Funding

This research received no external funding.

AI Declaration

The authors used generative AI tools to revise the manuscript, enhance its clarity and coherence, and correct typographical and grammatical errors. The authors reviewed and verified all contents and take full responsibility for the accuracy and integrity of the final manuscript.

References

- Di Cesare, M., Perel, P., Taylor, S., Kabudula, C., Bixby, H., Gaziano, T. A., Vaca McGhie, D., Mwangi, J., Pervan, B., Narula, J., Pineiro, D., & Pinto, F. J. (2024). The Heart of the World. *Global Heart*, 19(1). <https://doi.org/10.5334/gh.1288>.
- El-Sofany, H., Bouallegue, B., & 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(1), 1–18. <https://doi.org/10.1038/s41598-024-74656-2>.
- Hossain, S., Hasan, M. K., Faruk, M. O., Aktar, N., Hossain, R., & Hossain, K. (2024). Machine Learning Approach for Predicting Cardiovascular Disease in Bangladesh: Evidence from a Cross-Sectional Study in 2023. *BMC Cardiovascular Disorders*, 24(1). <https://doi.org/10.1186/s12872-024-03883-2>.
- Krishnan, S., Magalingam, P., & Ibrahim, R. (2023). Enhanced Recurrent Neural Network (RNN) for Heart Disease Risk Prediction Using Framingham Datasets. *Open International Journal of Informatics*, 11(1), 41–52. <https://doi.org/10.11113/oiji2023.11n1.242>.
- Mahmoud, W., Aborizka, M., & Amer, F. (2021). Heart Disease Prediction Using Machine Learning and Data Mining Techniques: Application of Framingham Dataset. *IDOSR Journal of Computer and Applied Sciences*, 6(1), 66–73. <https://www.idosr.org/wp-content/uploads/2021/11/IDOSR-JCAS-6166-73-2021>.

- Mienye, I. D., & Jere, N. (2024). Optimized Ensemble Learning Approach with Explainable AI for Improved Heart Disease Prediction. *Information*, 15(7), Article 394. <https://doi.org/10.3390/info15070394>.
- Mir, A., Rehman, A. U., Ali, T. M., Javaid, S., Almufareh, M. F., Humayun, M., & Shaheen, M. (2024). A Novel Approach for the Effective Prediction of Cardiovascular Disease Using Applied Artificial Intelligence Techniques. *ESC Heart Failure*. <https://doi.org/10.1002/ehf2.14942>.
- Nova, R. M., Fahim, J. A., & Seema, S. J. (2025). A study of machine learning models with population and swarm based optimization for cardiovascular disease prediction. *Journal of FST*, 3(1), 1–24. <https://doi.org/10.64494/JFST/v3i1/ss/2025/01/1-24>.
- O'Brien, R., & Ishwaran, H. (2019). A Random Forests Quantile Classifier for Class Imbalanced Data. *Pattern Recognition*, 90, 232–249. <https://doi.org/10.1016/j.patcog.2019.01.036>.
- Panattoni, G., Desimone, P., Toto, F., Meringolo, F., Jacomelli, I., Rebecchi, M., Cicogna, F., & Cal`o, L. (2025). Cardiovascular Risk Assessment in Daily Clinical Practice: When and How to Use a Risk Score. *European Heart Journal Supplements*, 27 (Supplement_1), i16–i21. <https://doi.org/10.1093/eurheartjsupp/suae100>.
- Patil, S. S., Dhumal, V., Gavale, S., Kulkarni, H., & Wadmalwar, S. (2022). Heart Disease Prediction Using Machine Learning. *International Journal of Scientific Research in Science and Technology*, 541–546. <https://doi.org/10.32628/IJSRST229676>.
- Prabu, S., Thiyaneswaran, B., Sujatha, M., Nalini, C., & Rajkumar, S. (2022). Grid Search for Predicting Coronary Heart Disease by Tuning Hyper-Parameters. *Computer Systems Science and Engineering*, 43(2), 737–749. <https://doi.org/10.32604/csse.2022.022739>.
- Setiyadi, D., Henderi, H., Suryaningrat, A., Swastika, R., Saludin, S., Mutoffar, M., & Yunianto, I. (2025). Prediction of Heart Disease Using Random Algorithm, Support Vector Machine, And Neural Network. *TELKOMNIKA (Telecommunication Computing Electronics and Control)*, 23(1), 129–138. <https://doi.org/10.12928/telkomnika.v23i1.25341>.
- Shipra, E. H. (2023). Unleashing the potential of machine learning algorithms for predicting strokes. *Journal of FST*, 2(1).

- Talaat, F. M., & Aly, W. F. (2025). Toward Precision Cardiology: A Transformer Based System for Adaptive Prediction of Heart Disease. *Neural Computing and Applications*, 37(19), 13547–13571. <https://doi.org/10.1007/s00521-025-11172-y>.
- Talukder, M. A., Talaat, A. S., & Kazi, M. (2025). HXAI-ML: A Hybrid Explainable Artificial Intelligence Based Machine Learning Model for Cardiovascular Heart Disease Detection. *Results in Engineering*, 25, Article 104370. <https://doi.org/10.1016/j.rineng.2025.104370>.
- Vu, T., Kokubo, Y., Inoue, M., Yamamoto, M., Mohsen, A., Martin-Morales, A., ... & Araki, M. (2025). Machine Learning Model for Predicting Coronary Heart Disease Risk: Development And Validation Using Insights from A Japanese Population–Based Study. *JMIR cardio*, 9(1), Article e68066.
- Yang, L., Wu, H., Jin, X., Zheng, P., Hu, S., Xu, X., Yu, W., & Yan, J. (2020). Study Of Cardiovascular Disease Prediction Model Based on Random in Eastern China. *Scientific Reports*, 10(1). <https://doi.org/10.1038/s41598-020-62133-5>.