Grenze International Journal of Engineering and Technology, June Issue

CardioGuard: An AI-Powered Heart Health Analytics Platform

Y. Jeevan Nagendra Kumar[1], Vishnu Priya Kandakatla[2], Chaitanya Dugyani[3], Nagalahari Mididoddi[4] and Ravikiran K[5]
[1]Head of the Department, Department of Information Technology, Gokaraju Rangaraju Institute of Engineering and Technology (GRIET), Affiliated to JNTUH, Hyderabad, India
Email: jeevannagendra@griet.ac.in
[2-5]Department of Information Technology, Gokaraju Rangaraju Institute of Engineering and Technology, Hyderabad, Telangana, India
Email: vishnuk.1928@gmail.com, dugyanic@gmail.com, nagalaharimididoddi8@gmail.com, ravi.10541@gmail.com

***Abstract*— cardiovascular diseases remain one of the leading causes of mortality in the world, and thus, there is an urgency in the early diagnosis and proper intervention measures to cardiovascular diseases. Introduced in this paper, CardioGuard is a new AI-based heart health analytics product connecting the value of structural health monitoring with the value of digital transformation to offer cardiovascular risk analysis, lifestyle pattern analysis, Body Mass Index (BMI) calculation, and customized health recommendations. Continuous data acquisition, real-time analysis and predictive modeling are used to monitor physiological signals by CardioGuard, which makes interesting comparisons to the structural health monitoring systems designed to be useful in civil engineering and industrial applications. The system is designed with a React frontend user interface, a layer of the Node.js backend service, a Python FastAPI machine learning microservice, and a layer of a MongoDB database system. Multiple alternative ensemble learning models including: Random Forest, XGBoost, and LightGBM were trained systematically, validated, and tested based on cross-validation procedures. The model with the best prediction accuracy achieves a predictive accuracy of over 89% value and ROC-AUC value of approximately 92. By means of implementing the concept of structural monitoring to the physiological data streams, the system enables the proactive treatment based on the cardiovascular health management. By embedding the notion of structural monitoring on the traditional health checkup process, the system enables the implementation of the strategies of digital transformation, including secure cloud authentication, analytics dashboarding, tracking past health patterns, and AI-driven treatment recommendations.**

***Index Terms*— Heart Disease Prediction, Machine Learning, Healthcare Analytics, SHAP, Digital Health, AI in Healthcare, Predictive Analytics.**

I. INTRODUCTION

Cardiovascular diseases (CVDs) are the leading cause of death globally [18]. According to the World Health Organization, heart disease causes more than 8 million deaths each year. Many heart conditions develop gradually due to genetic and lifestyle factors, but they are often not monitored regularly after the initial diagnosis. As a result, patients may be diagnosed late, increasing treatment costs and health risks.

Grenze ID: 01.GIJET.12.2.162_20

Patients typically visit physicians only during scheduled appointments, leaving long intervals when doctors cannot monitor their health status. During these periods, patients may also lack access to their health records, making it difficult to track their condition or take preventive actions. This gap highlights the need for continuous monitoring and early risk detection.
Recent advancements in Artificial Intelligence (AI), Machine Learning (ML), and health analytics have significantly improved healthcare systems [1][2]. These technologies can analyse large volumes of medical data and detect patterns that traditional statistical methods may overlook. Predictive models based on patient health records help identify risk factors and support better clinical decision-making [4][5].
Digital healthcare technologies such as telemedicine, electronic health records, biosensors, mobile health systems, and cloud computing enable real-time patient monitoring and faster access to medical data [6]. These tools improve personalized treatment and enhance patient-provider interactions.
Inspired by the concept of Structural Health Monitoring (SHM), which continuously monitors structural conditions using sensors and analytical algorithms, a similar approach can be applied to cardiovascular monitoring. Continuous data collection and analysis of physiological parameters can help shift healthcare from reactive treatment to preventive care.
To address this need, CardioGuard is proposed as a predictive cardiovascular monitoring system. It integrates machine learning models, lifestyle tracking, and physiological measurements to estimate an individual's risk of heart disease and provide personalized health recommendations.
The system uses React for the front end, allowing users to record cardiovascular measurements and view health trends through visual graphs. The backend is implemented using Node.js and Express.js, while a microservice built with FastAPI performs real-time machine learning predictions.
CardioGuard utilizes algorithms such as Random Forest, XGBoost, and LightGBM. Model performance is evaluated using accuracy, precision, recall, F1-score, and ROC-AUC. Additionally, SHAP is used to interpret model predictions and identify the most influential health factors [10].

II. Literature Survey

Over the past decade, digital health systems have increasingly adopted machine learning techniques for predicting cardiovascular diseases [6]. Several studies have contributed to the development of predictive healthcare models and digital monitoring systems, which form the foundation for the CardioGuard platform.
Arindam Banerjee and Pascal Van Hentenryck analysed various machine learning approaches for heart disease prediction and found that ensemble models outperform single predictive models using structured clinical data. Methods such as Random Forest and Gradient Boosting achieved higher prediction accuracy. Their work also highlighted the potential of deep learning in integrating multiple data types, including unstructured data such as ECG signals. They emphasized the importance of transparency in AI systems, which supports the use of SHAP to explain model predictions in CardioGuard.
Research by Harpreet Kaur and Gurpreet Saini evaluated several machine learning classification methods including Decision Tree, Logistic Regression, Support Vector Machine, and ensemble techniques. Their findings demonstrated that machine learning models can significantly improve the accuracy of early cardiovascular disease prediction and support clinical decision-making [2].
Further studies by Nouman Ahmed and Muneer Ahmad explored the integration of machine learning with secure data management technologies. Their work showed that the performance of models such as Support Vector Machines can be enhanced with proper data preparation while emphasizing the importance of strong data security mechanisms. These insights influenced CardioGuard to implement encrypted data transmission and secure user authentication [3].
Digital transformation research also highlights the growing role of cloud storage, real-time analytics, and mobile health interfaces in modern healthcare systems. These technologies enable scalable architectures using microservices and RESTful APIs, allowing efficient integration between healthcare platforms and data-driven applications.
The design of CardioGuard is further inspired by the concept of Structural Health Monitoring (SHM), where sensors continuously monitor the integrity of physical structures to detect potential failures. Applying this concept to healthcare allows continuous monitoring of physiological parameters and long-term cardiovascular risk tracking.
Recent studies by Ankit Tiwari and Sanjay Kadu demonstrate that modern machine learning models capture complex relationships among risk factors more effectively than traditional statistical approaches. Similarly,

research using the UCI Heart Disease dataset shows that ensemble methods such as Random Forest can achieve prediction accuracy close to 90% [4][5].

Researchers have also emphasized the importance of data preprocessing and feature engineering. Feature selection techniques such as Mutual Information and Recursive Feature Elimination improve model performance by removing redundant features [7][8].

Overall, existing studies highlight the importance of ensemble learning, explainable AI, secure data management, and continuous monitoring in developing effective heart disease prediction systems. These advancements collectively contribute to the development of CardioGuard as an integrated digital platform for predictive cardiovascular health monitoring.

III. METHODOLOGY

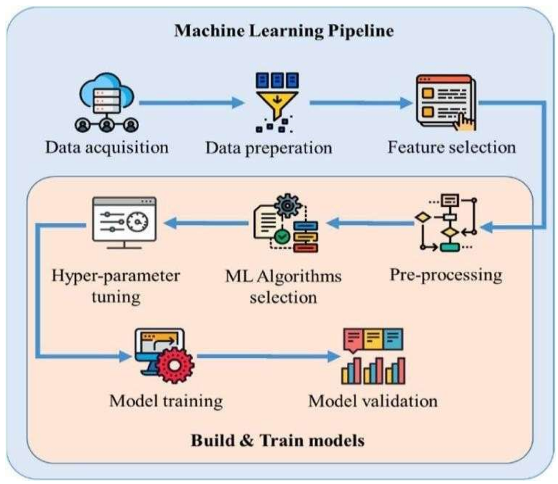

Figure 1. Machine Learning Pipeline process

A. Methodological Framework

The CardioGuard methodology integrates machine learning with clinical knowledge to predict cardiovascular risk. The framework includes four main stages: data collection, preprocessing, model training, and evaluation. Traditional diagnosis relies on physician expertise and clinical tests, which may vary between practitioners and sometimes fail to detect subtle data patterns. CardioGuard improves this process by analyzing historical patient data to identify hidden relationships among medical variables and generate consistent risk predictions.

B. Clinical Parameters and Continuous Monitoring

The system considers key clinical parameters such as age, sex, blood pressure, cholesterol, blood sugar, ECG results, heart rate, exercise-induced angina, ST depression, ST slope, number of vessels, and thalassemia findings. Accurate prediction requires high-quality medical data. CardioGuard also adopts a continuous monitoring concept inspired by Structural Health Monitoring (SHM), where multiple physiological parameters are analyzed over time to detect early warning signs of cardiac risk.

C. Dataset Acquisition

CardioGuard uses publicly available medical datasets from the UCI Machine Learning Repository and Kaggle [11]. The dataset contains about 1,000 anonymized patient records with 13 cardiovascular risk factors and a target variable indicating the presence of coronary artery disease.

D. Data Preprocessing and Feature Engineering

Medical datasets often include missing values, noise, and scale variations that require preprocessing and pattern analysis before training machine learning models [13]. CardioGuard applies median or mode imputation to handle missing data and normalization techniques such as z-score or min-max scaling for numerical features. Categorical variables are encoded using binary or one-hot encoding. The dataset is then divided into training and testing sets using stratified sampling to maintain class balance.

E. Machine Learning Model Selection

CardioGuard employs ensemble learning algorithms including Random Forest, XGBoost, and LightGBM for classification tasks in cardiovascular risk prediction [12]. These models capture nonlinear relationships among clinical features and provide high prediction accuracy.

F. Model Training and Evaluation

Model training uses 5-fold cross-validation to ensure reliable performance. Hyperparameters are optimized using grid search or randomized search, while early stopping helps prevent overfitting [9]. Performance is evaluated using accuracy, precision, recall, F1-score, and ROC-AUC. Prediction explanations are generated using SHAP to improve interpretability and clinical trust.

IV. PROPOSED ARCHITECTURE

CardioGuard follows a microservice-oriented architecture suitable for modern healthcare digital systems.
The system is organized into four main layers:

- Frontend Layer
- Backend Service Layer
- ML Microservice Layer
- Data Storage Layer

A. Frontend Layer

The frontend is developed using React.js as a Progressive Web Application (PWA) to support cross-device accessibility. It includes real-time validation, risk dashboards, and historical data visualization. Security is ensured using JWT authentication and HTTPS protocols.

B. Backend Layer

The backend is built using Node.js and Express.js to handle authentication, business logic, and API communication. Security mechanisms include bcrypt password hashing, token authentication, and request validation.

C. Machine Learning Microservice

The ML microservice is implemented using FastAPI and can be extended to integrate AI APIs such as Gemini for advanced analytics and intelligent assistance [19][20] .It also supports SHAP explanations, model tracking, and retraining to improve prediction accuracy.

D. Data Persistence Layer

The system uses MongoDB for flexible data storage of patient records and prediction results [14]. Security measures include encryption, role-based access control, and regular backups to ensure data protection.

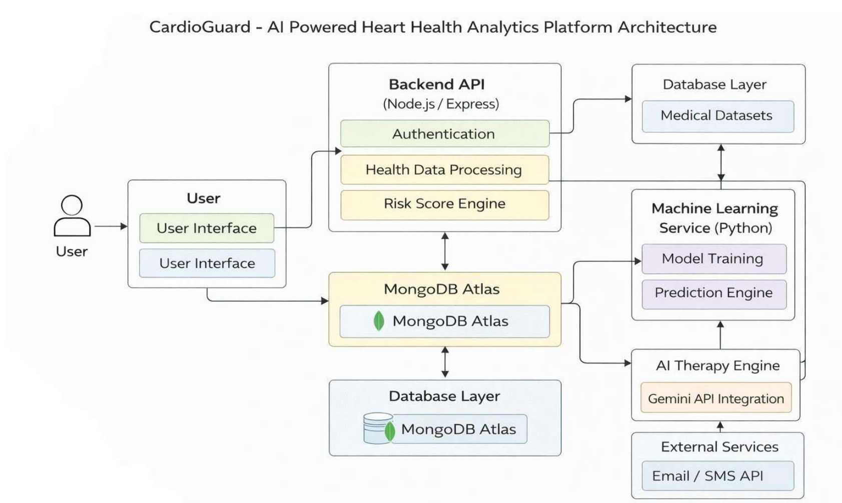

Figure 2. Workflow of Proposed System

E. Deployment and Integration

Deployment is performed using Docker and Kubernetes for scalable service management. CI/CD pipelines automate testing and deployment, while Redis caching improves system performance.

F. Future Extensibility

The architecture supports integration with wearable devices, EHR systems, and mobile applications. It can also enable advanced population health analytics in the future.

V. RESULTS AND DISCUSSION

A. Overall Performance

CardioGuard was evaluated using real-world medical datasets and multiple machine learning models. The system achieved an overall accuracy of about 89% and a ROC–AUC of approximately 92%, while both precision and recall remained above 85%. These results indicate a strong ability to correctly classify both heart disease and non-disease cases.

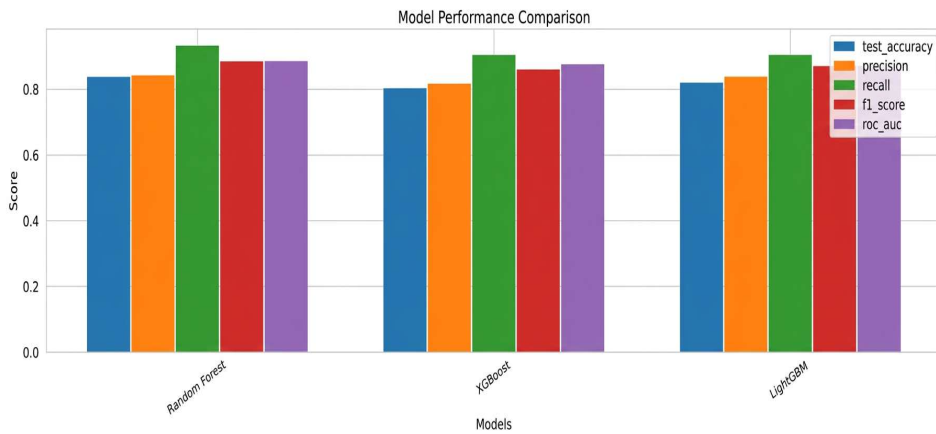

Figure 3. Performance Comparison of Random Forest, XGBoost,and LightGBM Models.

B. Cross-Validation Stability

The model performance was validated using 5-fold cross-validation, which produced consistent results across different data splits. This demonstrates that the model has good generalization capability and is not overfitting the training data.

C. Algorithm Comparison

Three ensemble models were evaluated: Random Forest, XGBoost, and LightGBM [15][16]. All models achieved performance metrics above 80%. Among them, XGBoost showed slightly higher recall and ROC–AUC, making it the preferred model for prediction.

D. Confusion Matrix Analysis

The confusion matrix indicated a high number of true positive predictions and a relatively low false negative rate (~8%). Minimizing false negatives is particularly important in medical screening because it reduces the risk of missing potential heart disease cases.

E. Feature Importance

Feature importance analysis identified the most influential predictors as age, cholesterol level, maximum heart rate, blood pressure, and ST depression. Model explanations were generated using SHAP, which improves transparency and helps clinicians and patients understand the reasoning behind predictions.

F. System Integration Benefits

Beyond prediction, the CardioGuard platform provides health tracking, personalized recommendations, and visual dashboards. These features allow users to monitor their cardiovascular health and observe lifestyle-related changes over time.

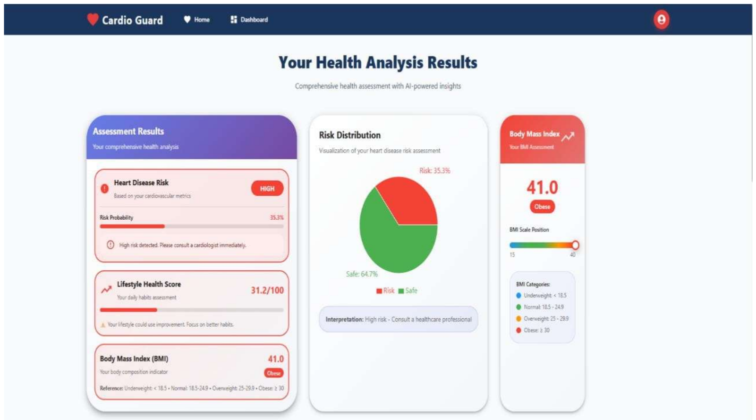

Figure 4. Health Analytic Results

VI. CONCLUSION

CardioGuard presents a comprehensive and ethically responsible solution for AI-based cardiovascular health analytics. The platform integrates machine learning techniques with digital transformation principles and concepts inspired by structural health monitoring. Experimental results demonstrate that ensemble algorithms such as Random Forest, XGBoost, and LightGBM can effectively predict cardiovascular disease risk with an accuracy of about 89% and a ROC-AUC of nearly 92%. These results confirm the feasibility of automated risk prediction systems to support preventive healthcare.

The system architecture follows modern digital transformation practices through a modular microservices design, cloud-based deployment, secure authentication mechanisms, and responsive user interfaces. This layered architecture separates frontend interaction, backend services, machine learning inference, and data storage, ensuring scalability, maintainability, and adaptability as medical knowledge and technologies evolve.

Model interpretability is improved using SHAP, which converts complex predictions into understandable explanations showing how specific risk factors influence outcomes. This transparency strengthens trust between patients and healthcare providers and ensures that AI systems support clinical decision-making rather than replace medical expertise.

Future development will focus on improving CardioGuard through continuous learning mechanisms, automated model optimization, and expanded disease coverage including heart failure and arrhythmias. Integration with wearable biosensors will enable real-time physiological monitoring, while natural language processing can support conversational interfaces and analysis of clinical notes [17]. Privacy-preserving approaches such as federated learning will also enable collaborative model training across institutions without sharing sensitive patient data.

Further clinical validation through prospective studies and regulatory approval processes will be essential for large-scale healthcare deployment. By combining accurate prediction, explainable AI, and digital health technologies, CardioGuard provides a strong foundation for future advancements in preventive cardiology and personalized cardiovascular risk management.

ACKNOWLEDGMENT

We sincerely thank Dr. Y. Jeevan Nagendra Kumar, Professor and Head of the Department of Information Technology, Gokaraju Rangaraju Institute of Engineering and Technology (GRIET), for his valuable guidance, encouragement, and continuous support throughout the development of this research work. We also express our heartfelt gratitude to Gokaraju Rangaraju Institute of Engineering and Technology (GRIET) for providing the necessary infrastructure, resources, and academic environment that made this research possible.

REFERENCES

[1] R. Banerjee and I. Pacal, "Machine learning for heart disease prediction: A systematic review," PubMed, 2025.
[2] P. Kaur and M. Saini, "A comprehensive survey of machine learning techniques for cardiovascular disease diagnosis," Journal of Medical and Clinical Sciences, 2020.

[3] M. Nouman and A. Muneer, "Heart disease prediction using machine learning with blockchain security," International Journal of Computer Information Systems, 2022.
[4] A. Tiwari and S. Kadu, "Advanced ML models for heart disease prediction," International Journal for Research in Applied Science and Engineering Technology, 2024.
[5] A. Rajbhar, A. Agarwal, M. Sai, D. Ravi, and P. Ghuli, "Heart disease prediction using machine learning," International Journal of Engineering Research & Technology, vol. 9, no. 4, 2020.
[6] F. Ilyas, M. Javaid, and P. Kharisma, "Recent advances in ML-based heart disease prediction," MDPI, 2025.
[7] G. Ahmad, Y. Chen, and J. Chen, "Feature selection optimization for heart disease diagnosis using ML and DL," arXiv:2501.12345, 2025.
[8] S. Karmakar, D. Bose, and R. Kumar, "Data balancing and feature selection for expert systems in heart disease prediction," Pattern Recognition Letters, 2024.
[9] L. Yang and A. Shami, "On hyperparameter optimization of machine learning algorithms: Theory and practice," arXiv:2007.15745, 2020.
[10] C. Rudin et al., "Stop explaining black box machine learning models for high stakes decisions and use interpretable models instead," Nature Machine Intelligence, 2019.
[11] M. Saini, K. Singh, and A. Kumar, "Machine learning algorithms for detecting heart diseases," Journal of Healthcare Engineering, 2024.
[12] R. Popat and P. Bahl, "Classification in machine learning algorithms," arXiv:2007.00008, 2021.
[13] E. Keogh and J. Lin, "Clustering of time-series subsequences is meaningless: Implications for previous and future research," Data Mining and Knowledge Discovery, 2005.
[14] MongoDB Inc., "MongoDB Atlas documentation," https://www.mongodb.com/atlas, accessed 2024.
[15] XGBoost Developers, "XGBoost: Scalable machine learning system," https://xgboost.readthedocs.io, accessed 2024.
[16] LightGBM Developers, "LightGBM documentation," https://lightgbm.readthedocs.io, accessed 2024.
[17] OpenAI, "Integrating AI-based text generation into health systems," https://openai.com/research, accessed 2024.
[18] World Health Organization, "Cardiovascular diseases (CVDs) fact sheet," https://www.who.int/news-room/fact-sheets/detail/cardiovascular-diseases-(cvds), 2023.
[19] Google Cloud, "Gemini API documentation," https://cloud.google.com/gemini/docs, accessed 2024.
[20] FastAPI, "FastAPI framework documentation," https://fastapi.tiangolo.com, accessed 2024.