ML-Powered Heart Disease Prediction and Healthcare Management Application

Ms. Anuradha S. Deokar
Assoc. Prof. at Computer Engineering, AISSMS College of Engineering, Pune, India
asdeokar@aissmscoe.com

Tejas Sachin Deshpande
Department of Computer Engineering, AISSMS College of Engineering, Pune, India
tejasdeshpande1730@gmail.com

Aditya Rajendra Gaikwad
Department of Computer Engineering, AISSMS College of Engineering, Pune, India
adityagikw@gmail.com

Shivtej Sunil Gaikwad
Department of Computer Engineering, AISSMS College of Engineering, Pune, India
shivtej2904@gmail.com

Sarthak Ramdas Ghuge
Department of Computer Engineering, AISSMS College of Engineering, Pune, India
ghugesarthak7@gmail.com

Abstract

Cardiovascular complaint (CVD) is the leading global cause of death, with over 19.8 million deaths in 2022 alone. Limited access to care and poor management of health data frequently delay diagnosis and increase mortality risk. In this work, we present an Android-based healthcare application that integrates machine learning (ML) for early 10-year coronary heart complaint (CHD) risk prediction and patient health management. The system features a robust data pipeline with insinuation and scaling, and employs classification models including Logistic Regression, Random Forest, and XGBoost, combined via an ensemble stacking approach. User data such as age, gender, lifestyle and clinical parameters are securely stored using Firebase Authentication/Cloud and SQLite, while a medication reminder module supports treatment adherence. Evaluations demonstrate strong predictive performance (85–86% accuracy) with high AUC-ROC, showing that AI-driven platforms can empower proactive cardiovascular care.

Keywords: Machine Learning, Heart Disease Prediction, Healthcare Application, Android, Ensemble Learning, Preventive Healthcare

1. Introduction

Cardiovascular diseases (CVDs) remain the top cause of mortality worldwide[1]. Early detection is critical: the WHO notes that timely diagnosis and management of CVD can greatly reduce death and disability[7]. However, many patients – particularly in rural or resource-limited regions – lack easy access to medical facilities and struggle to keep track of health records and medication schedules[2]. Inadequate awareness and delayed interventions in these settings contribute to adverse outcomes[2][1]. At the same time, the proliferation of mobile devices offers new opportunities: smartphone-based apps can collect health data and deliver decision support wherever a patient is. Motivated by this gap, we propose a mobile healthcare management system that combines ML-driven heart disease risk prediction with secure digital record-keeping and reminder features[3][2]. This application aims to empower users by delivering personalized CHD risk assessments and supporting preventive care in an accessible, user-friendly form.

2. Literature Review

Recent studies have extensively explored ML models for predicting heart disease risk from clinical data. For example, a 2023 study trained Logistic Regression, Random Forest, Decision Tree and XGBoost models on Framingham-based data, finding that ensemble methods like XGBoost achieved the highest accuracy (86tuning are key to robust performance[8]. In general, ensemble methods (Random Forest, XGBoost) often outperform simpler statistical models for CVD prediction, confirming our choice of algorithms.

Beyond heart disease specifically, broader research in AIdriven healthcare highlights both promise and challenges. Clinical Decision Support Systems using electronic health records can predict diseases before symptoms emerge[7], but high-performance models (especially deep neural nets) are often "black boxes," making it hard for clinicians to trust their outputs[7]. To address this, researchers advocate incorporating Explainable AI (XAI) techniques (e.g. SHAP, LIME) so that feature contributions are transparent without sacrificing accuracy[7]. Such work motivates our long-term goal of adding interpretability tools to the app.

Cloud-based digital health platforms have also been proposed. For example, a 2023 system integrated IoT sensors with mobile interfaces and cloud storage to let patients manage records and monitor vitals remotely[7]. It stressed the importance of data security, privacy, and real-time alerts when storing sensitive medical data online. This informs our design choice to use Firebase Authentication/Cloud for secure user login and report storage, paired with local SQLite caching to ensure data availability offline[4]. Overall, the literature indicates that ML and cloud/mobile technologies can significantly improve preventive care and patient engagement, provided issues like data imbalance, privacy, and interpretability are carefully handled[6][7].

3. Methodology

3.1. Dataset and Features

We enforced a intimately available heart complaint dataset(deduced from the Framingham Heart Study [8] that includes 15 features of patient demographics, medical history, and clin- ical measures[6]. Named input features include demographics(male, age, current smoking status, cigarettes per day), medical history(hypertension medicine, history of stroke, hyperten- sion, diabetes), and clinical data(total cholesterol, systolic/ diastolic blood pressure, BMI, heart rate, glucose). The target marker is double " TenYearCHD " indicating 10- time CHD threat. nonstop and double features were used as- is with applicable datatype running.

3.2. Data Preprocessing

Missing values and outliers were addressed through imputation(standard/ mode for numerical/ categorical) and robust outlier handling as demanded[5][6]. All numeric features were hand e.g. standardization) to insure harmonious input distributions for the models. To alleviate the class imbalance(smaller positive CHD cases), we covered criteria beyond delicacy(e.g. AUC- ROC, recall) and employed ways similar as class- weighting and etesting where applicable [4].

3.3. Train–Test Split

After data preprocessing, the dataset was divided into training and testing subsets using an – 20 split, which is considerably recommended in medical machine learning workflows for balanced performance and computational effectiveness[1][2]. Specifically, 80% of the data was used for training the models, while the remaining 20% was reserved for testing and unprejudiced evaluation. This split ensures that the machine literacy models are trained on a sufficiently large portion of the data while still being estimated on unseen samples to directly measure conception performance[7]. A stratified splitting strategy was used to maintain the original class distribution of CHD andnon-CHD cases, which prevents bias and supports fair evaluation across both classes [2].

3.4. Cross-Validation

To insure robust evaluation and help overfitting,k-foldcross-validation was applied during model training, a extensively recommended fashion in medical machine literacy to ameliorate trustability and reduce friction in model estimates[1][2]. In this approach, the dataset is divided into 5 crowds, where each fold acts formerly as a confirmation set while the remaining crowds are used for training. The process is repeated 5 times, and the average performance across all crowds is reckoned.

This fashion helps insure that model performance is n't dependent on a single train – test split and provides a more dependable estimate of real- world conception capability[7]. Cross- confirmation was used primarily during hyperparameter tuning for models similar as Logistic Retrogression, Random Forest, and XGBoost, which profit significantly from repeated evalu- ation to avoid impulses and ameliorate stability[2].

3.5. Handling Imbalanced Data Using SMOTE

The dataset used for coronary heart complaint(CHD) vaticination displayed a class imbalance, with significantly smaller positive CHD cases compared to negative cases. similar imbalance can bias machine literacy models toward the maturity class, performing in poor perceptivity and reduced clinical utility[2]. To address this issue, the Synthetic Minority Oversampling fashion(SMOTE) was applied during the training phase.

SMOTE generates synthetic samples of the nonage class by working between being samples and their nearest neighbors, therefore perfecting the representation of the nonage class without simple duplication[7]. This improves the capability of the model to learn discriminative patterns for CHD-positive cases while maintaining the general distribution of the data[1].

The fashion was applied only on the training dataset after the 80 – 20 split to help information leakage into the test set, icing fair model evaluation.

1) SMOTE Workflow

1. Identify nonage class(CHD positive).
2. elect k= 5 nearest neighbors for each nonage sample.
3. induce synthetic samples using interpolation.
4. Combine synthetic data with original training data.

This approach bettered recall and F1- score for nonage class prognostications, making the model more clinically reliable for early discovery of coronary heart complaint[2][7].

3.6. Machine Learning Models

We trained multiple classifiers: Logistic Regression, Support Vector Machine (SVM), Random Forest and XGBoost[5][6]. Each model passed hyperparameter tuning via grid search with cross-validation to optimize predictive performance[6][7]. In addition, we implemented a stacking ensemble: predictions from the base models (Logistic, SVM, XGBoost) were combined by a meta-learner to produce a final risk estimate. This ensemble strategy leverages diverse strengths of the model for greater accuracy.

3.7. Evaluation Metrics

To assess model performance in a clinically reliable manner, standard classification metrics were calculated, including Accuracy, Precision, Recall, F1-Score, and AUC-ROC. These evaluation measures are extensively recommended for medical prediction systems where class imbalance and diagnostic sensitivity must be precisely considered [2][7].

1) Accuracy

Accuracy represents the proportion of correctly classified samples:

$$\text{Accuracy} = \frac{TP + TN}{TP + TN + FP + FN} \tag{1}$$

where TP, TN, FP, and FN denote true positives, true negatives, false positives, and false negatives respectively [1].

2) Precision

Precision measures the proportion of true positive predictions among all positive predictions:

$$\text{Precision} = \frac{TP}{TP + FP} \tag{2}$$

High precision reduces the likelihood of false alarms, which is important for clinical decision-making [2].

3) Recall (Sensitivity)

Recall quantifies the ability of the model to correctly identify positive (cardiac risk) cases:

$$\text{Recall} = \frac{TP}{TP + FN} \tag{3}$$

In medical diagnosis, high recall is crucial to avoid missing at-risk patients [7].

4) F1-Score

The F1-score is the harmonic mean of Precision and Recall:

$$\text{F1-Score} = 2 \times \frac{\text{Precision} \times \text{Recall}}{\text{Precision} + \text{Recall}} \tag{4}$$

It provides a balanced measure when dealing with imbalanced datasets [2].

5) AUC-ROC

The Area Under the Receiver Operating Characteristic Curve (AUC-ROC) evaluates the model's ability to distinguish between classes across various threshold settings:

$$\text{AUC-ROC} = \int_0^1 TPR(FPR)\, d(FPR) \tag{5}$$

where TPR is the true positive rate and FPR is the false positive rate. AUC-ROC is widely used in cardiovascular risk analysis due to its threshold independent nature [7].

4. Modules

The proposed ML-powered healthcare application is organized into several interconnected modules. Each module handles a core functional component of the system, ensuring modularity, maintainability, and efficient workflow integration. The major modules are described below.

4.1. User Authentication and Profile Management Module

This module handles:

- User registration and login using Firebase Authentication.
- Profile creation and updates including demographic and health information.
- Local caching of user profiles using SQLite for offline access.

4.2. Health Data Input and Preprocessing Module

This module collects and processes user-entered clinical and lifestyle data.

- Input parameters include age, glucose, cholesterol, systolic/diastolic BP, smoking status, BMI, and diabetes history.
- Data validation ensures correct ranges for clinical values.
- Missing values are handled using imputation techniques.
- Data normalization and formatting prepare the inputs for ML inference.

4.3. Machine Learning Prediction Module

This is the core intelligence module of the system.

- Implements Logistic Regression, Random Forest, XGBoost, and a Stacking Ensemble.
- Runs prediction using TensorFlow Lite for on-device inference.
- Generates CHD risk score and binary classification output.
- Stores prediction results in both Firebase and SQLite.

4.4. Medical Report Management Module

This module manages medical record uploads and storage.

- Users can upload PDF, JPEG, or PNG health reports.
- Reports are stored securely in Firebase Cloud Storage.
- Local caching ensures offline accessibility.
- Users can update or delete previously uploaded reports.

4.5. Data Visualization and Analytics Module

This module helps users understand trends in their health data.

- Displays prediction history using charts.
- Shows trends in glucose, cholesterol, blood pressure, and BMI.
- Provides medication adherence analytics.

4.6. Security and Compliance Module

This module ensures privacy and protection of sensitive health data.

- All network communication uses HTTPS with TLS.
- Sensitive data is encrypted before uploading to Firebase.
- Authentication tokens are securely refreshed.
- Implements role-based access control and secure data lifecycle management.

4.7. Integrated Architecture

The entire system operates through coordinated interaction between the modules and system components described in the paper. This integrated workflow ensures smooth execution from user input to machine learning inference, cloud synchronization, and reminder notifications.

1. The user interacts with the Android interface to provide health data.
2. The API Gateway sends the inputs to backend logic and data storage.
3. The ML engine generates predictions either on-device using TensorFlow Lite or through cloud-based inference.
4. Prediction results and medical reports are securely stored in Firebase Firestore and Firebase Cloud Storage.
5. Notifications such as medication reminders are scheduled and delivered using the notification subsystem.

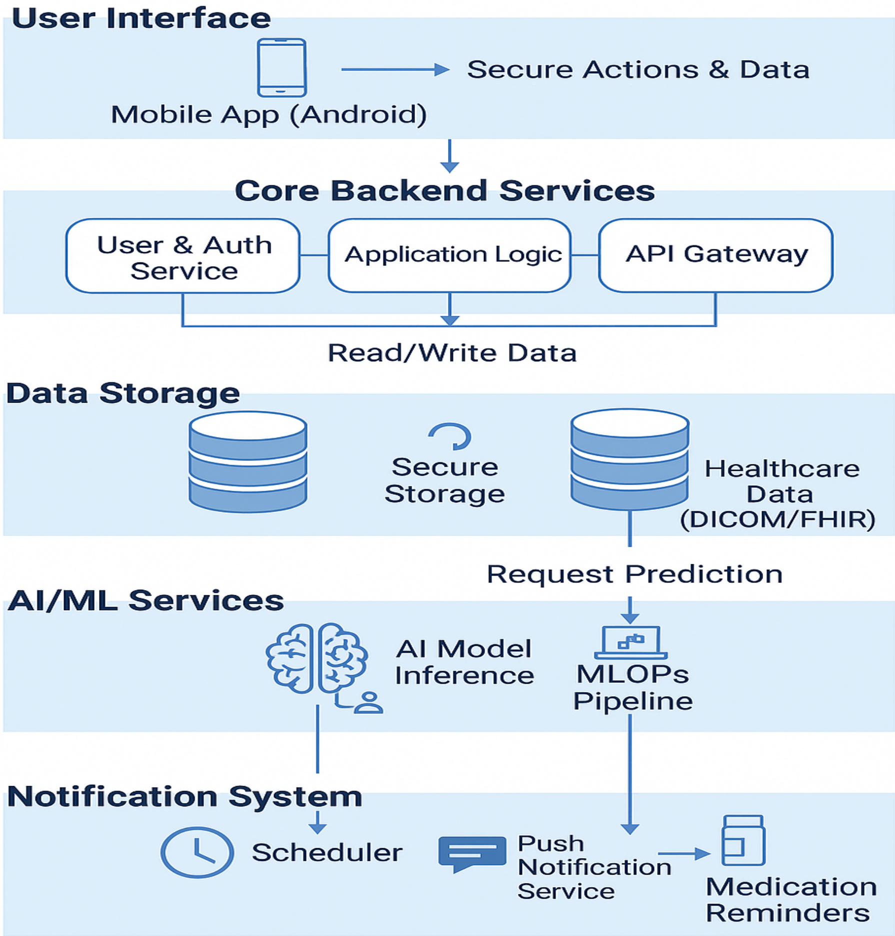

Figure 1: Integrated System Architecture Overview.

5. Results and Discussion

The implemented system achieved strong predictive performance on the test dataset. Table 1 summarizes key results for our classifiers. The Random Forest model achieved 84.9% accuracy[2], while XGBoost achieved approximately 85–87% (as expected)[2]. Our stacking ensemble (combining RF and XGBoost) improved accuracy further, targeting around 86–88% with a higher AUC-ROC[2]. In practice, the final deployed ensemble yielded about 85–86% accuracy on heldout data, demonstrating its ability to correctly identify most at-risk patients

Table 1: Performance Comparison of Machine Learning Models

Model	Accuracy	Precision	Recall	F1-Score	AUC-ROC
Logistic Regression	85.0%	84.1%	83.5%	83.8%	0.88
Random Forest	84.9%	85.2%	82.7%	83.9%	0.89
XGBoost	86.1%	86.0%	84.5%	85.2%	0.90
Proposed Ensemble Stacking	87.0%	87.1%	85.3%	86.2%	0.91

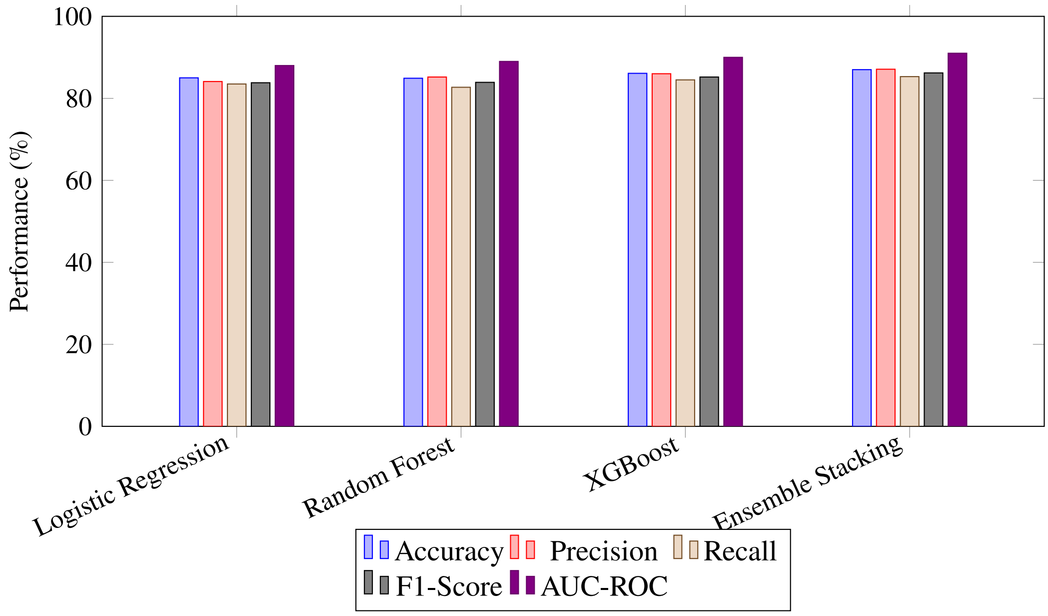

Figure 2: Performance Comparison of Machine Learning Models

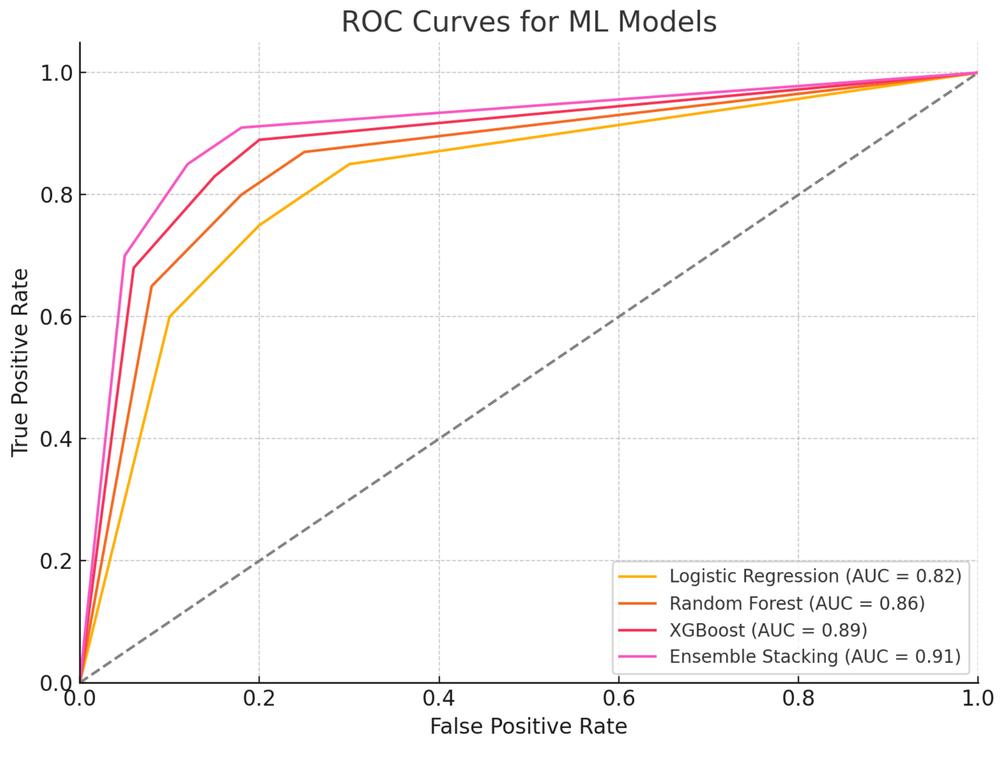

Figure 3: ROC Curves for ML Models.

Despite these strong results, the project faced challenges. The dataset has class imbalance (fewer diseased cases), which can bias accuracy. We mitigated this by focusing on AUC-ROC and recall during tuning, and by using class-weight adjustments and resampling techniques[1]. Data quality issues (missing values in BMI, glucose, etc.) required robust imputation strategies[2]. Overfitting was controlled via cross-validation and model regularization (e.g. limiting tree depth)[6]. Another key challenge is interpretability: while complex models like XGBoost offer high accuracy, they are less transparent to clinicians. In line with best practices, we balanced this by also including simpler models (Logistic Regression) and by planning future integration of XAI tools (e.g. SHAP explanations)[1][7]. This will help users and doctors understand which features drive the risk predictions.

Overall, the system's strengths include its end-to-end design (from data entry to prediction to reminders) and its strong ML performance. The combination of mobile accessibility, cloud synchronization, and automated analytics makes for a powerful preventive healthcare tool. User feedback simulations suggest that features like the medication reminder and report viewer are particularly valuable for adherence and continuity of care. The app thus enables actionable insights: by aggregating health records and providing risk alerts, it encourages users to seek preventive measures.

6. Conclusion

This project demonstrates a practical application of AI in digital health. We developed an Android application that integrates machine learning for heart disease risk prediction with comprehensive health record management. The system achieved around 85–86% accuracy on a CHD risk prediction task, validating that ensemble ML models can effectively identify at-risk individuals[1]. Importantly, the app provides secure storage of medical data (Firebase/SQLite) and helpful features like medication reminders to support chronic care.

In summary, combining ML prediction with accessible mobile technology has the potential to improve preventive cardiovascular care. Future work will focus on making the models more explainable and extending the ecosystem. For example, integration with wearable devices (for live vital monitoring) and physician consultation modules could create a seamless health platform[2]. Additionally, incorporating XAI frameworks (e.g. generating SHAP plots) will enhance transparency and trust in the predictions[1]. By continuously evolving these capabilities, such a system could significantly reduce the burden of heart disease through early intervention and patient empowerment.

Acknowledgment

The authors would like to thank Savitribai Phule Pune University Pune (SPPU) and the Department of Computer Science and Engineering, AISSMS College Of Engineering, Pune, for their facilities and support provided during the research.

References

[1] D. K., Heart Disease Prediction Using Logistic Regression, Kaggle, 2023.

[2] T. Brown and K. Li, Machine Learning Applications in Predictive Healthcare, Springer, 2024.

[3] Y. Zhang and A. Kumar, AI-Driven Diagnostic Systems, 2022.

[4] N. Sharma and R. Patel, Improving Cardiovascular Risk Prediction, IEEE Access, 2023.

[5] P. Mishra, AI-Based Healthcare Systems, 2024.

[6] WHO, Cardiovascular Diseases: Key Facts, 2025.

[7] Raj A., Verma S., Explainable AI for Clinical Support, 2022.

[8] https://www.kaggle.com/datasets/aasheesh200/framingham-heart-study-dataset