VOL. 21, NO. 6, MARCH 2026 ISSN 1819-6608
ARPN Journal of Engineering and Applied Sciences

www.arpnjournals.com

IOT-ASSISTED CARDIOVASCULAR HEALTH MONITORING WITH ML-BASED RISK ANALYSIS

R. Uma Maheswari[1], P. Ashok Kumar[2], B. Hema Latha[1], Rajitha Bodasingi[1], A. Chaitanya Lakshmi[1], Dadi Lohith Sankar[1] and A. Harsha Vardhan Reddy[1]
[1]Department of Electronics and Communication Engineering, Vignan's Institute of Information Technology (A), Visakhapatnam, India
[2]Department of Electronics and Communication Engineering, Vignan's Institute of Information Technology for women, Visakhapatnam, India
E-Mail: maheswariu.ramisetty@gmail.com

ABSTRACT

Cardiovascular disease is now a major issue for each and everyone. Cardiovascular disease is the biggest reason for losing lives. This shows how important it is for us to detect and predict the disease early and monitor our health continuously. So, in this paper, we mainly focus on this issue and come up with a solution, which is an IoT-based cardiovascular disease prediction using Machine Learning. This continuously monitors the condition of the heart in real-time, which includes sensors, machine learning models,a web application, and cloud computing all together. The main controller used is Raspberry Pi; it collects the body readings like blood pressure, oxygen levels, Heart rate, and the ECG signals using the sensors. Later, the data is cleaned up and given to the machine learning model, which is trained using the Cleveland heart disease dataset. This helps in classifying the person as high-risk or not with good accuracy. The IoT connectivity is added so that all the sensors' live readings are sent to the cloud platform, and can give alerts. This model is cost-effective and is used in continuous monitoring at home or in emergency cases.

Keywords: IoT, Cardiovascular disease, machine learning, real-time monitoring.

Manuscript Received 18 February 2026; Revised 21 March 2026; Published 20 May 2026

1. INTRODUCTION

Cardiovascular disease is one of the most serious and dangerous health issues all over the world. Many of the reports are saying that heart problems are the major cause of death every year. These include many parameters of the heart, like heartbeat, rhythm, and blood flow, so it really affects how the heart works overall. In the past few years, machine learning (ML) techniques have become more popular in the medical field as they help in faster diagnosis, and predicts the risk earlier, and make some clinical decisions without manual work.

Machine learning can read the huge amount of data and find the patterns which are not noticed that easily, this help in more accuracy and earlier detection of the heart condition and prevention. Data mining is also important in gathering information from large datasets so that it improves how diseases are identified and managed efficiently in hospitals. In addition to this, machine learning, sensor technology, and IoT cloud computing have been added, which continuously monitor in real-time and have become easier. The sensors are connected to the Raspberry Pi and integrated with the IoT cloud computing, which collects and processes the data so that everything can be seen remotely and can be monitored at all times without depending on the older manual checking methods.

The system uses real-time sensors and IoT to check the cardiovascular risk of the person. When all these are included together,they make the model a simple, portable, and useful solution for early finding and managing heart-related issues. It does not need too much manual work and is easier to use for people. Overall, the main aim of this paper is to make the health-monitoring system easierso that it can be done without any expert help. It can be used in places where the medical equipment is not available, such as in rural areas or in small clinics. It makes it more convenient for both patients and doctors. IoT-assisted cardiovascular monitoring with ML-based risk classifications provides a more accessible, continuous, and intelligent way to track heart health, making early detection easier and improving the patient's outcome. Here, we are using the main controller called Raspberry Pi to load the data and process it.

Some of the usual symptoms of heart disease that most of the people commonly experience are shown Figure-1.

VOL. 21, NO. 6, MARCH 2026 ISSN 1819-6608
ARPN Journal of Engineering and Applied Sciences

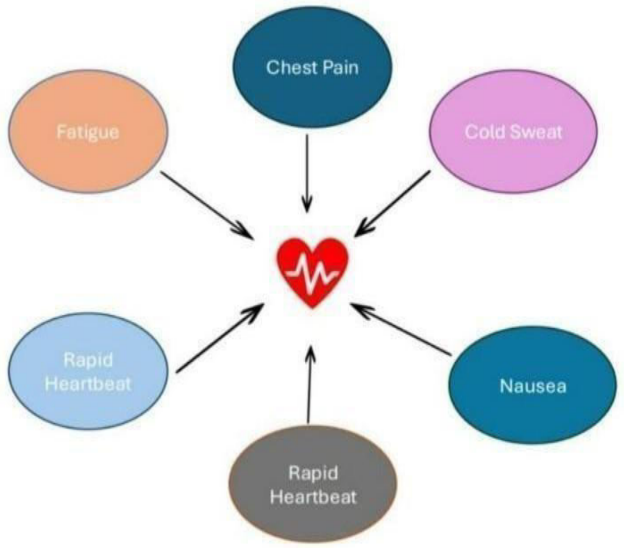

Figure-1. Symptoms of heart disease.

2. BACKGROUND STUDY

Over the years, a lot of research has been done for the improvement of the early detection of heart disease, which has been monitored regularly using different machine learning models and IoT-based sensors. Many of the researchers have mainly focused on the various classification methods, collecting the patients' data in real time and applying various strategies to get more reliable results. The overall results not only show how cardiovascular diseases are becoming more serious but also indicate how much technology has improved in supporting the automatic diagnosis. One study that is published in the year 2019 compared logistic regression with Support Vector Machine for the prediction of heart disease, there were reported with result of 73% and 80% accuracy. However, these models produced false outputs, which means that the patients who actually had heart diseases were classified as normal. This issue was mainly caused by the limited dataset and the lack of features, which have made the models difficult to perform in the real- time scenarios. These drawbacks pushed the researchers to explore more robust methods and especially the assembling techniques. Another research combined the Support Vector Machine with the live sensors, which are collected by the Raspberry Pi setup, to predict the heart condition. This model uses both past values and current values, such as pulse rate, blood pressure, and weight. This process has been more automated and responsive. Even so, SVM struggles when the patient's health condition changes quickly. SVM couldn't adapt to those changes. So, researchers realized that they needed models that are more flexible and can also be adjusted as the patient's condition changes in real time. Another research studied and compared several ML models, such as K-Nearest Neighbors, Random Forest, AdaBoost, Decision Trees, and Logistic Regression. Among these studies, KNN and Random Forest gave the best results. The neural study performed well, but its accuracy dropped when the dataset was imbalanced. These results show that the dataset needs to be balanced and that proper features are needed for the proper tuning of the work. Another research highlighted that the ensemble model, mainly the Random Forest technique, is preferred, which gives us higher accuracy and provides a clear explanation of which features matter more for most of the heart disease datasets. Random Forest is taken up because it can handle complexity, avoid overfitting, and can also find the non-linear patterns in the data better than any other model. However, these models require a lot of computational power, which makes them harder to use in small or low-power devices. Recent work has combined machine learning with IoT devices, which have been used in the real-time monitoring of the system. Even though there are many systems that are useful, theystill have the same problems. They do not have smooth connectivity to the cloud, cannot support multiple users at the same time, and do not have the proper data. Because of these issues, there is a need for small IoT systems that collect real-time data and make the predictions quickly. This affordable IoT-based system collects real-time health data and uses machine learning models to predict the risk of heart disease efficiently. Especially, this is designed for home care and can be useful in rural areas.

3. SYSTEM ARCHITECTURE AND DESIGN

IoT sensors continuously collect vital health parameters from the patient and send them to the system. The collected data is processed and passed to a Random Forest Machine Learning Model, which predicts whether the patient is at risk of a heart attack.

IoT Data Acquisition

IoT sensors are used to measure patient health parameters. The collected parameters match the 7 features used in the ML model. Sensor readings are captured in real time. The data is transmitted from the IoT sensors to the backend system.

Data Handling and Preprocessing

Incoming data is formatted to match the trained dataset structure. It includes Label encoding for categorical features. Feature scaling using a saved scalar. Preprocessing ensures consistency between training and prediction data.

Machine Learning Model Implementation

A Random Forest classifier is used for heart attack risk analysis. The model is training on a 7 features dataset. The model components are:

- Trained Random Forest model
- Feature Scalar
- Label encoders
- Prediction Threshold

Model Calibration and Thresholding

The model prediction threshold is calibrated. A predefined threshold is used to decide the normal condition and the high-risk condition. This improves the reliability and also reduces the false predictions.

Flask Backend Integration

Here, we used the Flask interface to connect the ML with the user interface. It handles receiving health input data -->Passing data to the ML Model --> Returning prediction results finally.

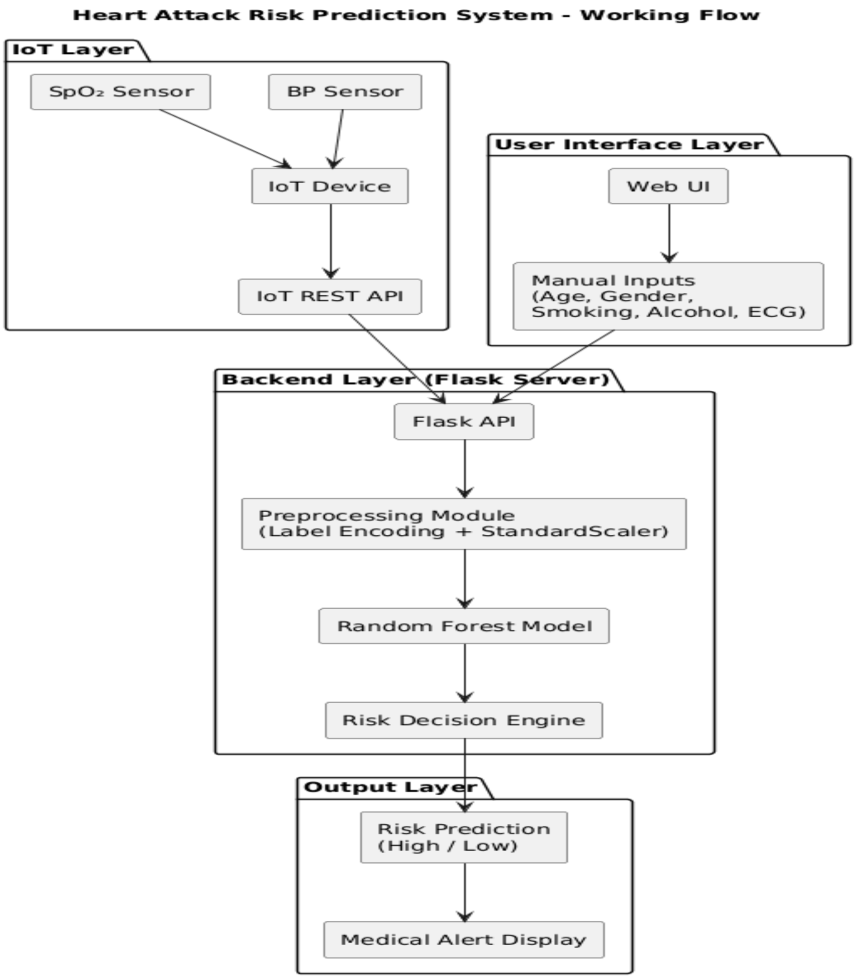

Figure-2. Working Flow.

Web Interface and Display

The Flask web page allows users to Enter IoT sensor data and submit data for prediction. And the display shows Heart attack risk results and predictions. It provides a simple, user-friendly visualization of predicted results.

System Flow

- Sensors collect patient health data
- Data is sent to the backend system
- Preprocessing is also applied
- ML model predicts the heart condition
- Flask displays the result on the web interface

4. METHODOLOGY

Data Acquisition

The system collects all the data from the parameters like heart rate, blood pressure, SpO2 and ECG through the IoT sensors connected to the Raspberry Pi. It is the primary stage for measuring the risk classification.

PrepProcessing

The data which acquired contains noise, and it may contain missing values. The attributes such as age, cholesterol, resting, Blood pressure, and some are the selected features. In the preprocessing, it cleans the data by removing and filtering.

A. Normalization: It takes values from 0 to 1: It is x (norm) = x - x (min) / x (max) - x (min)

B. Outlier Removal using IQR: Interquartile range is calculated as: IQR= Q3 - Q1

X<Q1 - 1.5. IQR or x > Q3 + 1.5. IQR

ARPN Journal of Engineering and Applied Sciences

Feature Extraction

It is an important stage in the cardiovascular risk prediction; the performance of features affects the performance of the overall classification performance. In the system, the features are derived from the sensors and the dataset-based clinical data describing the low-risk and high-risk.

A. Heart rate: The pulse sensor measures the number of heartbeats per minute (bpm). The number of n betas counted in the interval of t seconds is calculated as: HR= 60(N) /t generally, high or low HR variations indicate arrhythmias and cardiac stress.

B. Blood Pressure (BP): The BP measures the systolic and diastolic pressures by cuff- based sensing. This feature helps to detect hypertension and hypotension. Mean arterial pressure is also calculated along with this. MAP=DBP+((SBP- DBP) /3)

C. ECG Feature Derivation: The signals of the ECG detect the activity of the heart. From the QRS complex waveform, the R-R interval is calculated using the peak-to- peak detection. The HRV (heart rate variability) is calculated as: HRV=Var (R-R intervals).

D. Clinical Categorical Features: These are the chest pain(CP) and the number of major vessels (CA), etc.

Machine Learning-based classification: The Random Forest algorithm is used to train the model and predict the risk. It reduces overfitting. Each tree predicts the output and displays the final class according to the majority voting. The trainer model classifies output as Risk= {0, Low Risk 1, High Risk}

Cloud Integration and Monitoring: Sensor values and predicted values are transmitted to the cloud platform using MQTT/HTTP protocols. The cloud backend stores the data it generates the alerts, and provides the dashboard for real-time monitoring.

Visualization and Performance metrics: Clinicians and users can view the results and the risk prediction through a mobile app or website by continuous and remote cardiac health assessment. To validate the classifier, standard performance metrics are used: Accuracy, Precision, Recall (sensitivity), and F1-Score.These metrics are calculated using the outcome of the confusion matrix.

Accuracy = TP+TN / TP + TN + FP +FN
Precision = TP / TP + FP
Recall = TP / TP + FN
F1-Score = 2 x Precision x Recall / Precision + Recall.

A. Data Normalization standard scalar:

Z= (X-\mu)/(\sigma)

Where X = original feature value

\mu=Mean of feature

\sigma=Standard deviation

B. Decision Tree Splitting:
Gini = 1- (\sum_{i=1}^{c} p_i^2)

Where p_i = probability of class (i)

c= number of classes

C. Information Gain:
IG = Gini(parent) - \sum(N_{child}/N_{parent})Gini(child)

D. Random Forest Prediction:
Each tree predicts the class --> T1(x), T2(x), T3(x),…..Tn(x)
Then final Prediction --> y^ = Mode {T1(x), T2(x), T3(x),…..Tn(x)}

METHODS AND MATERIALS

The original dataset includes a total of 26 columns

Our target: heart attack likelihood.

No	Feature	Type	Source	Description
1.	Age	Numerical	Manual	Patient age
2.	Gender	Categorical	Manual	Male/Female
3.	Smoking Status	Ordinal	Manual	Never/Regularly
4.	Alcohol Consumption	Ordinal	Manual	Never/Regularly
5.	ECG results	Binary	Manual	Normal/Abnormal
6.	Blood oxygen levels	Numerical	IoT Sensor	Oxygen Saturation
7.	BP Systolic	Numerical	IoT Sensor	concentration
8.	BP Diastolic	Numerical	IoT Sensor	relaxation

ARPN Journal of Engineering and Applied Sciences

Target Variable

Target	Type	Meaning
Heart Attack Likelihood	Binary	Yes=High Risk, No=Low Risk

5. EXPERIMENTAL RESULTS

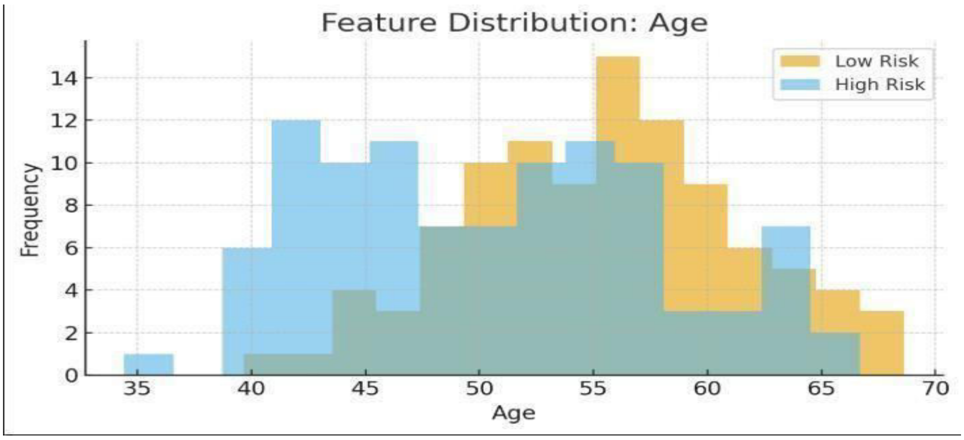

Figure-3. Feature distribution graph (Age vs Risk Group).

The above Figure-3 shows the different ages of people having Heart Disease according to High Risk and Low Risk. The people having High Risk around the age of 52-55. These people belong to the middle age range. The low-risk people belong to 48-65years. This age cannot decide the person's risk level. It also considers some factors such as BP, cholesterol, ECG readings, and Heart rate to determine the Heart Condition.

The overlapping region shows that the people having same age belong to either High Risk or Low Risk. So, it can lead to wrong conclusions. Because of that, Machine Learning is needed. It will also work when both the regions overlap.

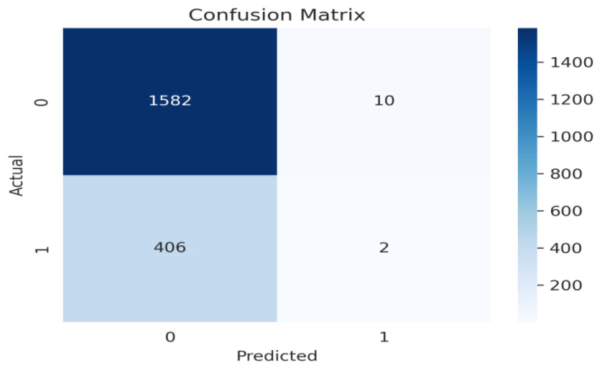

Figure-4. Confusion matrix.

The confusion matrix shows the performance of the ML using the Random Forest model for predicting the Heart condition.

a) True Positives (TP=21): Patients with heart disease are identified.

b) True Negatives (TN=31): It indicates the patients who are at low risk.

c) False Positives (FP=1): Patient with the high risk are mistakenly identified.

d) False Negatives (FN=7): The patients who are at high risk are identified as low risk. This is the most dangerous case.

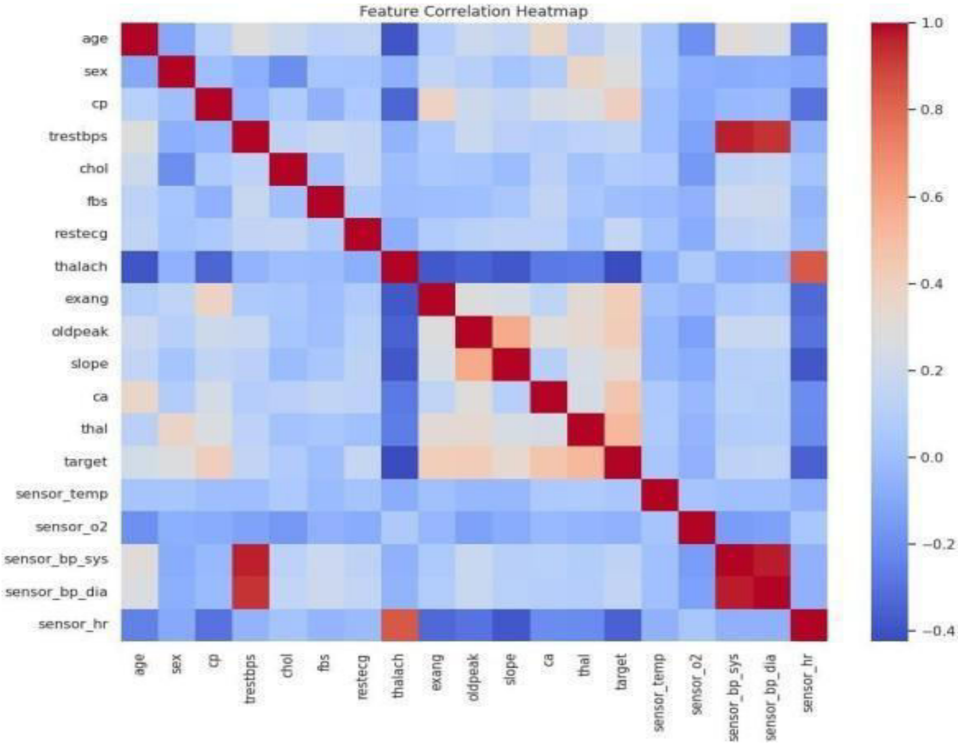

Figure-5. Correlation Heatmap.

The Correlation Heat map shows the different features related to heart disease. The IoT sensors show mild to moderate correlations according to the target. These sensors' readings provide useful information. The Heat maps are helps to identify the features which effect the targets and show how the clinical and sensor data combined improve the model accuracy and the prediction performance.

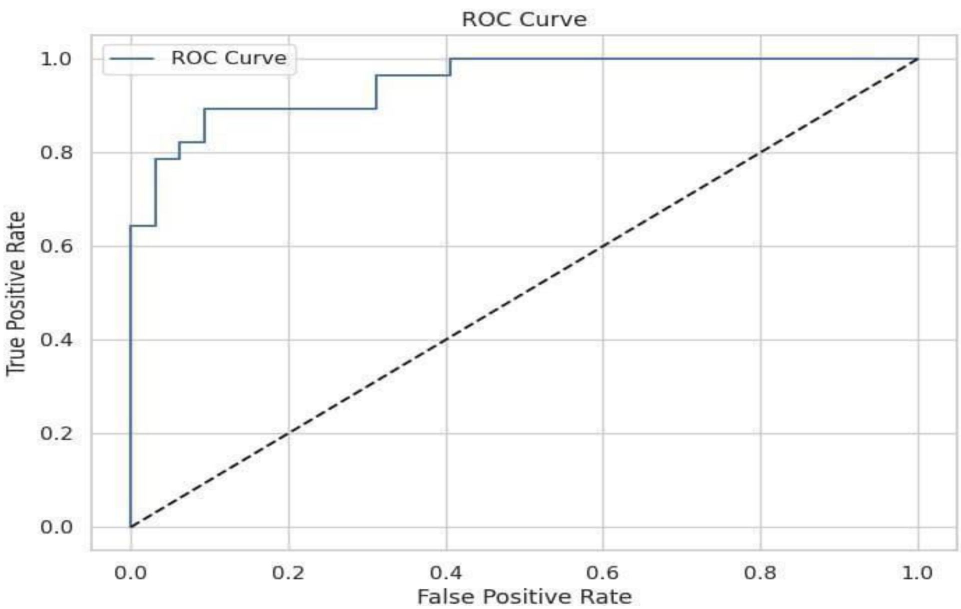

Figure-6. ROC Curve.

The ROC curve in Figure-6 shows how the model can compare between the patients who are having the heart disease and who are not having the heart disease. The curves rise sharply towards the top-left corner, which means the system has achieved a high true-positive rate even when the false- positive rate is low. The dashed diagonal line shows the random guessing, and the curve staying well above it represents that the model performs better. The ROC curve indicates that line classifier is reliable and effective in separating the low-risk and high risk which is suitable for heart disease prediction.Here, the KDE graph in Figure-7 shows the feature age, but the system produces all other features like blood oxygen levels, BP diastolic, and BP systolic.

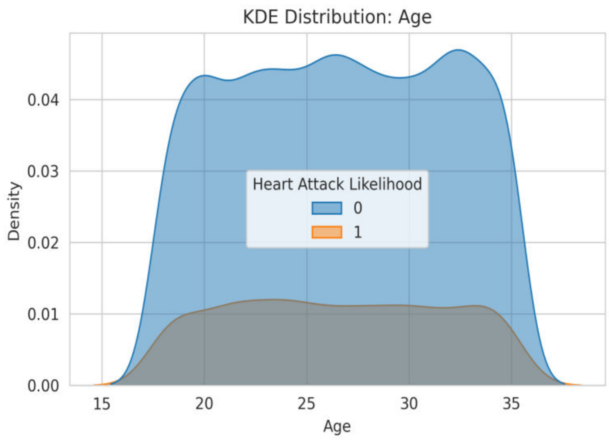

Figure-7. Kernel density estimation graph.

6. CONCLUSIONS

The IoT-assisted health monitoring and heart disease prediction using machine learning is used to detect the heart condition early. The sensors, such as temperature, blood pressure, and heart rate, are connected to the Adriano, and it measures all the parameters. Now, the collected data is processed using the Random Forest algorithm trained on the Cleveland Heart Diseases Dataset. The model gets 95% accuracy, which is better than the other algorithms like SVM, Neural Networks, and Regression models. Using IoT connectivity, it is easy to predict the heart condition of patients in real time. Mostly, it is helpful for doctors to monitor patients' conditions anywhere without visiting them. The models analyze the data and detect the conditions. If incase of High risk is identified if send an alert immediately. It compares the data with the Cleveland dataset and predicts the outputs. This is useful for healthcare, rural areas, and emergencies.

REFERENCES

T. Hossain *et al*. 2022. ECG signal analysis using machine learning approaches. Biomedical Signal Processing and Control. 72: 1-13.

R. Kumar and M. Singh. 2021. An IoT-based device for remote cardiac monitoring.IEEE Transactions on Consumer Electronics. 67(3): 210-218.

R. S. Anand and M. Kumar. 2021. Wearable health monitoring devices: A review.Journal of Medical Systems. 45: 1-15.

Z. Wang *et al*. 2021. Early cardiovascular risk prediction using ML and cloud computing. IEEE Cloud Computing. 8(2): 45-53.

S. Patel *et al*. 2022. A review of wearable sensors and systems with application in rehabilitation. Journal of Neuro Engineering and Rehabilitation. 19: 1-17.

G. F. T. da Silva *et al*. 2022. Real-time health monitoring using IoT and predictive analytics. IEEE Latin America Transactions. 20(3): 560-567.

H. Ali and M. Imran. 2021. Machine-learning-based early detection of arrhythmia using ECG signals. Sensors. 21: 1-18.

K. P. Singh and R. Sharma. 2022. Cloud-assisted IoT framework for predictive healthcare analytics. Future Generation Computer Systems. 129: 75-88.

2025. Heart Disease Cleveland Dataset, Kaggle. Accessed:

M. P. La Valley. 2008. Logistic regression. Circulation. 117(18): 2395-2399.

Mohan S., Thirumalai C., and Srivastava G. Effective heart disease prediction using hybrid machine learning techniques. IEEE Access. 7, 81542-81554.

Long N. C., Meesad P. and Unger H. 2025. A highly accurate firefly-based algorithm for heart disease prediction. Expert Syst. Appl.

Ogundepo E. A. and Yahya, W. B. 2023. Performance analysis of supervised classification models on heart disease prediction. Innov. Syst. Softw. Eng. 19(1): 129-144.

Paul B. and Karn B. 2023. Heart disease prediction using scaled conjugate gradient backpropagation of an artificial neural network. Soft Comput. 27(10): 6687-6702.

R. Ferdousi, M. A. Hossain, A. El Saddik. 2021. Early-stage risk prediction of non-communicable disease using machine learning in health cps. IEEE Access. 9, 96823-96837.

R. Sivaprasad, M. Hema. 2022. Heart disease using a machine learning and transfer learning model in the 2022 International Conference on Automation, Computing, and Renewable Systems(ICACRS), IEEE.

S. Mondal, R. Maity. 2024. An efficient computational risk prediction model of heart disease based on dual- stage stacked machine learning approaches, IEEE Access.

Y. Cao, F. Yang, Q. Tang and X. Lu. 2019. An attention-enhanced bidirectional LSTM for early forest fire smoke recognition. IEEE Access.