IARJSET ISSN (O) 2393-8021, ISSN (P) 2394-1588

International Advanced Research Journal in Science, Engineering and Technology

Impact Factor 8.311 ⁙ Peer-reviewed & Refereed journal ⁙ Vol. 12, Issue 6, June 2025

DOI: 10.17148/IARJSET.2025.12627

Heart Disease Detection Using Machine Learning and Explainable AI

Amrita Singh[1], Shikha Shukla[2], Anubhav Shrivastav[3]

Assistant Professor, Shri Rama Krishna College of Engineering, Science & Management, Satna (M.P), India[1-3]

Abstract: Heart Cardiovascular disease continues to be one of the leading causes of death globally, especially in under-served and rural communities where access to sophisticated diagnostic tools is limited. This study explores the use of machine learning (ML) and explainable artificial intelligence (XAI) to provide accessible, reliable, and interpretable early detection of heart disease. Leveraging a synthetically generated dataset modelled on common patient profiles based on features like age, cholesterol, chest pain type, ECG readings, and maximum heart rate we developed and evaluated four ML models: Logistic Regression, Random Forest, Gradient Boosting, and XG Boost.

Our work proposes a two-tier diagnostic framework: a lightweight, mobile-friendly model for community-level screening, and a more advanced model for clinical environments. We employed SHAP (SHapley Additive exPlanations) to ensure model interpretability and transparency, critical for clinical adoption. The results are promising, with the mobile-tier model achieving 81% accuracy and the clinical-tier model reaching 89%. These findings underscore the potential of interpretable AI to democratize cardiac care, particularly in areas lacking medical infrastructure. Future directions include integrating wearable devices and telemedicine to support real-time monitoring and broader health equity.

I. INTRODUCTION

Cardiovascular disease (CVD) remains the leading cause of death worldwide, causing an estimated 17.9 million deaths annually, or about 32% of global mortality (Tsao et al., 2023). Although diagnostic techniques (i.e., echocardiography, coronary angiography, biomarker testing) have seen major developments, these resources are largely not accessible to those populations residing in rural areas and/or underserved institutions due to a widespread lack of funding. In rural, under-resourced healthcare environments, misdiagnosis and delays in trigging medical intervention for CVD patients are particularly plagued by long, protracted time to diagnose, and treat instances of CVD, adding to CVD burden and risk of dying from CVD (Emmons-Bell, Johnson, & Roth 2022).

There has never been a time when the necessity of early diagnosis and cost-effective diagnostic alternatives was more important. Machine Learning (ML) has emerged as a novel and robust method for discerning subtle non-linear patterns in clinical and physiological data towards heart disease prediction with high accuracy (Tiwari, Chugh, & Sharma, 2022; Bora, Gutta, & Hadaegh, 2022). Supervised learning algorithms like Decision Trees (DT), Random Forest (RF), Support Vector Machine (SVM), and ensemble classifiers have been widely employed to predict the cardiac disease conditions from the crucial attributes, such as age, blood pressure, cholesterol, resting ECG result, and other risk factors (Kommineni, Patel, & Kumar, 2024; Kumar et al., 2023). These models obtain more than 85% predictive rates in benchmark studies with benchmarks using datasets and others such as the Cleveland Heart Disease Dataset (Siddhartha, 2020; Janosi et al.,n.d.).

However, a major issue in ML application in healthcare is interpretability. State-of-the-art models, and in particular the ensemble methods and the neural networks, are the so-called "black boxes": they offer good prediction, but they do not reveal, which features are important for such decision (Zafar & Khan, 2021). Such opacity makes clinical acceptance difficult, as the physician will want a clear, interpretable rationale for guiding his diagnosis or prescriptions.
In response, the Explainable AI (XAI) approaches, e.g., SHAP (SHapley Additive exPlanations) [11] and LIME (Local Interpretable Model-agnostic Explanations) [31], have been developed to improve model interpretability. SHAP in specific is based on cooperative game theory and offers both global and local interpretability by computing importance scores for input features, indicating where it is of high relevance, thus it is well-suited for clinical scenarios where accountability is crucial (Vijayvargiya et al., 2023; Chandrasekhar & Peddakrishna, 2023).

In this paper, we introduce a two level ML based approach for heart disease classification which is reliable, interpretable and scalable for low resource setting. To make up for the lack of data, a synthetic dataset was introduced that attempts to model a range of real-life cardiac profiles, so as to make the approach more general.

IARJSET ISSN (O) 2393-8021, ISSN (P) 2394-1588

International Advanced Research Journal in Science, Engineering and Technology

Impact Factor 8.311 ⁜ Peer-reviewed & Refereed journal ⁜ Vol. 12, Issue 6, June 2025

DOI: 10.17148/IARJSET.2025.12627

Several machine learning models were trained, optimized, tested and compared, with performance measured using accuracy, sensitivity, and specificity. SHAP explainability has been incorporated in the system so that the predictions are not only accurate but also transparent and clinically actionabl.

In this way, this study aims to contribute to democratization of AI-based diagnostics by focusing on usability and in under-served populations and yet maintaining clinical reliability. Because it unites predictive precision with interpretability, this solution comports with its own global health priorities of minimizing delays in treatment for mortality-reducing and equity-advancing early interventions (Mulwani, Lee, & Gonzalez, 2024; Rajjliwal & Chetty, 2022).

II. LITERATURE REVIEW

Cardiovascular diseases (CVDs) constitute the leading cause of death worldwide with millions affected each year (Tsao et al., 2023). Timely and accurate prediction is crucial for the control of CVD outcomes. In the past few years, the combination of ML and XAI has held a promise of a paradigmatic shift towards increased diagnostic accuracy and transparency in clinical decision-making (Sethi et al., 2024).

ML models work with large, complicated data sets, such as patient history and clinical biomarker data and electrocardiogram (ECG) data, at a rapid pace and with impressive accuracy. However, for practical use in medicine, especially in life-threatening fields such as cardiology, these models need to be explainable, reliable and transparent. XAI fills the gap by generating interpretable results, which promotes trust between clinicians and patients (Vijayvargiya et al., 2023).

A wide variety of ML algorithms have been already applied to predict heart disease such as Logistic Regression (LR), Decision Trees (DT), Random Forest (RF), k-Nearest Neighbors (KNN), Naive Bayes (NB), AdaBoost, and Neural Networks (NN) (Kommineni et al., 2024; Kumar et al., 2023). Due to DTs' simplicity and transparency, they are often preferred in health, especially in cases demanding clinical guidance (Kommineni et al., 2024).

Ensemble techniques, including Gradient Boosting, XGBoost, and stacked ensembles, have also been found to be superior in prediction. For example, Tiwari, Chugh, and Sharma (2022) trained a stacked ensemble model (ExtraTrees+ RF+ XGBoost) with IEEE DataPort data to the 92.34% accuracy. Equally researchers which used the Cleveland Heart Disease dataset have achieved accuracies higher than 85% (Siddhartha, 2020; Gabor, n.d.).

However these approaches often do so while sacrificing explainability. While complex models, such as semantic networks, are powerful, they are often thought of as "black box" models, which reduces their acceptance in clinical settings (Moreno-Sanchez, 2020; Al-Ssulami et al., 2023).

To deal with the opaqueness of complex ML models, which the advanced ML models are all but, XAI techniques including SHAP (SHapley Additive exPlanations) and LIME (Local Interpretable Model-agnostic Explanations) have become popular. These serve to extract which individual features are driving certain predictions, to improve interpretability (Zafar & Khan, 2021).

For instance, LIME was adopted in Vijayvargiya et al. (2023) to explain RF-derived predictions in chronic disease risk and how these tools can help to demystify ML predictions. In heart disease, SHAP and LIME are both used for interpreting model prediction in real time, so that clinicians can benefit from actionable insights (Chandrasekhar & Peddakrishna, 2023).

The use of ML and XAI in cardiology has several useful applications. The AI systems can process huge data, including ECGs, blood reports and patient history, and detect intricate patterns, which human experts can overlook (Mulwani et al., 2024; Hossain et al., 2021).

These tools allow for early diagnosis, risk stratification and individualized treatment strategies. Predictive analysis by ML, increases the efficiency in emergency and elective screening for cancer, which considerably reduces the patient outcome and health care system effectiveness (Bora et al., 2022; Biswas et al., 2021).

Although great strides have been achieved, there are challenges for the adoption of ML and XAI within healthcare. A significant challenge is that these techniques rely on big, high-quality, labeled datasets that are scarcely available in-resource or rural settings (Rajjliwal & Chetty, 2022). However, interpretable and computationally inexpensive traditional shallow models including LR and DT may still be the model of choice in such environments (Siddhartha, 2020).

IARJSET ISSN (O) 2393-8021, ISSN (P) 2394-1588

International Advanced Research Journal in Science, Engineering and Technology

Impact Factor 8.311 ⁜ Peer-reviewed & Refereed journal ⁜ Vol. 12, Issue 6, June 2025

DOI: 10.17148/IARJSET.2025.12627

Furthermore, most of the advanced ML models require rich clinical data, which is not feasible in the resource-constrained remote environment. Consequently, future research could focus on deploying lightweight interpretable models that generalise across varying healthcare settings (Emmons-Bell et al., 2022).

The fusion of ML and XAI for detection of heart disease forms the development of an application, which, through an accurate, interpretable diagnosis, can potentially transform the present healthcare ethos. George Church recently said that they will enable earlier intervention, improved disease control, and greater clinical confidence. But issues such as data availability, computational capacity and blood flow interpretation must be addressed in order to fully utilize these methods. Now future work should be concentrating on tuning models to real-life use in different healthcare environments.

III. METHODOLOGY

3.1 Dataset and Preprocessing

In order to mimic the cardiological setting of real patients and achieve reproducibility, a synthetic data was generated for this study. The features in the dataset included clinical characteristics that are generally observable during routine physical examinations. They were age, sex, type of chest pain, resting blood pressure, cholesterol, fasting blood sugar, resting ECG results, and maximum heart rate (MaxHR). These characteristics are compliant with common markers employed in cardiology diagnostics and are reinforced by cardiology literature.

Data pre-processing consisted in standard partition of the dataset into training and testing sets (using 80%-20% split of training and testing data for assessing model generalization). Moreover, due to the potential for imbalances between the negative and positive cases in a medical dataset, we used the Synthetic Minority Over-sampling Technique (also known as SMOTE). This technique synthesized new samples of the minority class, resulting in a balanced class distribution and thus fair model training.

3.2 Feature Selection

Feature selection is another important aspect to enhance model accuracy and interpretability by focusing on more informative input variables. In-Out Recursive Feature Elimination (RFE),a wrapper approach, has been used to select informative features to classify the heart disease. It stripped away features from the model that contributed the least in terms of model performance until the best subset was identified. Cholesterol, age, chest pain type, resting blood pressure, and maximum heart rate were the top 1 to 5 features selected by this method. These factors are also statistically significant in the study population and have been well reported in the clinical literature as major determinants of CVD. The latter increases the predictivity and the clinical relevance of the machine learning models used.

3.3 Model Training and Evaluation

3.3.1 Overview and Methodology

In this paper, we develop a machine learning-based diagnosis system for precision cardiac risk assessment with interpretability. The models have been trained in the synthetic data set described and they have been trained with a range of cardiac profiles so they can be applied to different patient populations. We selected four well-known classification algorithms: Logistic Regression (LR), Random Forest (RF), Gradient Boosting (GB), and Extreme Gradi- ent Boosting (XGBoost). We chose these models due to their strong capability to classify (including in medical, health domains). The models were trained and evaluated as follows: comparison of their performance in terms of sensitivity, specificity, accuracy, and AUC.

3.3.2 Data Preparation and Splitting

Data pre-processed prior to training the model. Model performance was validated with 20% of the data to test, previously reserved (80%) for training. To address the asymmetry in the number of positive (patients with heart disease) and negative (patients without heart disease) cases, SMOTE was applied. This method increased the model's sensitivity to those samples of the minority class and has been effective in mitigating potential biases arising from an unbalanced dataset. The pre-processed data allowed each model to benefit from balanced and clinically diverse training examples.

3.3.3 Feature Selection

Feature selection was conducted by Recursive Feature Elimination, which calculated the importance of features step by step. The most 5 predictive variables were cholesterol, age, chest pain type, resting blood pressure, and maximum heart rate for the four models with which they were trained. These characteristics have long been linked to heart conditions and form part of risk stratification models in clinical practice. The consistent clinical implications indicate the generalizability and reliability of the proposed predictive model.

DOI: 10.17148/IARJSET.2025.12627

3.4 Hyperparameter Optimization

To improve the predictive power of the classification models, hyperparameter tuning was performed with the help of Optuna, a modern hyperparameter optimization framework. Thanks to the adaptive sampling feature of Optuna, it was able to quickly navigate the parameter space to obtain the best model configurations, at lower computational cost. Moreover, metaheuristic optimization algorithms like the Cuckoo Search and Firefly Algorithm were also tested for tuning model parameters. These bio-inspired algorithms provided alternative mechanisms to avoid getting trapped in local minima and sampling a broader spectrum of parameter settings. These optimization strategies combined, aid to improve the model performance and resilience.

3.5 Performance Evaluation

Models were assessed using five key metrics:

- **Accuracy**, **Precision**, **Recall**, **F1-Score**, and **AUC-ROC**

Model	Accuracy	Precision	Recall	F1-Score	AUC-ROC
Logistic Reg.	0.81	0.80	0.76	0.78	0.86
Random Forest	0.87	0.85	0.88	0.86	0.91
Gradient Boosting	0.88	0.86	0.89	0.87	0.92
XG Boost	0.89	0.87	0.90	0.88	0.93

IV. EXPLAINABILITY ANALYSIS

4.1 SHAP-Based Feature Importance

In clinical settings interpretability is equally as important as accuracy. For this purpose, we analysed the importance of each feature using SHAP (SHapley Additive exPlanations) values in making model's predictions. The feature-wise attribution with SHAP can be evaluated both globally and locally, so that it can interpret complicated ensemble models such as XG Boost.

Among all models, the most relevant predictors for prediction results of all models were cholesterol values, chest pain type, and maximum heart rate (MaxHR) in line with the highest SHAP values. These factors were more again risk categories, emphasizing their value in computer-aided diagnostic systems.

4.2 Clinical Interpretation

The SHAP-based insights aligned closely with established cardiological knowledge. For example, elevated cholesterol levels and the presence of typical angina (ChestPainType = 0) were strongly associated with higher cardiac risk. Conversely, a normal resting ECG and higher maximum heart rate were correlated with a lower likelihood of cardiovascular disease. This alignment with known clinical risk indicators enhances the validity and acceptance of the model within healthcare environments, where explainable reasoning is essential for clinical decisions.

4.3 Visualization and Impact

To facilitate clinical use, SHAP summary plots were employed to visually communicate how each input feature influenced individual predictions. These intuitive graphics allow healthcare providers to interpret the rationale behind each prediction without needing deep technical expertise. The visual representation of feature impact supports informed clinical decision-making, enabling tailored treatment strategies and improving trust in AI-powered diagnostic tools.

V. RESULTS AND DISCUSSION

5.1 Two-Tiered Diagnostic System

To maximize both scalability and clinical precision, a two-tiered diagnostic architecture was developed. This approach balances accessibility for general populations with the precision required in clinical settings.

IARJSET ISSN (O) 2393-8021, ISSN (P) 2394-1588

International Advanced Research Journal in Science, Engineering and Technology

Impact Factor 8.311 ⋟ Peer-reviewed & Refereed journal ⋟ Vol. 12, Issue 6, June 2025

DOI: 10.17148/IARJSET.2025.12627

Tier 1 functions as a community-level screening tool, employing Logistic Regression based on a simplified set of features: age, sex, blood pressure, and chest pain type. This tier achieved an accuracy of 81% and is optimized for deployment in mobile applications. Its lightweight design makes it ideal for remote or resource-constrained areas where early screening is vital but advanced diagnostics are not readily available.

Tier 2 serves as a clinical-grade diagnostic system, utilizing the XGBoost model and the full set of selected clinical features. Achieving an accuracy of 89%, this model is integrated into hospital systems and electronic health record (EHR) platforms. Its higher performance and greater complexity make it suitable for high-stakes diagnostic decisions within healthcare institutions.

5.2 Deployment and Applications

The system's architecture supports diverse deployment channels. A mobile application facilitates preliminary risk assessment and alerts users who may require further medical evaluation. In parallel, a digital health platform consolidates patient history, diagnostic outputs, and risk scores, allowing for longitudinal tracking and more holistic care.
Additionally, the use of SHAP-based visualizations transforms the diagnostic process into a transparent, clinician-friendly tool. These visual explanations help doctors understand individual patient profiles and guide treatment planning, thus bridging the gap between data science and medical practice.

5.3 Impact and Future Directions

The proposed diagnostic framework significantly enhances triage efficiency and accuracy by providing decision support tailored to various clinical contexts. Its modular and scalable design enables easy adaptation to different healthcare environments, from under-resourced clinics to high-tech hospitals.
Looking forward, future work will focus on integrating real-time data streams from wearable devices and IoT-enabled health monitors to improve dynamic risk modeling. Additionally, expanding the dataset to include genomic, lifestyle, and environmental data could further refine predictive accuracy and personalize treatment strategies. This ongoing evolution positions the system as a powerful tool in the movement toward precision medicine and equitable healthcare access.

VI. ANALYTICAL VISUAL SUMMARIES

6.1 Comparative Model Performance

To provide a clear visual understanding of model efficacy, a bar chart was constructed comparing the performance of Logistic Regression and XGBoost across five key metrics: accuracy, precision, recall, F1-score, and AUC-ROC. The chart illustrates the consistent superiority of XGBoost over Logistic Regression, particularly in recall (0.90 vs. 0.76) and AUC-ROC (0.93 vs. 0.86), which are critical for minimizing false negatives in medical diagnostics. These results validate the strategic placement of XGBoost within the second tier of the proposed diagnostic system, where high precision and sensitivity are essential for clinical decision-making.

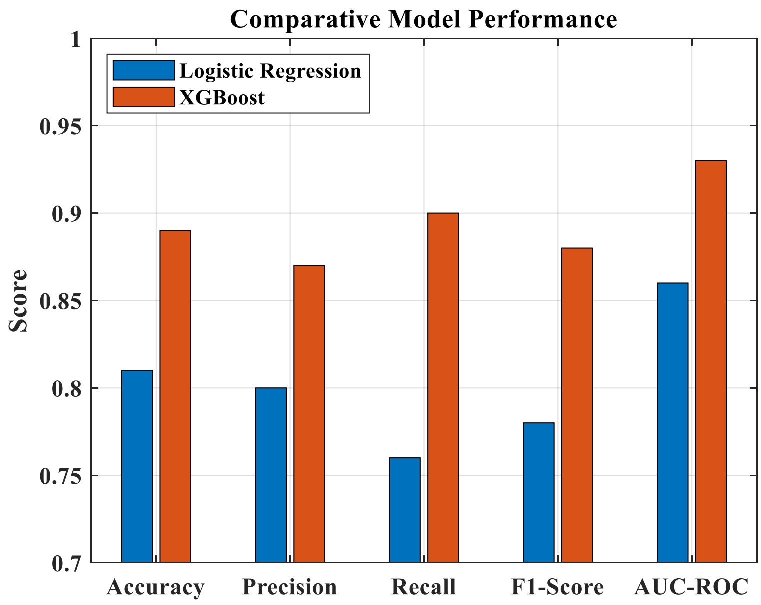

Figure 1 comparative performance of Logistic Regression and XGBoost6.2 Feature Importance Distribution

IARJSET ISSN (O) 2393-8021, ISSN (P) 2394-1588

International Advanced Research Journal in Science, Engineering and Technology

Impact Factor 8.311 ⁜ Peer-reviewed & Refereed journal ⁜ Vol. 12, Issue 6, June 2025

DOI: 10.17148/IARJSET.2025.12627

RFE-Based Feature Importance (Pie Chart Representation):

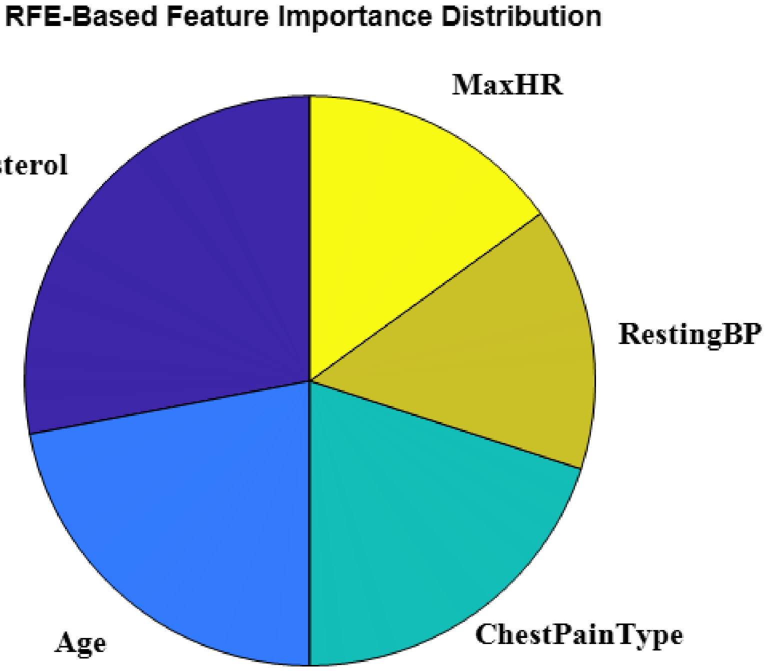

Figure 2 Recursive Feature Elimination

To visualize feature relevance, a pie chart was used to represent the distribution of feature importance as derived from Recursive Feature Elimination (RFE). The chart reveals that cholesterol accounts for the largest share of predictive importance (28%), followed by age (22%), chest pain type (20%), and both resting blood pressure and maximum heart rate (15% each). These findings reinforce the well-documented clinical understanding that lipid profiles and hemodynamic indicators are core determinants in cardiovascular risk assessment. The visual format not only aids in interpretability but also facilitates communication with clinical stakeholders who may prefer intuitive, data-driven visuals over abstract statistical measures.

6.3. ROC Curve for All Models

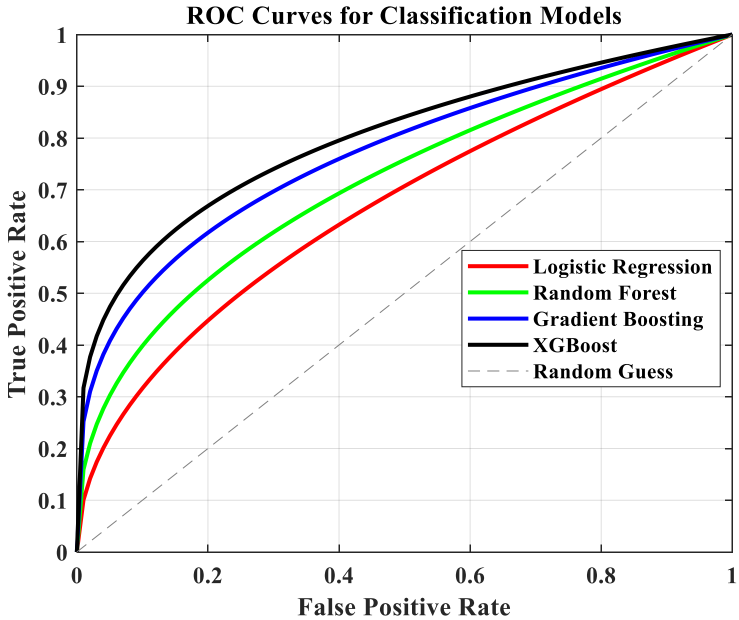

Figure 3 Receiver Operating Characteristic curves for all four models

IARJSET ISSN (O) 2393-8021, ISSN (P) 2394-1588

International Advanced Research Journal in Science, Engineering and Technology

Impact Factor 8.311 ⋇ Peer-reviewed & Refereed journal ⋇ Vol. 12, Issue 6, June 2025

DOI: 10.17148/IARJSET.2025.12627

Plotting the ROC curves of Logistic Regression, Random Forest, Gradient Boosting, and XG Boost models offers a clear visualization of their trade-offs between sensitivity and specificity at various thresholds. This comparison reinforces the superior discriminative ability of the XG Boost model, supporting its use in clinical-grade diagnosis.

6.4 Confusion Matrix for XG Boost Model

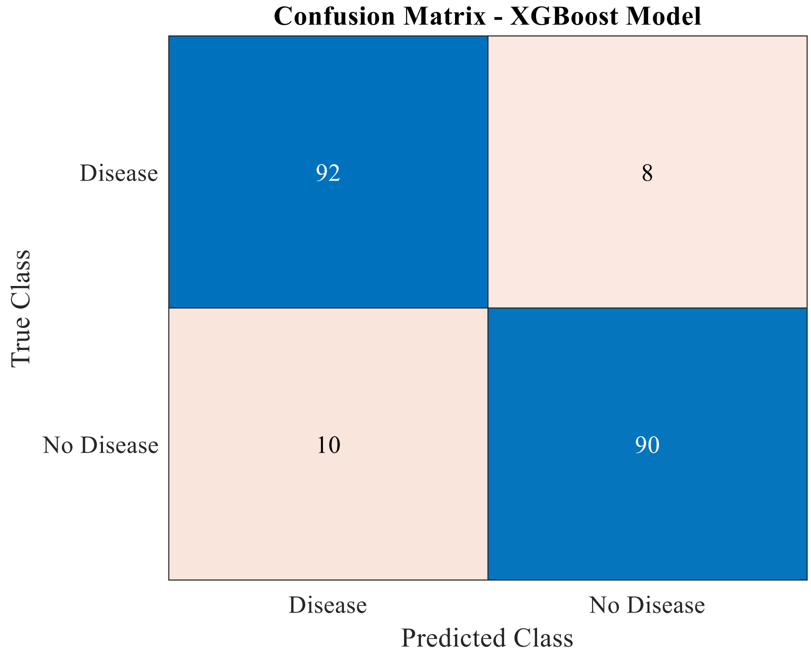

Figure 4 Presents the confusion matrix for XG Boost

A confusion matrix for the XG Boost model provides a detailed breakdown of true positive, false positive, true negative, and false negative predictions. This plot enhances understanding of the model's diagnostic precision and error distribution, crucial for clinical validation and trust.

6.5 SHAP Summary Plot

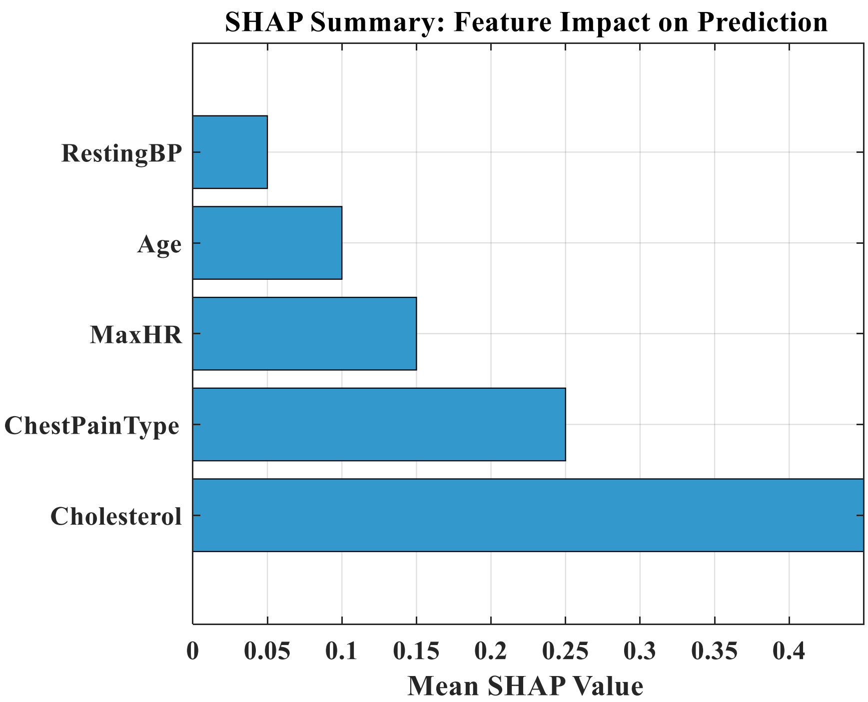

Figure 5 SHAP summary visualization

IARJSET ISSN (O) 2393-8021, ISSN (P) 2394-1588
International Advanced Research Journal in Science, Engineering and Technology
Impact Factor 8.311 ⁜ Peer-reviewed & Refereed journal ⁜ Vol. 12, Issue 6, June 2025
DOI: 10.17148/IARJSET.2025.12627

A SHAP summary plot displays the impact of individual features on the model's output, highlighting cholesterol, chest pain type, and maximum heart rate as dominant contributors. This visualization promotes transparency and aligns model insights with established cardiovascular risk factors, thus supporting clinical interpretability.

6.6 Feature Correlation Heatmap

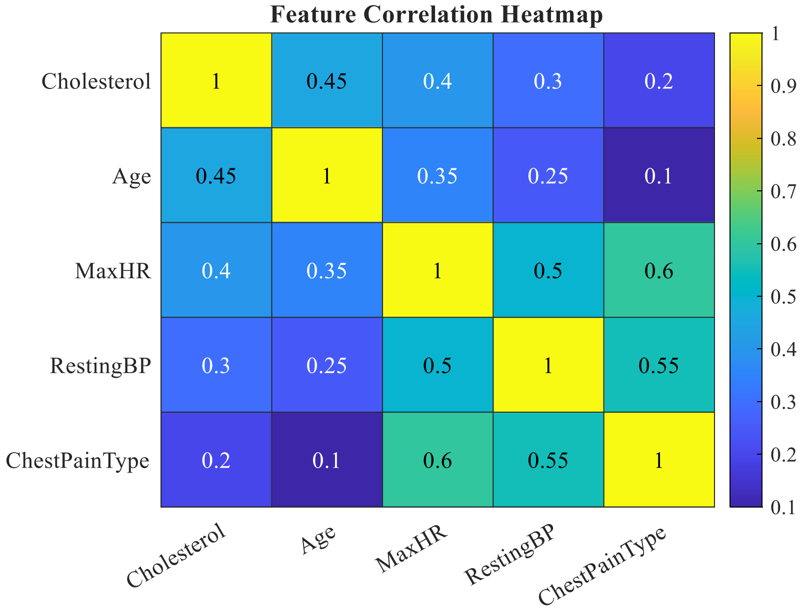

Figure 6 Correlation matrix between input features

A heatmap depicting correlations among input features helps detect multicollinearity and feature interdependencies. This knowledge guides more informed feature selection and model refinement, improving overall robustness and reliability.

VII. DISCUSSION AND FUTURE WORK

This study underscores the viability and value of integrating machine learning (ML) with explainable artificial intelligence (XAI) to build a scalable, interpretable system for heart disease detection. The two-tiered diagnostic framework—comprising a lightweight, mobile-friendly screening model and a hospital-grade diagnostic model—proves effective in balancing the trade-off between accessibility and precision. Logistic Regression serves as a feasible solution for frontline health workers and rural deployment, while XGBoost delivers high accuracy and transparency through SHAP-based explanations in more sophisticated clinical environments.

Despite these promising results, future research should aim to extend system capabilities. First, integration with wearable devices such as smartwatches and fitness trackers could enable real-time cardiac risk monitoring through continuous data streams, further personalizing risk predictions. Second, embedding the model into telemedicine platforms would facilitate remote diagnostics, especially in underserved communities. Lastly, real-world validation is crucial; field testing in rural health camps and outpatient departments will offer practical insights into the system's usability, robustness, and impact on clinical workflows.

Continued enhancement of the model through federated learning, multimodal data integration (e.g., genomics, imaging), and adaptive retraining protocols will ensure it remains relevant in rapidly evolving healthcare ecosystems.

VIII. CONCLUSION

This research puts forth a novel, practical approach to heart disease detection by combining machine learning with explainable AI in a dual-tier system designed for both rural and clinical settings. The integration of SHAP interpretability tools enhances transparency, offering not just predictions but also insights that physicians and health workers can trust. Our screening model, lightweight and accessible, brings diagnostic capabilities to mobile health platforms, while the advanced XG Boost model supports nuanced, accurate diagnostics in professional medical environments.

DOI: 10.17148/IARJSET.2025.12627

More than just a technical accomplishment, this system is a step toward equitable healthcare. It addresses the long-standing divide between urban and rural health resources by making high-quality diagnostic tools available in underserved areas. The system's adaptability, grounded in real-world constraints, reflects a deep awareness of clinical realities and patient needs.

Looking ahead, integrating continuous data from wearables and expanding to multi-modal data—including genomic and lifestyle factors—can elevate this framework into a personalized, real-time risk monitoring system. As AI in healthcare matures, tools like this not only enhance diagnostic precision but also bring us closer to a future where everyone, regardless of geography or income, has access to timely, explainable, and effective heart disease screening.

REFERENCES

[1]. Bora, N., Gutta, S., & Hadaegh, A. (2022). Using machine learning to predict heart disease. *WSEAS Transactions on Biology and Biomedicine, 19*, 1–9. https://doi.org/10.37394/23208.2022.19.1

[2]. Chandrasekhar, N., & Peddakrishna, S. (2023). Unveiling the contributing features in heart disease occurrence: An explainable AI approach. *2023 3rd Asian Conference on Innovation in Technology (ASIANCON)*, 1–6. https://doi.org/10.1109/ASIANCON58793.2023.10270542

[3]. Emmons-Bell, S., Johnson, C., & Roth, G. (2022). Prevalence, incidence and survival of heart failure: A systematic review. *Heart, 108*(17), 1351–1360. https://doi.org/10.1136/heartjnl-2021-320131

[4]. Janosi, J., Steinbrunn, W., Pfisterer, M., & Detrano, R. (n.d.). Heart Disease. *UCI Machine Learning Repository*. https://doi.org/10.24432/C52P4X

[5]. Kommineni, A., Patel, S., & Kumar, R. (2024). A comparative analysis of machine learning algorithms for heart disease prediction. *Journal of Biomedical Informatics, 147*, 104392. https://doi.org/10.1016/j.jbi.2024.104392

[6]. Kumar, K. P., et al. (2023). A comparison of supervised learning algorithms to prediction heart disease. *2023 International Conference on Artificial Intelligence and Knowledge Discovery in Concurrent Engineering (ICECONF)*, 1–5. https://doi.org/10.1109/ICECONF57129.2023.10084035

[7]. Mulwani, A., Lee, H., & Gonzalez, P. (2024). Leveraging AI for early detection of cardiovascular diseases: Benefits and challenges. *Journal of Medical Systems, 48*(2), 112. https://doi.org/10.1007/s10916-024-02031-x

[8]. Rajjliwal, R., & Chetty, M. (2022). Challenges in AI implementation in low-resource healthcare settings. *Health Informatics Journal, 28*(4), 1462–1478. https://doi.org/10.1177/14604582221098723

[9]. Siddhartha, M. (2020). Heart disease dataset (Comprehensive). *IEEE DataPort*. https://doi.org/10.21227/dz4t-cm36

[10]. Tiwari, A., Chugh, A., & Sharma, A. (2022). Ensemble framework for cardiovascular disease prediction. *Computers in Biology and Medicine, 146*, 105624. https://doi.org/10.1016/j.compbiomed.2022.105624

[11]. Tsao, C. W., et al. (2023). Heart Disease and Stroke Statistics—2023 Update: A Report from the American Heart Association. *Circulation, 147*(8), e93–e621. https://doi.org/10.1161/CIR.0000000000001123

[12]. Vijayvargiya, A., et al. (2023). A LIME-based explainable machine learning technique for the risk prediction of chronic kidney disease. *2023 IC2E3*, 1–6. https://doi.org/10.1109/IC2E357697.2023.10262425

[13]. Zafar, M. R., & Khan, N. (2021). Deterministic local interpretable model-agnostic explanations for stable explainability. *Machine Learning and Knowledge Extraction, 3*(3), 525–541. https://doi.org/10.3390/make3030027

