Industrial Engineering Journal
ISSN: 0970-2555
Volume : 54, Issue 9, September : 2025

A MACHINE LEARNING FRAMEWORK FOR REAL-TIME ECG SIGNAL CLASSIFICATION AND HEART ANOMALY DETECTION

NEERATI SWATHI SREE, Student, Dept. of Computer Science & Engineering, MVR College of Engineering and Technology (A), Email: swathisree@example.com

Dr. D. Srinivas, Ph.D, Assistant Professor, Dept. of Computer Science & Engineering, MVR College of Engineering and Technology (A), Email: srinivas@example.com

Abstract—

Cardiovascular diseases continue to be a leading cause of mortality worldwide, frequently progressing undetected until critical events occur. Accurate detection of anomalies in Electrocardiogram (ECG) signals traditionally requires cardiology experts, making diagnosis both time-consuming and error-prone. Recent advancements in Machine Learning (ML) offer opportunities for automated ECG interpretation. This paper proposes a real-time ML-based ECG anomaly detection framework utilizing the XGBoost classification algorithm. The system is capable of processing ECG signals in under 30 milliseconds. Using a dataset of approximately 40,000 ECG recordings annotated by cardiologists across hospitals and countries, seven signal categories were classified: Normal, Atrial Fibrillation (AF), Tachycardia, Bradycardia, Arrhythmia, Other, and Noisy. The proposed XGBoost model achieved an accuracy of 97%, with F1-scores ranging between 0.93–0.99, outperforming classical ML baselines. This demonstrates the framework's scalability, robustness, and applicability across diverse clinical environments, particularly for real-time healthcare monitoring systems.

Index Terms— Machine Learning, Electrocardiogram (ECG), Heart Anomaly Detection, XGBoost, Real-Time Classification.

I. INTRODUCTION

Cardiovascular disorders, including atrial fibrillation, tachycardia, bradycardia, and arrhythmias, represent one of the most severe global health concerns, contributing to nearly 12 million deaths annually . Early detection remains critical, yet traditional ECG analysis relies heavily on human interpretation, leading to potential delays and diagnostic errors.

Recent progress in machine learning and artificial intelligence has enabled scalable approaches for analyzing ECG signals efficiently. Unlike static clinical evaluation, ML models can be trained on large datasets to detect underlying anomalies with high precision.

This study proposes a real-time ECG anomaly classification system that leverages the XGBoost algorithm, known for its superior performance in structured data modeling. Our approach processes ECG signals in less than 30 ms, making it suitable for wearable health devices and telemedicine applications.

II. RELATED WORK

Several researchers have explored ML techniques for cardiovascular prediction. Early studies applied data mining techniques to medical repositories , while hybrid ensemble classifiers (Decision Trees, KNN, and Naïve Bayes) demonstrated improved diagnostic accuracy . Association rule mining and decision trees for survival prediction have also been applied in cardiology.

Palaniappan et al. developed the Intelligent Heart Disease Prediction System (IHDPS) using Naïve Bayes, Decision Trees, and Neural Networks. Parthiban et al. designed a Co-Active Neuro-Fuzzy Inference System (CANFIS) integrating neural networks, fuzzy logic, and genetic algorithms. Vijayarani et al. reported near-perfect accuracy with neural networks for heart disease classification.

While these works confirm the potential of ML in cardiology, most relied on limited datasets with uniform recording standards. Our contribution focuses on real-time classification using a large, cross-hospital dataset of 40,000 ECGs, ensuring robustness across diverse clinical variations.

III. PROPOSED SYSTEM

The architecture of the proposed system consists of sequential modules for ECG signal acquisition, preprocessing, feature extraction, classification, and real-time prediction. Fig. 1 outlines the system workflow.

System Pipeline:

1. Input Data: Raw ECG signals gathered from hospitals and wearable devices.
2. Preprocessing: Noise reduction, QRS complex detection, P-wave/T-wave identification, normalization, and heart rate variability analysis.
3. Dataset Splitting: 80% for training, 20% for testing.

Industrial Engineering Journal
ISSN: 0970-2555
Volume : 54, Issue 9, September : 2025

4. Model Training: Using XGBoost as the primary model, compared against Decision Trees, Neural Networks, Naïve Bayes, and KNN.
5. Prediction: Real-time diagnosis into one of seven anomaly classes.

XGBoost's gradient boosting strategy enables efficient handling of high-dimensional ECG signals while maintaining computational efficiency. Hyperparameter tuning further ensured convergence within a 30 ms inference time.

IV. IMPLEMENTATION

The framework was built in Python 3.7 on a Windows 8 platform with minimum requirements of an Intel i3 processor, 4GB RAM, and a 40GB hard disk. Core libraries:

- XGBoost: Model training and optimization.
- Scikit-learn: Baseline model implementation.
- NumPy/SciPy: Signal processing and numerical feature extraction.

Modules:

- Data Preprocessing (noise filtering and feature extraction).
- Model Training and Evaluation.
- Real-Time Web Interface for prediction.

V. RESULTS

A. Performance Comparison

The model was evaluated using precision, recall, accuracy, and F1-score metrics. Table I summarizes the comparative results.

The classification performance of five algorithms—XGBoost, Decision Tree, Neural Network, Naïve Bayes, and KNN—was evaluated on ECG anomaly detection. XGBoost achieved the best results with approximately 97% accuracy, 97% precision, 99% recall, and 97% F1 score. Its gradient boosting approach effectively handles complex ECG features and noise.

Decision Tree and Neural Network models showed moderate performance with accuracies near 90% and 92%, respectively. Naïve Bayes and KNN performed lower, around 85% and 80% accuracy, due to their limitations in modeling complex dependencies and sensitivity to noise.

This confirms that ensemble models like XGBoost are most suited for real-time ECG classification tasks, providing robust and accurate detection of cardiac anomalies.

XGBoost outperformed all baseline models across all evaluation metrics, confirming its suitability for real-time classification.

B. Test Data Validation

A secondary validation was performed on 12 unseen ECG samples (80-dimensional feature vectors). The system produced correct classifications: Normal (2), Tachycardia (1), Bradycardia (1), Arrhythmia (2), Other (4), and Noisy (2). These results emphasize the robustness of the system in practical deployment scenarios.

VI. CONCLUSION

This study presented a real-time ECG anomaly classification model using XGBoost, achieving an overall accuracy of 97%, with inference times under 30 milliseconds. Compared to conventional algorithms, the proposed framework delivers superior predictive performance and scalability across heterogenous ECG datasets.

Such a system can be integrated into wearable monitoring devices and telemedicine applications, significantly improving accessibility to cardiac diagnosis. Future extensions will focus on additional anomaly classes, deep learning-based hybrid models, and deployment over IoT-enabled medical devices.

REFERENCES

1. World Health Organization, "Cardiovascular Diseases," 2025.
2. S. Tsumoto, “Data Mining in Clinical Databases,” 2000.
3. Y. A. Aslandogan et al., “Combining Classifiers for Heart Disease Prediction,” 2004.
4. C. Ordonez, “Association Rule Mining for Heart Disease Prediction,” 2004.
5. F. Le Duff, “Decision Trees for Survival Prediction in Cardiology,” 2004.
6. S. Palaniappan et al., “Intelligent Heart Disease Prediction System,” 2008.
7. L. Parthiban et al., “CANFIS for Heart Disease Prediction,” 2008.
8. S. Vijayarani et al., “Classification Approaches for Heart Disease Prediction,” 2013.

9. P. Groves et al., “The Big Data Revolution in Healthcare,” 2016.
10. M. Chen et al., “Big Data: A Survey,” Mobile Netw. Appl., vol. 19, no. 2, pp. 171–209, 2014.