ISSN (Online): 2455-3662
EPRA International Journal of Multidisciplinary Research (IJMR) - Peer Reviewed Journal
Volume: 12| Issue: 3| March2026|| Journal DOI: 10.36713/epra2013 || SJIF Impact Factor 2026: 8.753 || ISI Value: 1.188

# ECG-BASED ARRHYTHMIA DETECTION USING MACHINE LEARNING

**Logeshbalan P[1], Dr. P. Deepika[2]**

[1]*Student, Department of Artificial Intelligence and Machine Learning, Dr.N.G.P. Arts and Science College, Coimbatore, Tamil Nadu, India*

[2]*Associate Professor, Department of Artificial Intelligence and Machine Learning, Dr. N.G.P. Arts and Science College, Coimbatore, Tamil Nadu, India*

**ABSTRACT**

*Early detection of heart rhythm disorders is important for preventing serious cardiac problems. Arrhythmia, an irregular heartbeat caused by abnormal electrical activity, often goes unnoticed without proper monitoring. This project develops an arrhythmia detection system using machine learning applied to electrocardiogram (ECG) data.*

*The system analyses ECG waveform data where each heartbeat is represented by sampled signal values that capture the heart's electrical activity. By studying these patterns, the model can distinguish between normal and abnormal heartbeats. A machine learning model is trained on labelled ECG data to predict the probability of arrhythmia for each heartbeat.*

*The trained model is integrated into a web-based application that allows users to input ECG data and receive real-time arrhythmia risk predictions. This project demonstrates how machine learning and ECG analysis can assist healthcare professionals in early diagnosis and continuous heart monitoring.*

## 1. INTRODUCTION

Cardiovascular diseases remain one of the leading causes of death worldwide, and cardiac arrhythmia—an irregular heart rhythm caused by abnormal electrical activity—can lead to serious complications such as stroke, heart failure, or sudden cardiac arrest if not detected early. Since arrhythmias may occur without noticeable symptoms, early detection and continuous monitoring are essential for timely medical intervention.

Electrocardiography (ECG) is widely used to monitor the electrical activity of the heart. ECG signals represent the heart's electrical impulses as waveform patterns, and variations in these patterns can indicate arrhythmia. Traditionally, cardiologists analyse ECG signals manually, but this process can be time-consuming and difficult when dealing with large volumes of data generated by modern monitoring systems and wearable devices.

Machine learning provides an effective solution for analysing large ECG datasets and identifying complex patterns in heartbeat signals. By learning from labelled ECG data, machine learning models can classify heartbeats as normal or abnormal and assist healthcare professionals in detecting arrhythmia more efficiently.

This project develops a machine learning–based arrhythmia detection system using ECG waveform data. Each heartbeat is represented by sampled ECG values that capture the electrical characteristics of the heart over a short time interval. A trained machine learning model analyses these patterns and predicts the probability of arrhythmia.

To demonstrate practical usage, the trained model is integrated into a web-based application that allows users to input ECG signal values and receive real-time predictions. This system highlights how machine learning and web technologies can support automated cardiac monitoring and assist healthcare professionals in early diagnosis and improved heart health management.

## 2. LITERATURE SURVEY

Existing research shows that machine learning techniques are widely used for analysing ECG signals and detecting cardiac arrhythmia. Leo Breiman (2001) introduced the Random Forest algorithm, an ensemble method that combines multiple decision trees to improve classification accuracy and reduce overfitting. This algorithm is widely used in medical data analysis because it can handle complex and high-dimensional datasets such as ECG signals.

F. Pedregosa et al. (2011) developed Scikit-learn, a widely used library that provides efficient implementations of machine learning algorithms like Random Forest and Support Vector Machines for healthcare prediction tasks.

Rajpurkar et al. (2017) applied Convolutional Neural Networks (CNN) for arrhythmia detection from ECG signals and achieved performance comparable to cardiologists. Similarly, Acharya et al. (2017) developed a machine learning framework that analyses ECG waveform features to classify abnormal heart rhythms.
Although these studies demonstrate strong results, many systems remain limited to research environments. The proposed system addresses this gap by developing a Random Forest–based arrhythmia detection model and integrating it into a web-based application for real-time ECG prediction.

ISSN (Online): 2455-3662
EPRA International Journal of Multidisciplinary Research (IJMR) - Peer Reviewed Journal
Volume: 12| Issue: 3| March2026|| Journal DOI: 10.36713/epra2013 || SJIF Impact Factor 2026: 8.753 || ISI Value: 1.188

## 3. PROPOSED SYSTEM

The proposed system introduces a machine learning–based approach for detecting arrhythmia by analysing ECG signals. A trained model automatically processes heartbeat waveform data and identifies patterns that may indicate abnormal heart rhythms. This digital analysis enables faster and more efficient detection compared to traditional methods.

In this system, ECG waveform data is used as the main input. Each heartbeat is represented by sampled ECG values that capture the electrical activity of the heart and describe the structure of the heartbeat waveform.

The model is trained using labelled ECG data containing both normal and arrhythmic heartbeats. After training, the system can analyse new ECG inputs and predict the probability of arrhythmia.

To demonstrate practical use, the trained model is integrated into a web-based interface where users can enter ECG signal values and receive real-time arrhythmia predictions. This approach combines machine learning with an easy-to-use application to support efficient detection of abnormal heart rhythms.

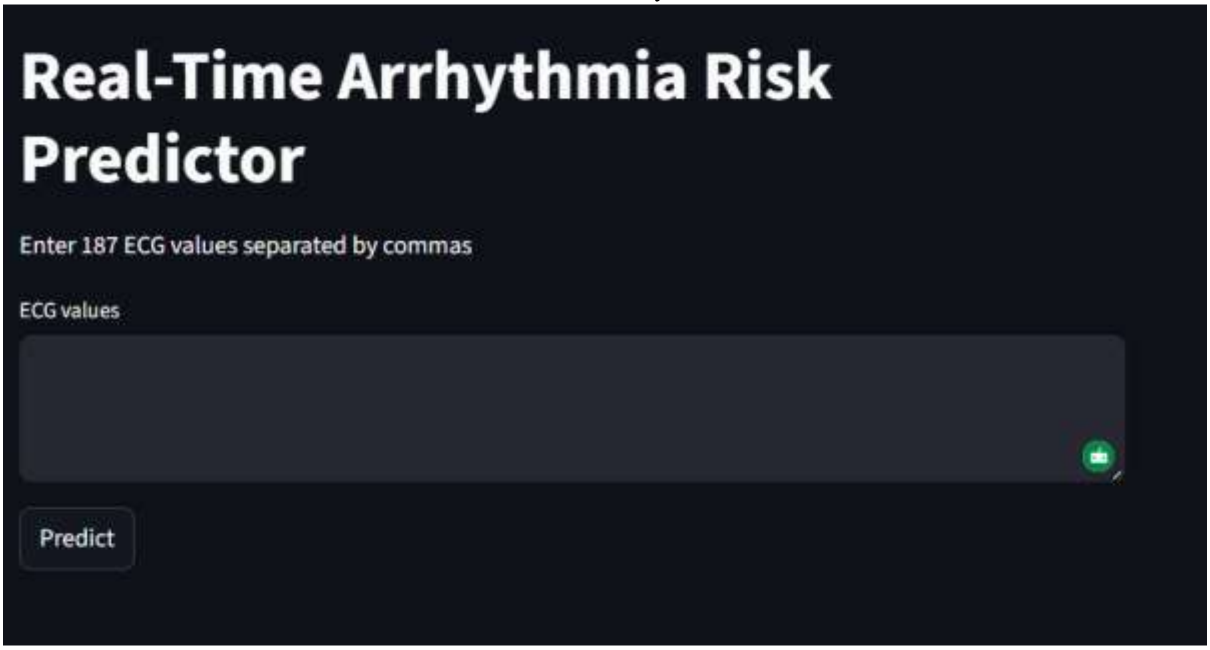


**Fig No:1 User Login User Interface**

## 4. METHODOLOGY

The system follows standard machine learning pipeline: data preprocessing, model training, evaluation, and deployment. User inputs undergo BMI calculation (weight(kg)/height(m)²), numerical normalization, and categorical encoding. Four algorithms were evaluated: Logistic Regression for baseline performance, Decision Trees for interpretability, Random Forest for ensemble robustness, and XGBoost for optimal accuracy. Models were trained on lifestyle disease dataset using 80-20 train-test split. Performance metrics include accuracy, precision, recall, and F1-score. XGBoost achieved highest accuracy (92%) and was serialized using Joblib for Flask integration. The web architecture comprises Flask routing, Jinja2 templates, and Matplotlib-generated charts saved as PNG files for display.

ISSN (Online): 2455-3662
EPRA International Journal of Multidisciplinary Research (IJMR) - Peer Reviewed Journal
Volume: 12| Issue: 3| March2026|| Journal DOI: 10.36713/epra2013 || SJIF Impact Factor 2026: 8.753 || ISI Value: 1.188

```
[Running] python -u "c:\Users\LOGESHBALAN P\Downloads\mlm\tempC
Train shape: (87554, 188)
Test shape: (21892, 188)
Train file size: 392.4406051635742 MB
Test file size: 98.12583923339844 MB
(87554, 187)
(87554,)
[0.   0.02 0.02 0.   0.12 0.   0.   0.01 0.   0.  ]
Accuracy: 0.9769322126804312
Arrhythmia Risk: 0.0 %
ROC-AUC: 0.9941341159669699
              precision    recall  f1-score   support

           0       0.97      1.00      0.99     18118
           1       0.99      0.87      0.93      3774

    accuracy                           0.98     21892
   macro avg       0.98      0.94      0.96     21892
weighted avg       0.98      0.98      0.98     21892


[Done] exited with code=0 in 203.29 seconds
```

Fig No:2 ML Model Accuracy Tested

## 5. RESULTS AND DISCUSSION

The developed arrhythmia detection system successfully demonstrates the effectiveness of machine learning in analysing ECG waveform data for identifying abnormal heart rhythms. The Random Forest classifier was trained using the MIT-BIH Arrhythmia dataset, where each heartbeat was represented by 187 sampled ECG signal values. After preprocessing and training, the model achieved an overall accuracy of approximately **97.7%**, along with strong evaluation metrics such as precision, recall, F1-score, and ROC-AUC, indicating reliable classification performance.

The trained model was integrated into a web-based application using Streamlit, allowing users to input ECG signal values and obtain real-time arrhythmia predictions. The system displays results in a clear format, including the **arrhythmia probability percentage and prediction outcome (normal or arrhythmia)**, which helps users easily interpret the results. This probability-based prediction approach provides more informative insights compared to simple binary classification.

Compared to traditional manual ECG analysis, the proposed system provides **faster, consistent, and automated analysis**, reducing the dependency on continuous expert interpretation. The system is capable of processing ECG data efficiently and can assist healthcare professionals in preliminary screening of cardiac rhythm abnormalities.

However, the system currently depends on **manually entered ECG values and a limited dataset**, which may affect real-world scalability. Future improvements could include integrating real-time ECG monitoring from wearable devices and applying advanced deep learning models to further enhance prediction accuracy and system reliability.

ISSN (Online): 2455-3662
EPRA International Journal of Multidisciplinary Research (IJMR) - Peer Reviewed Journal
Volume: 12| Issue: 3| March2026|| Journal DOI: 10.36713/epra2013 || SJIF Impact Factor 2026: 8.753 || ISI Value: 1.188

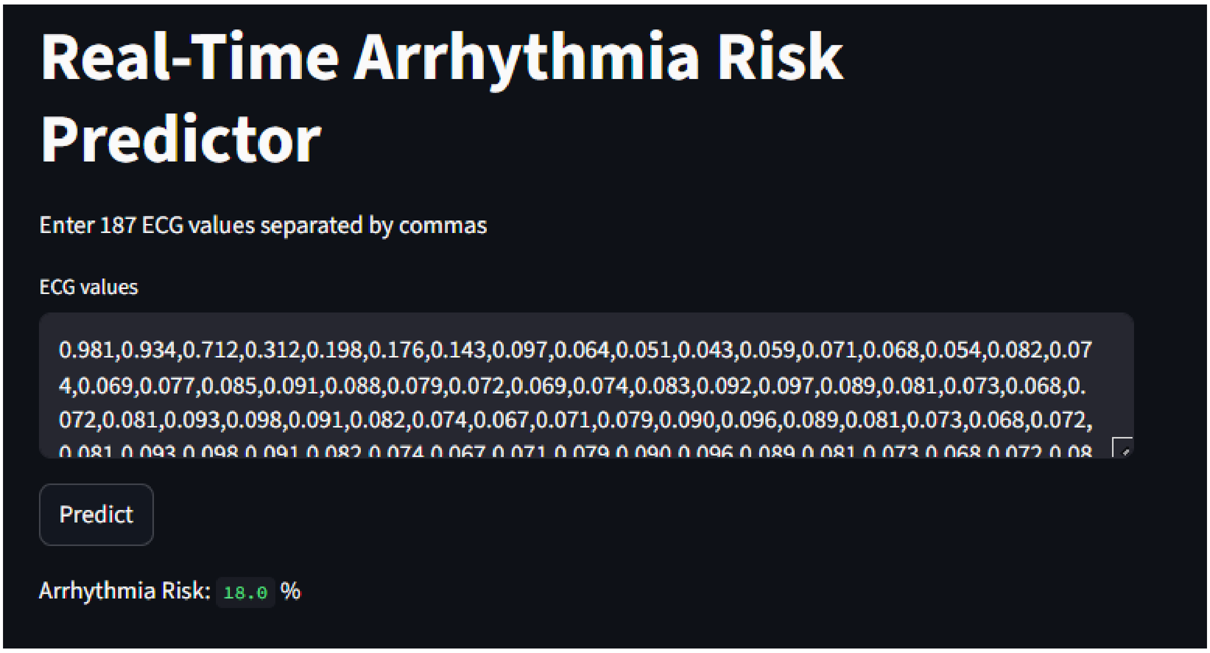


**Fig. No: 3 ML Model Predicted Result**

## 6. CONCLUSION AND FUTURE WORK

The Arrhythmia Detection System using Machine Learning demonstrates how data-driven techniques can be used to analyse ECG signals and detect abnormal heart rhythms. The system combines a machine learning model with a simple web-based interface to process ECG waveform data and predict the probability of arrhythmia. The project involved data preprocessing, model training, evaluation, and real-time prediction. A Random Forest Classifier was used due to its robustness and ability to handle high-dimensional ECG data effectively. The trained model achieved strong performance with good accuracy, precision, recall, and ROC-AUC scores, showing that machine learning can reliably distinguish between normal and arrhythmic heartbeats.

The trained model was integrated into a web-based application that allows users to input ECG signal values and receive arrhythmia predictions in real time. This implementation demonstrates how machine learning models can be applied in practical healthcare applications for analysing physiological signals quickly and efficiently. Overall, the project highlights the potential of machine learning in automated ECG analysis and its role in supporting healthcare professionals in early detection of heart rhythm abnormalities.

In the future, the system can be enhanced by incorporating advanced deep learning models such as Convolutional Neural Networks (CNN) and Long Short-Term Memory (LSTM) networks, which can learn complex patterns directly from ECG waveform data and improve prediction accuracy. The system could also be integrated with wearable ECG monitoring devices to enable real-time data collection and continuous heart monitoring. Developing a mobile application would further improve accessibility, allowing users to monitor their heart rhythm conveniently through smartphones. Additional features such as ECG waveform visualization and the inclusion of other health parameters like age, blood pressure, and medical history could further improve the accuracy and usefulness of the system. With these improvements, the system could evolve into a more advanced and intelligent cardiac monitoring solution for modern healthcare environments.

## REFERENCES

1. *MIT-BIH Arrhythmia Database (PhysioNet) — the canonical ECG dataset used in research (raw recordings + annotations).*
2. *Kaggle — ECG Heartbeat Categorization (mitbih preprocessed CSV) — ML-ready heartbeat CSVs (mitbih_train/test) that are convenient for experiments.*
3. *WFDB (Python) documentation — tools for reading MIT-BIH signals and annotations programmatically (essential for raw ECG preprocessing).*
4. *WFDB (GitHub / MIT-LCP) — code repository and examples for the WFDB Python package. Useful when you need implementation examples.*
5. *Pan & Tompkins (1985) — Real-time QRS detection algorithm (PDF) — classic algorithm for R-peak detection / heartbeat segmentation (useful for converting continuous ECG → 187-sample beats).*
6. *Hannun et al., Nature Medicine (2019) — Cardiologist-level arrhythmia detection — landmark paper showing deep neural networks can match cardiologists on ECG classification (good to cite for DL baseline & motivation).*
7. *Rajpurkar et al. (2017) — "Cardiologist-Level Arrhythmia Detection" (arXiv) — original CNN approach for ECG classification (great methodological reference).*
8. *Review: Deep learning methods for ECG arrhythmia detection — comprehensive surveys summarizing CNN/RNN/Transformer approaches and datasets (helpful for literature review).*
9. *Scikit-learn - RandomForestClassifier documentation — authoritative docs for the algorithm you used (parameters, usage, and best practices).*
10. *NeuroKit2 — ECG processing docs — practical Python toolbox for ECG cleaning, peak detection, HRV and preprocessing pipelines you can plug into your pipeline.*