Forecasting of Heart Attack syndrome using Logistic Regression Algorithm with Blynk App Assimilation

G.Vijaybaskar[1], R.Pugazendi[2],

1Research Scholar, Department of Computer Science, Government Arts College (Autonomous), Salem-7, Affiliated to Periyar University, Salem-636011,Tamil Nadu. India.

2Assistant Professor, Department of Computer Science, Government Arts College (Autonomous), Salem-7, Affiliated to Periyar University,Salem-636011,Tamil Nadu. India

Abstract:

Heart and stroke diseases continue to be a predominant cause of mortality globally, where heart attacks are a big problem to public wellness. Early predictions & timely interventions in cases of a heart attack can reduce the fatality rate and enhance patient outcomes.This research proposes an Internet of Things systemreal-time heart attack prediction system using the logistic regression algorithm, which classifies the risk of a patient by considering age, cholesterol levels, BP, and some electrocardiographic measures. The dataset is preprocessed to make it reliable for study purposes. Logistic regression was selected because of its interpretability andsuitable for health applications. In order to make the predictive model more accessible and user-friendly, the same has been integrated with the Blynk IoT mobile application, which can visualize patient health data seamlessly collected from sensors and microcontrollers. The system architecture combines IoT-based data acquisition with cloud-based analytics, enabling real-time alerts and risk assessment on the user's smartphone. The study outcomes show that the suggested model provides a high level of precision inpredicting heart attack risk while maintaining transparency in decision-making, which is essential for clinical adoption. Integration of ML with IoT platforms underlines the potentials of intelligent healthcare systems in supporting preventive care, empowering patients, and assisting medical professionals in early diagnosis.

Keyword: *Heart attack, Machine learning, Logistic regression, Blynk app*

1. Introduction

Cardiovascular illnesses consistently rank as the leading cause of mortality worldwide. Cardiovascular diseases and strokes account for nearly 1.7×10^7 deaths worldwide every year, as reported by the WHO. Over 75% of fatalities attributed to cardiac disorders mostly occur in nations with middle- and low-income demographics. Moreover, myocardial infarctions and cerebrovascular accidents responsible for nearly eighty percent for all deaths attributable to Heart and stroke diseases [13]. Developing and underdeveloped nations lack infrastructure, technology, and medical professionals to detect diseases at an early stage, hence preventing complications and reducing death rates. The advancement of information and telecommunications technology has benefitted both affluent and impoverished patients by delivering real-time information at a reduced cost for diagnosis and health monitoring. This has significantly enhanced the detail of patients' health data. The extensive medical records are accessible for research purposes. The healthcare sector has significant obstacles in using vast medical data. A substantial volume of data is rapidly converted into useful and precise information by machines. Consequently, machine learning is a significant domain. The extremely effective machine learning methods used to identify concealed patterns and correlations among characteristics in the dataset[4]. The medical dataset exhibits inconsistencies and redundancies; hence, effective preprocessing is a pivotal step [5].

The diagnosis of cardiovascular disease is frequently contingent upon the assessment of the recipient's conditions and a comprehensive health check up. Cardiovascular disease risk factors encompass tobacco consumption, advancing age, familial history of heart disease, elevated cholesterol levels, sedentary behaviour, hypertension, obesity, diabetes, and psychological stress [6]. Modifications to lifestyle, including cessation of tobacco use, weight reduction, increased physical activity, and stress management, may mitigate certain risk factors. Heart disease is diagnosed using medicinal history, physical examination, and imaging techniques including electrocardiograms, echocardiograms, cardiac MRIs, and laboratory tests. Lifestyle modifications, pharmacological treatments, and surgical interventions such as angioplasty, coronary artery by pass grafting, or the implantation of devices like pacemakers or defibrillators may alleviate heart disease. The advancement of contemporary healthcare systems has facilitated the creation of predictive models for cardiovascular disorders through the extensive availability of clinical data. Machine learning is regarded as a data-analysis methodology that examines extensive datasets from multiple perspectives and subsequently converts the findings into actionable insights.

The analysis of cardiovascular syndrome often depends on the evaluation of the patients conditions and a thorough health examination. Risk factors for cardiovascular disease include tobacco use, increasing age, family history of heart disease, high cholesterol levels, physical inactivity, hypertension, obesity, diabetes, and psychological stress [6]. Alterations to lifestyle, such as the discontinuation of tobacco consumption, weight loss, enhanced physical activity, and stress regulation, may alleviate certain risk factors. Heart disease is diagnosed using medical history, physical examination, and imaging modalities such as electrocardiograms, echocardiograms, cardiac MRIs, and laboratory assessments. Lifestyle adjustments, pharmaceutical therapies, and surgical procedures such as angioplasty, coronary arteries bypass grafting, or the installation of devices like pacemakers or defibrillators may mitigate heart disease. The progress of modern healthcare systems has enabled the development of predictive models for cardiovascular diseases due to the abundant availability of clinical data. Machine learning is considered a data-analysis technique that analyses large datasets from several angles and later transforms the results into usable insights.

Motivation

This study is motivated by the increasing need to enhance reliability andeffectiveness of heart attack prediction by using new technologies like machine learning and the IoT. Cardiovascular disorders, especially myocardial infarctions, continue to be a most important source of Global ill health and mortality significantly burden healthcare systems. Timely identification of heart attack risk is essential for averting serious consequences and enhancing patient survival rates. Nonetheless, precise diagnosis continues to pose a significant difficulty owing to the complex nature of the illness and the heterogeneity in individual patient traits and symptoms. Machine learning (ML) has demonstrated remarkable potential in healthcare applications by analyzing complex medical data, identifying hidden patterns, and generating reliable predictions. Among these, logistic regression stands out as an interpretable and computationally efficient algorithm capable of binary classification, making it suitable for heart attack risk assessment. Despite this, many existing ML models operate in offline environments and lack real-time implementation or user interaction. This research is motivated by the need to bridge that gap by integrating a logistic regression–based heart attack prediction model with the Blynk IoT mobile application, enabling real-time data collection, risk prediction, and alert generation. By leveraging ML techniques alongside IoT-enabled monitoring, this study aims to create an intelligent, readily available and simple to use system that assists clinicians and individuals in early diagnosis, timely intervention, and continuous heart health management.

2. Review of Literature

Saryadi et al., (2025) [18]formulated a machine learning-driven prediction model to aid clinicians in identifying and forecasting heart attacks with enhanced accuracy. Data from patients at an Iranian hospital were gathered and pre-processed, including the management of missing values, normalization, and the elimination of outliers. Analysis conducted using Rapid Miner software found seven key predictors of heart attack from an original set of thirteen parameters, with a high accuracy rate of 97.86%, as corroborated by hospital doctors. Among several machine learning methods, the decision tree had the best accuracy (97%) and offered interpretable decision routes for clinical application. The choice ruled derived from the decision tree identified chest discomfort, trait anxiety, age, smoking, and resting electrocardiographic findings as the most significant criteria. Fuzzy clustering was used to corroborate these results, demonstrating a 76% alignment rate for the two most significant factors—chest pain and trait anxiety.

Ahmed et al., (2025) [19]presented an enhanced machine learning methodology using Gradient Boosting Machine (GBM) and Adaptive Inertia Weight Particle Swarm Optimization (AIW-PSO) to forecast heart failure survival rates. A total of 299 individuals who were diagnosed with heart failure were included in the dataset, which was obtained via Kaggle. The dataset contains clinical information include age, ejection fraction, and serum creatinine levels.In comparison to more conventional ML models, the improved performance of the enhanced GBM model was shown by the achievement of a test accuracy of 94%.

Stonier et al., (2024) [20]provided an overview to predict the probability of a heart attack, using several machine learning approaches and techniques. Within the scope of this research, a multitude of techniques, such as RF, regression models, K-NN imputation (KNN), and the NB algorithm, were tested with the aim of strengthening medical diagnosis. The RF algorithm demonstrates an effective accuracy of 88.52 when used to predict heart attack probability and this can transform the diagnostic and treatment of cardiovascular diseases.

Hasan et al., (2024)[21]discovered key factors connected with heart failure and use various machine learning techniques to forecast its incidence, thereby facilitating early assessment of mortality rates related to heart failure. A comprehensive model construction was carried out by them using a dataset including information about heart failure. When applied to this dataset, the model shown outstanding performance, with an veracity of 85.23% when RF was used across the full data-set and 86.36% when Scalable Discriminate Evaluationwas used to the XGBoost dataset.

Khan et al., (2023)[22]employed the machine learning (ML) method to carry out a key explanation for accurate prediction and decision making for CVD patients. Several different situations were used to assess the efficiency of the recommended ML algorithm in order to find out which approach within the model class was the most suitable. When it came to cardiovascular disease (CVD), the RF technique had the highest values for predicting accuracy, sensitivity, and receiver operating characteristic curves, with numbers of 85.01%, 92.11%, and 87.73%, respectively. As far as cardiovascular disease (CVD) was concerned, it had the lowest specificity and misclassification errors, with a respective percentage of 43.48% and 8.70%.

Bhatt et al., (2023)[23]presented a k-modes clustering algorithm with Huang initialization to enhance classification accuracy. This methodology utilizes various models, including random forest (RF), decision tree (DT), multi-layer perceptron (MP), and extreme gradient boosting (XGB). GridSearchCV was employed to adjust the model's parameters for improved overall performance. The study findings demonstrate that the multi-layer perceptron utilizing cross-validation surpasses all alternative methods regarding accuracy. The maximum accuracy achieved was 87.28%.

Hasan et al., (2022) [24]examined coronary heart disease is a leading cause of mortality worldwide. Predicting cardiac disease is among the most formidable challenges in clinical data analysis. Diagnostic assistance may greatly benefit from machine learning (ML) because it makes it easier to make decisions and predictions using data supplied by the healthcare sector worldwide. They have encountered the implementation of ML techniques in the medical field for disease prediction. Numerous research has shown that a machine learning classifier is an effective instrument for predicting heart illness. This study employed eleven machine learning classifiers to improve heart disease prediction by identifying essential factors. The predictive model was developed utilizing several feature combinations and established classification techniques. Utilizing multilayer perceptrons and gradient boosted trees, they achieved a 95% accuracy rate in the model forecasting cardiac complications. The Random Forest model exhibits a 96% accuracy rate, rendering it the most effective for predicting heart disease.

Kaur et al., (2022) [25]aimed to forecast the probability of coronary heart disease in patients by the use of an enhanced machine learning algorithm. The input data undergoes several steps, including preprocessing, aggregation, and the identification of pertinent features prior to classification. Four algorithms—RF, K-means, GA, and LR—are used to identify heart disease. This approach eliminates extraneous features from the cardiac dataset to enhance performance and reduce training duration. This method is executed using the random forest approach. The evolutionary method optimizes K-means clusters to aggregate all outlier data points. Finally, logistic regression is used to categorize individuals according to heart disease. The performance of several current methodologies has been examined based on certain performance metrics. The computed accuracy rose to 95%.

Garg et al., (2021) [26] examined that ML, a leading application of AI, is achieving remarkable advancements in research. This study used to ascertain the presence of heart disease in individuals. heart disease (CVD) affects a large population and has the potential to kill people all over the world. A person's age, cholesterol levels, chest pain, and other symptoms may be able to be used by machine learning to detect the existence of cardiovascular disease. Supervised learning classification techniques could be a better fit for problems with diagnosing heart disease. The Random Forest approach achieved an accuracy of 81.967 percent, while the K-NN method achieved 86.385 percent.

Bharti et al., (2021) [27]examined an accurate prognosis of cardiac disease may avert life-threatening situations, whereas erroneous predictions may be deadly. To achieve the aim of this work, multiple deep learning methods and ML solutions are applied to Analyse and compare the outcomes of the UCI ML Heart Disease dataset. Accuracy and confusion matrix analysis help to obtain some promising results and confirm them.To achieve improved performance, dataset is normalized and Isolation Forest performs to eliminate any irrelevant attributes. The use of multimedia tools such as mobile devices in the research is also discussed. The precision was 94.2, which was achieved through a deep learning method.

Reddy et al., (2021) [28]According to research, cardiovascular diseases (CVDs) result in about 20.5 million deaths annually. However, forecasting ahead of time helps individuals to adjust their life choices and provides access to necessary medical assistance when needed. This study integrates the complete attribute set from the Cleveland heart dataset with the most favourable attributes subset chosen by 3 evaluators. The optimized subsets are utilized to train 10 distinct machine learning classifiers, including Bayesian, function based, slothful, meta, rule based, and tree based models, which effectively forecast the probability of getting heart disease. The effectiveness of these algorithms was evaluated by a ten fold cross validation method. The researchers ultimately improved the hyper parameter 'k,' which denotes the number of nearest neighbours in the instance based (IBk) classifier. A chi-squared(x^2) attribute evaluator determined the optimal feature set, achieving a maximum accuracy of 86.468%, in contrast to 85.148% when utilizing all features for SMO. Upon setting the hyper parameter ‘k’ to 9 using the chi-squared feature set, IBk demonstrated an 8.25% enhancement in prediction accuracy; still, the SMO classifier emerged as the most efficacious predictive technique.

Padmaja et al., (2021) [29]cardiovascular disease has become a leading health concern affecting many millions of people globally. Timely diagnosis of the disease can result in many lives saved. This study developed a predictive model for predicting cardiovascular disease based upon a ML model. The model was developed using a classification model, which is critical for predicting. The RF Model uses a pragmatic methodology to identify cardiac abnormalities and could be used in the healthcare industry to drastically change the practice of cardiology. With the introduction of technology into the healthcare sector, automation of tasks using machine learning methodologies has become possible. These machine learning methodologies can identify and/or predict early-stage cardiac disease.

Rani et al., (2021)[30]The hybrid decision support system introduced by these authors aims to provide a tool for assisting with the early detection of heart disease by analyzing a patient's clinical variables. The authors employed the Multivariate Imputation by Chained Equations approach to estimate the absent features or variables from the raw input data. The authors devised a distinctive Hybridized Feature Selection Method, integrating Genetic Algorithms and Recursive Feature Elimination techniques, to identify the most significant characteristics from the dataset for this system. During the Data Preprocessing phase, the authors employed the Synthetic Minority Over Sampling Technique alongside several conventional scaling methods. Ultimately, the authors employed five distinct AdaBoost classifiers and algorithms, including RF, SVM, Naive Bayes Classifier, and LR, during the final development phase of the proposed hybrid system to achieve optimal classifier accuracy. The RF Classifier enabled the system to have the highest accuracy in classifying heart illness to date. The experimental assessment of the proposed hybrid system was conducted in a simulated environment developed with Python, utilizing the Cleveland Heart Disease dataset from the UCI Machine Learning Repository. The hybrid model attained an overall accuracy of 86.6%, surpassing numerous previous Heart Disease Prediction Systems documented in the literature.

Yaswanth et al., (2020) [31]determined that the healthcare sector produces a significant amount of unstructured data that is unintelligible to machines. With the progression of modern technology, healthcare organizations are structuring data in a style that is intelligible to machine learning systems. The utilization of machine learning algorithms in this context enables the early detection of cardiac problems, so encouraging patients to seek improved treatment options for the condition. This study utilized seven different supervised learning algorithms to forecast heart illness. The algorithms included K-Nearest Neighbours, Decision Trees, Naive Bayes, LR, RF, SVM, and Neural Networks. This study produced performance metrics such as veracity, precision, recall, score of F1, and ROC values to assess the system's prediction veracity. The findings of this study demonstrate that, among the seven algorithms, Neural Networks attained the best accuracy of 92.30%. The findings provide empirical evidence that supports the model's effectiveness in predicting heart disease.

Obasi et al., (2019) [32]cardiovascular diseases are deadly yet stealthy conditions that annually increase the mortality rate among affected persons. In 2016, WHO estimated that 17.9 million deaths per year worldwide are due to heart disease. In the healthcare dept., substantial volumes of data are generated daily, encompassing all types of information, and deriving insights from this data is crucial. The suggested solution employs existing methodologies, including Random Forest Bayesian Classification and Logistic Regression. It is a decision support system that allows medical professionals to recognize and predict cardiac illnesses and heart attacks in individuals based on associated risk factors. The methodology was executed on the RStudio platform to assess the probability of patients developing cardiac disease. Analysis of the comparative data revealed that the system's performance and accuracy are satisfactory. The Random Forest achieved a classification accuracy of 92.44% for heart disease, while the Naïve Bayes Classifier and LR recorded veracity rates of 61.96% and 59.7%.

Mohan et al., (2019) [33]cardiovascular disease is a predominant cause of mortality globally at present. A major challenge in clinical data analysis is forecasting cardiovascular disease. Machine learning (ML) has proven effective in enhancing decision-making and generating predictions by leveraging the extensive data generated by the healthcare sector. Recent breakthroughs in several businesses related to the Internet of Things (IoT) have showcased the implementation of machine learning methodologies. Many research utilizing machine learning techniques provide just a superficial understanding of cardiac illness prediction. This research study introduces an innovative method to improve the precision of cardiovascular disease prediction by identifying critical characteristics using machine learning techniques. The predictive model employs several feature combinations and a spectrum of established classification techniques. The heart disease prediction model based on the hybrid random forest with a linear model (HRFLM) demonstrates improved performance, with an accuracy of 88.7%.

3. Problem statement

Cardiovascular disease is seriousglobal wellness challenge, with risk factors like hypertension, obesity, and sedentary lifestyle leading to severe consequences. In spite of the fact that machine learning has shown some promise to predict cardiac disease, most of the existing models are dependent on offline datasets and hence are devoid of real-time monitoring, with most of the algorithms having complex mathematical expressions that deter clinical adoption. The absence of IoT-based integration and mobile applications further restricts accessibility and timely intervention. To address these gaps in an efficient manner, this study proposes a lightweight and interpretable logistic regression-based predictive model integrated with the Blynk IoT mobile application. This allows for real time data visualization, risk prediction, and generation of alerts. Patient data from sensors or manually input will be fed to the model, and the output will be sent to the Blynk Cloud. Results will be displayed as interactive widgets on the mobile app. In the case of high risk-the probability is greater than 0.7-the app triggers alerts and notifications to the caregivers, enhancing preventive healthcare with timely intervention.

4. Research Methodology

The research methodology undertakes the structured design of a real-time heart disease prediction system by integrating ML with IoT-based monitoring. A clinical dataset collected is preprocessed for normalization and encoding, ensuring preparedness of the model. A number of algorithms are tested, and Logistic Regression is considered owing to its accuracy in binary classification with interpretability. The model trained has been deployed on a smartphone app to enable real-time risk prediction. ESP32/Arduino hardware interlinks biological sensors like ECG, heart rate, and blood pressure, while patient data is sent to the cloud for processing. Prediction and sensor readings visualization has been designed using the Blynk mobile app, while a display monitor and a buzzer avail immediate alerts. Accuracy, responsiveness, and usability are tested, proving that the system is valid for intelligent, accessible, and clinically relevant heart disease monitoring. An illustration of the block diagram of the whole system process may be seen in Figure 1.

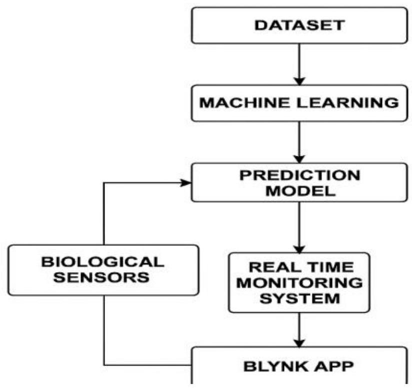

Figure 1. Workflow of Suggestedframework

4.1 Data source and Data pre-processing

This study employs the UCI heart disease dataset. The dataset comprises patient records containing individual profiles, which include age, sex, kind of chest discomfort, blood pressure, cholesterol levels, ECG results, maximum heart rate, and thalassemia status. The dependent variable is the presence or absence of heart illness.

In the preprocessing phase, missing data were detected and rectified, discrete attributes(cp, restecg, slope, thal) were encoded using one-hot encoding, and quantitative attributes(age, trestbps, chol, thalach, oldpeak, ca) were standardized to a consistent scale. The dataset was subsequently partitioned into Training & Testing groups in an 80:20 ratio via stratified sampling to preserve class equilibrium. These methodologies guaranteed pristine, uniform data appropriate for training LR, SVM, and RF models.

Feature selection

The data collection for software consists of thirteen distinct input parameters pertaining to various attributes of the cardiovascular system. A crucial component was identified for the improvement of the cloud-based categorization and forecasting system: the patient's mobile number, functioning as the patient's unique identifier. Quantifiable attributes are allocated a numerical scale to distinguish between the two designated categories (i.e., Healthy and Heart Disease). Table 1 delineates the input features, consisting of 13 physiological markers extracted from the two previously described datasets employed for predicting various forms of heart disease. Following feature selection, the data-set is divided into Training (80%) and Testing (20%) subsets.

Table 1. Clarification of thirteen input variables utilized for modelling development and authentication.

S. No	Attribute	Description
1	Age	Age in years
2	Sex	Male =1, Female =0
3	Cp	Chest pain type (Typical angina =1, atypical angina =2, no-anginal pain =3, asymptomatic =4)
4	Trestbps	Resting blood sugar (in mm Hg in case of admission to hospital)
5	Chol	Serum cholesterol in mg/dl
6	Fbs	Feasting blood sugar>120 mg/dl (True-1, false-0)
7	Restecg	Resting electrocardiographic results (normal =0, having ST-T wave abnormality =1, left ventricular hypertrophy =2)
8	Thalach	Maximum heart rate
9	Exang	Exercise-induced angina
10	Old peak	ST depression induced by exercise comparative to rest
11	Slope	Slope of the peak exercise ST segment (upsloping =1, flat =2, down sloping = 3)
12	Ca	Number of major vessels which are colored by fluoroscopy
13	Thal	Normal =0, fixed defect =2, reversible defect = 3

Classification using ML algorithms

4.1.1 Logistic Regression

In the realm of ML classification algorithms, LR stands out as one of the most straightforward. The method of supervised machine learning known as LR is used widely in a variety of applications. It is used for binary classification. A categorical dependent variable is used in its operation, and the outcomes it produces may be discrete or binary, which are denoted by the numbers 0 and 1. There is a cost function that is represented by the sigmoid function.An anticipated actual value is transformed into a probabilistic value ranging from 0 to 1 by the sigmoid function.

Logistic sigmoid function:

$$P(x) = 1/(1 + e^{(-x)}) \qquad (1)$$

A probability estimation function is denoted by the notation P(x), which produces values ranging from zero to one. The parameterx serves as the parameters for the probability function, representing value predicted by the algorithm. As can be seen in Equation (1), the value of the mathematical constant e, sometimes referred to as Euler's number, is around 2.71828. A logistic regression machine learning model is used in order to provide predictions about heart illness. In the beginning, the logistic regression model is trained under five distinct splitting circumstances. After that, it is assessed using a test dataset in order to reach the highest possible level of accuracy and evaluate the performance abilities of the modelling.

4.1.2 Random forest

It is a reputable and strong ML technique. It is a ML method which is also referred to as bagging or Bootstrapping Aggregation. The support structureis an effective statistical method of estimating data sample parameters, including the mean [36].Numerous samples of data are collected, mean is calculated and then, all the means are averaged to give a more accurate determination of the actual mean value. The process is also the same in bagging but the mean of each data specimens is not calculated; Decision trees are predominately applied instead. In this regard, a number of specimens of the training-data are analyzed and modelings are built regarding each specimen of data. Prediction of every dataset is also essential; every model provides a prediction, and they are averaged to obtain a more accurate assessment of the true value of the output.

4.1.3 Support Vector Machine

It is a categorizing algo., which can process nonlinear as well as lineardata. It makes use of non-Linear mapping operation that associates the Training-data on a non-linear basis to a greater dimensional realm that is easier to make a distinct difference between different classes [37]. A hyperplane, in SVM, is the decision boundary which separates the input variable space using class labels (0 or 1). This hyperplane can be illustrated by a line in a two-dimensional setting which separates data points of different classifications.The perimeter denotes the distance from the hyper plane to the nearest point of information in each of the classes. The hyperplane with the highest margin is the best, and it ensures maximum efficiency in making a distinction between the two classes. The points of the dataset that are nearest Vectors next to theHyperspaces are designated as support vectors, as they are crucial in the description and stabilization of the hyperplane.The configurations that give the maximum margin are determined without an optimization approach.

4.4 IoT Sensor Integration with Hardware setup

The physical layer consists of wearable medical sensors including ECG sensors for cardiac electrical activity monitoring, pulse oximeters for heart rate and blood oxygen saturation, and blood pressure monitors.These sensors interface with an ESP32 microcontroller through appropriate signal conditioning circuits and analog-to-digital converters.The ESP32 is programmed to collect sensor data at specified intervals, perform preliminary signal processing, and establish Wi-Fi connectivity for cloud communication. Power management circuits ensure continuous operation, and data validation checks are implemented at the hardware level to ensure signal quality.

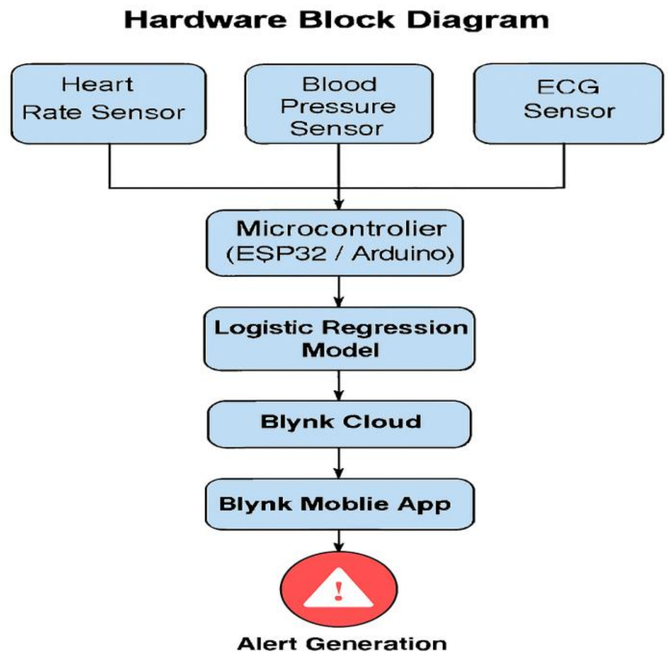

Figure2: Hardware block diagram

1. **ESP32 Microcontroller:** The ESP32 is a highly interconnected, energy-efficient,low-power, dual-core System on Chip (SoC) microcontroller that serves as the central processing unit in modern IoT-based healthcare monitoring systems as shown in Figure 4. Developed by Espressif Systems, it features two 32-bit LX6 microprocessors operating at frequencies up to 240 MHz, providing substantial computational capability for real-time data processing from multiple medical sensors simultaneously. The microcontroller incorporates both Wi-Fi (802.11 b/g/n) and dual-mode Bluetooth (Classic and BLE 4.2) connectivity, enabling seamless wireless communication with cloud platforms like Blynk and direct data transmission to mobile devices. Its rich peripheral interface includes multiple analog-to-digital converters (ADCs) with 12-bit resolution for precise sensor data acquisition from physiological monitors, pulse oximeters, and environmental sensors, along with digital interfaces such as I2C, SPI, UART, and PWM that facilitate communication with various medical sensors and output devices.

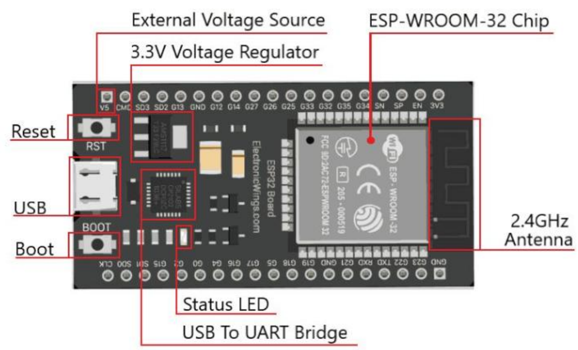

Figure 3. Diagram of ESP32 Microcontroller

- **DHT11:** It is a digital sensor for temp and humidity that delivers calibrated digital signal output via an integrated analog-to-digital converter.It measures the surrounding temperature with a reliability of ±2°C and moisture content with a reliability of ±5%, operating within a temperature range of 0-50°C and a humidity range of 20-90% with respect to humidity.
- **Blood Pressure:** This is a Non-invasive blood pressure monitoring system typically using the Oscillo metric method. It consists of an inflatable cuff, pressure sensor, and pump system thatassesses systolic pressure, diastolic pressure, and mean arterial pressure.
- **Pulse&Oxygen:** This integrated sensor typically uses photoplethysmography (PPG) technology with two primary measurements:
- **Heart Rate/Pulse:** Measures beats per minute using infrared light absorption patterns caused by blood volume changes in peripheral arteries.
- **Oxygen Saturation (SpO_2):** Calculates the percentage of oxygen-saturated hemoglobin in blood using red (660nm) and infrared (940nm) LEDs with photodetectors.
- **Buzzer:**It is an electrical apparatus utilizedto provide alert sounds. It is utilized to notify the caregiver in critical situations. This auditory phenomenonsignifies that the patient's condition is in critical hazard.
- **LED:** A visual alert system typically using LEDs that provides color-coded status information:
  a) Green LED: Normal conditions, low risk prediction.
  b) Yellow/Amber LED: Moderate risk, requires attention.
  c) Red LED: High risk prediction, immediate action needed.

Table 2 clarifies the associationamong temp, humidity, heart rate, and diverse action alternatives about danger levels and human comfort pertinent to the established continuous patient monitoring system.

Table 2.Boolean table for evaluating the status of a patient's environment.

Temp. sensor	Humidity	Human perception	Pulse rate sensor	Action taken	Risk Level
<37 ℃	31%-41%	Comfortable	60 to 100	No action	0/1
37 ℃-38 ℃	41%-52%	Comfortable for most people	40-60 or 100-120	Inform family	2
>38 ℃	46%-52%	Somewhat uncomfortable for most people at upper age	40-60 or 100-120	Inform local doctor	3
>38 ℃	>52%	Very humid, extremely uncomfortable	<40 or > 120	Inform emergency	4

4.5 Cloud serverand Blynk App design

The cloud server has a core, a substantial database capable of storing extensive data from various sensors, facilitating the seamless tracking of the system user's history. A diverse array of data analysis methods is integrated utilizing the database and application programming interfaces, including Google Sheets, enabling data visualization.

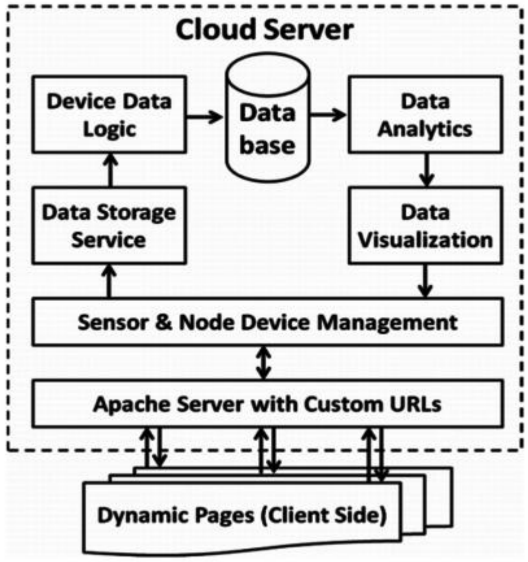

Figure 4. Diagram of Cloud server

4.6 Cloud (Blynk) App design

The Blynk platform configuration initiates with project instantiation in the Blynk Cloud ecosystem, wherein a unique authentication token is generated to establish secure bidirectional communication between the microcontroller unit (typically ESP32) and the cloud infrastructure. Subsequent interface development involves strategic widget placement within the mobile application framework, incorporating data visualization components for real-time physiological parameters (cardiac rhythm, hemodynamic measurements) and alert notification systems for clinical decision support. Hardware integration necessitates firmware implementation utilizing the Blynk IoT libraries, configuring wireless network parameters, and mapping sensor data streams to designated virtual pins that synchronize with corresponding application widgets. When it comes to monitoring the person's physiological temperature and moisture levels, the humidity and temperature sensors are quite necessary, especially in situations when the patient's health is serious. When a situation like this arises, the patient is unable to communicate his or her feelings to other people. Over the course of that time period, this apparatus will do an analysis of the various components of the atmosphere, such as the temperature and humidity, and it will also offer an evaluation of the patient's mental state in relation to those variables.It has a buzzer fitted into this system that would raise an alert when the heart-beat, temp, humidity, or any of these parameters surpass the standard minimum value.

This will notify the closest caregiver to the patient. This will also enable the doctor to keep an eye on the patient through the live video streaming application.

4.7 Performance measure

In terms of parameters, such as TP signifies "True positive," TN signifies"True Negative," FP signifies "False positive," and FN signifies"False negative values," the performance measure, Accuracy, Precision, f-measure, & recall are all computed.

Before calculating these metrics, it's essential to understand the four outcomes defined in a confusion matrix for each class:

- **True Positive (TP):** Technique accurately anticipateda positive outcome for a sample that is actually positive.
- **False Positive (FP):** Technique inaccurately anticipated a false positive result for a sample that is authentic negative.
- **True Negative (TN**Technique accurately anticipateda false negative result for a sample that is authentic negative.
- **False Negative (FN):** Technique inaccurately anticipated a false negative result for a sample that is authentic positive.
- **Precision:**It is the proportion of recovered documents that are pertinentto the query.

$$Precision = \frac{TP}{TP + FP}$$

Recall: It is defined as the proportion of pertinent documents that are retrieved

successfully.

$$Recall = \frac{TP}{TP + FN}$$

F-measure: It represents the harmonic mean of precision & recall.

$$F - measure = \frac{2TP}{(2TP + FP + FN)}$$

Accuracy: It is articulatedas the fraction of true positive & true negatives to that of the positive and negative observations.

$$Accuracy = \frac{(TP+TN)}{(TP+TN+FP+FN)}$$

5. Result and discussion

Experiments with the heart disease dataset were performed, testing various machine learning algorithms and doing comparisons. Logistic Regression showed the best accuracy and gave very interpretable results on risk factors, being appropriate for a clinical environment. This was implemented with IoT hardware, ESP32/Arduino with biological sensors, which allowed real-time acquisition and Transferring information to the public internet. The Blynk mobile app was able to display predictions, and immediate alerts were provided through a buzzer and monitor. The system was found to perform with respect to accuracy and responsiveness and had better usability, proving it has the potential to be useful in practice for prompt identification andcontinuous monitoring of cardiac disorder risk.

5.1 Proposed Model Results

LR is the Commonly used A classification algorithm that utilizes the logistic function to describe the probable outcome of a binary result. In this research, For the anticipation of the probable outcome of heart disease, the algorithm will be applied to patient health attributes such as (Age, Sex, chest pain type, blood pressure, cholesterol& ECG) results. The model computes a probable outcome score between 0 and 1, and values above the threshold, for example, 0.7, indicate high risk. The dataset is preprocessed by scaling continuous variables and encoding categorical features, thereafter divided into training and testing sets. After its training, the model has attained an accuracy of 93%, as evaluated using a confusion matrix and ROC curve. It integrates prediction result with the Blynk IoT mobile application for real-time visualization of patient vitals and automated generation of alerts for timely medical intervention.

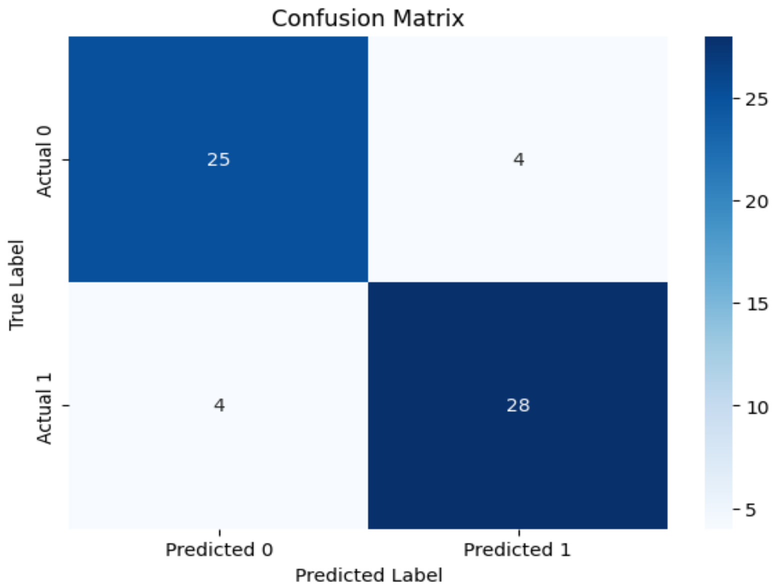

Figure5: prediction matrix using LR

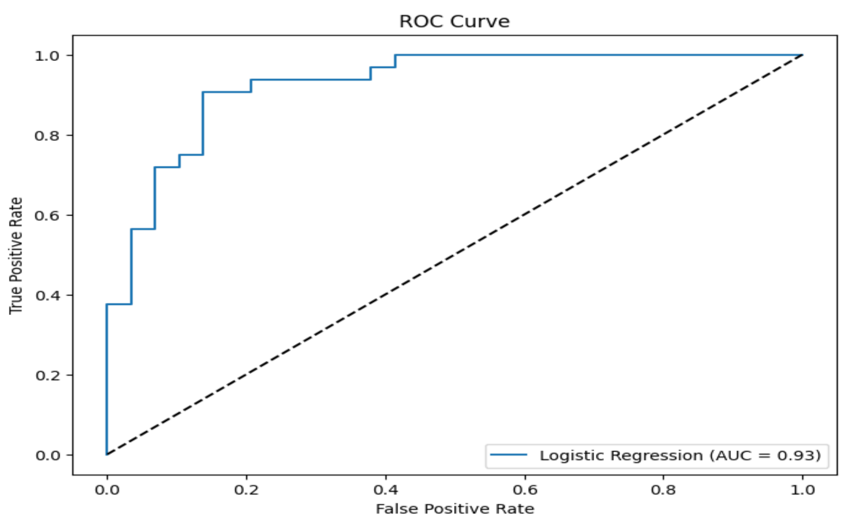

Figure6: ROC curve using LR

5.2 Base Modal Result

5.2.1 Support Vector Machine (SVM)

It was executed on the heart disease dataset to determine its predictive performance. This model reached averacity of 84.2%, aCorrectnessof 0.85, a recall of 0.82, and aScore of F1 is 0.83. These results identify that SVM is quite efficient in classifying the risk factors in patient data, though still somewhat bound by complexity of dataset. The confusion matrix below shows the categorizationoutcomes and the model's capacity to difference among high risk & low risk patientsas represented by the ROC curve as shown in Figure.

5.2.2 Random forest (RF)

The RF algorithm was also employed to test its performance regarding heart disease prediction. For this model, accuracy reached 86.7%, with acorrectnessof 0.87, recall of 0.85, and Score of F1 is 0.86. RF ensembles several decision chainsto boost robustness and reduce variance, hence enhancing the predictive capability of the modal. The Performance matrix and ROC curve (Figure) provide a visual presentation regarding the system's efficacy in classification and demonstrate that this classifier is reliable in detecting heart attack risk.

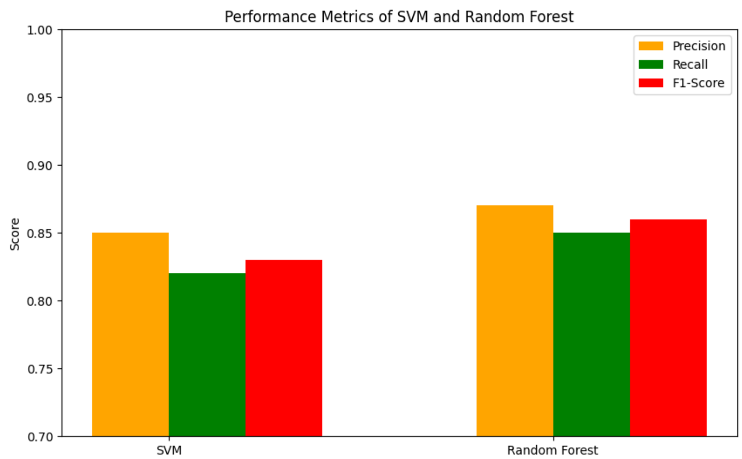

Figure7:Performance matrix of SVM and Random Forest

5.3 Evaluation

The dataset was evaluated using three Classification methods include LR,SVM, and RF.As shown in Table 3, Logistic Regression achieved the highest veracity(88.5%), Correctness (0.89), recall (0.87), Score of F1 (0.88), outperforming both SVM and Random Forest. These results demonstrate that Logistic Regression not only provides superior predictive performance but also maintains interpretability, rendering it the most appropriate method for coronary artery disease detection in this research.

Table3: Modal Evaluation using different types of methods

algorithm	accuracy	precision	recall	f1-score
LR	93	0.89	0.87	0.88
SVM	84.2	0.85	0.82	0.83
RF	86.7	0.87	0.85	0.86

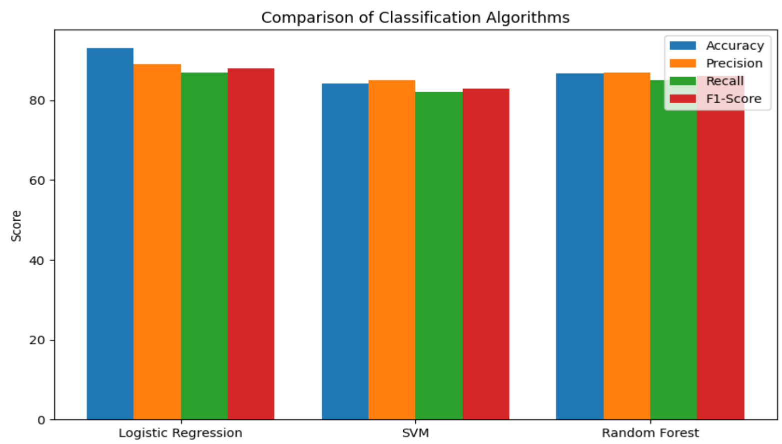

Figure8: comparison of classification Algorithm

6. Conclusion

The result presented the progress andperformance analysis of ML models for heart attack prediction by considering patient health parameters, including age, BP, cholesterol level, ECG results, and other clinical features. It implemented and compared LR, SVM, & RF methods on its dataset. Among these, LR had the highest Efficiency on the dataset, with veracity: 88.5%, Correctness: 0.89, recall: 0.87, Score of F1: 0.88, outperforming the other approaches, namely Support vector machine and Random Forest. Apart from the high veracity of the predicted output, this model was selected to be the primary model, because it is interpretable and suitable for clinical adoption. The proposed model was combined effectively with the Blynk IoT mobile application, providing visualization in real time of a patient's vitals, predicting the risk, and generating automatic alerts. This fills the gap between machine learning model prediction and actual deployment at an end-user level in healthcare by providing a lightweight, interpretable framework that is efficient for preventive cardiology.

7. Future work

- **Expanded Dataset:** Incorporating larger datasets from more diverse patient populations across different demographics for better generalizability.
- **Additional Features:** More physiological signals will be integrated (e.g., continuous ECG, oxygen saturation, lifestyle factors) for richer prediction.
- **Hybrid Models:** Exploring ensemble or hybrid approaches that combine Logistic Regression with advanced algorithms to balance accuracy and interpretability.
- **IoT Hardware Integration:** Integrating real sensors like heart rate, BP, ECG modules with ESP32/Arduino for continuous monitoring in natural environments.
- **Enhancing the features of the mobile application:** Expanding the Blynk Dashboard to include advanced visualizations, customized alerts, and integration with EHR.
- **Clinical Validation:** This involves cooperation with healthcare professionals, providing system validation in clinical settings, and assurance of compliance with medical standards.
- **Edge Computing:** Directly deploying lightweight ML models on microcontrollers can result in faster and offline predictions, independent of cloud connectivity

References

1. Alom, Zulfikar, Mohammad Abdul Azim, Zeyar Aung, Matloob Khushi, Josip Car, and Mohammad Ali Moni. "Early-stage detection of heart failure using machine learning techniques." In *Proceedings of the International Conference on Big Data, IoT, and Machine Learning: BIM 2021*, pp. 75-88. Singapore: Springer Singapore, 2021.
2. World Health Organization and J. Dostupno, cardiovascular diseases: key facts, vol. 13, no. 2016, p. 6, 2016. [Online]. Available: https://www.who.int/en/news-room/fact-sheets/detail/cardiovascular-diseases- (cvds).
3. Uyar, Kaan, and Ahmet İlhan. "Diagnosis of heart disease using genetic algorithm based trained recurrent fuzzy neural networks." *Procedia computer science* 120 (2017): 588-593.
4. Kausar, Noreen, Sellapan Palaniappan, Brahim Belhaouari Samir, Azween Abdullah, and Nilanjan Dey. "Systematic analysis of applied data miningbased optimization algorithms in clinical attribute extraction and classification for diagnosis of cardiac patients." In *Applications of Intelligent Optimization in Biology and Medicine: Current Trends and Open Problems*, pp. 217-231. Cham: Springer International Publishing, 2015.

5. Amin, Mohammad Shafenoor, Yin Kia Chiam, and Kasturi Dewi Varathan. "Identification of significant features and data mining techniques in predicting heart disease." *Telematics and Informatics* 36 (2019): 82-93.
6. Gour, Sanjay, Punita Panwar, Divya Dwivedi, and Chetan Mali. "A machine learning approach for heart attack prediction." In *Intelligent Sustainable Systems: Selected Papers of WorldS4 2021, Volume 1*, pp. 741-747. Singapore: Springer Nature Singapore, 2022.
7. Gupta, Chiradeep, Athina Saha, NV Subba Reddy, and U. Dinesh Acharya. "Cardiac Disease Prediction using Supervised Machine Learning Techniques." In *Journal of physics: conference series*, vol. 2161, no. 1, p. 012013. IOP Publishing, 2022.
8. Shameer, K., B. M. Smith, J. Kodysh, M. Yonker, B. S. Glicksberg, J. A. Udell, and J. T. Dudley. "Machine learning predictions of cardiovascular disease risk in a multi-ethnic population using electronic health record data." *Int. J. Med. Informatics* 146 (2021): 104335.
9. Liu, M., X. Sun, Y. Liu, X. Yang, Y. Xu, and X. Sun. "Deep learning-based prediction of coronary artery disease with CT angiography." *Japanese J. Radiol* 38, no. 4 (2020): 366-374.
10. Zakria, N., A. Raza, F. Liaquat, and S. G. Khawaja. "Machine learning based analysis of cardiovascular disease prediction." *J. Med. Syst* 41, no. 12 (2017): 207.
11. Yang, M., X. Wang, F. Li, and J. Wu. "A machine learning approach to identify risk factors for coronary heart disease: a big data analysis." *Comput. Methods Programs Biomed* 127 (2016): 262-270.
12. Ngufor, C., A. Hossain, S. Ali, and A. Alqudah. "Machine learning algorithms for heart disease prediction: a survey." *Int. J. Comput. Sci. Inform. Secur* 14, no. 2 (2016): 7-29.
13. Shoukat, A., S. Arshad, N. Ali, and G. Murtaza. "Prediction of Cardiovascular diseases using machine learning: a systematic review." *J. Med. Syst* 44, no. 8 (2020): 162.
14. Waigi, D., Dr Sonali Choudhary, Dr Punit Fulzele, and D. Mishra. "Predicting the risk of heart disease using advanced machine learning approach." *Eur. J. Mol. Clin. Med* 7, no. 7 (2020): 1638-1645.
15. Breiman, Leo. "Random forests." *Machine learning* 45, no. 1 (2001): 5-32.
16. Chen, Tianqi, and Carlos Guestrin. "Xgboost: A scalable tree boosting system." In *Proceedings of the 22nd acm sigkdd international conference on knowledge discovery and data mining*, pp. 785-794. 2016.

17. Gietzelt, Matthias, Klaus-Hendrik Wolf, Michael Marschollek, and Reinhold Haux. "Performance comparison of accelerometer calibration algorithms based on 3D-ellipsoid fitting methods." *Computer methods and programs in biomedicine* 111, no. 1 (2013): 62-71.
18. Saryazdi, Mohammad Dehghani, and Ali Mostafaeipour. "Identification and validation of key predictive factors for heart attack diagnosis using machine learning and fuzzy clustering." *Engineering Applications of Artificial Intelligence* 142 (2025): 109968.
19. Ahmed, Marzia, Mohd Herwan Sulaiman, Md Maruf Hassan, and Touhid Bhuiyan. "Predicting the classification of heart failure patients using optimized machine learning algorithms." *IEEE Access* (2025).
20. Stonier, Albert Alexander, Rakesh Krishna Gorantla, and K. Manoj. "Cardiac disease risk prediction using machine learning algorithms." *Healthcare Technology Letters* 11, no. 4 (2024): 213-217.
21. Hassan, Md Mehedi, Sadika Zaman, Md Mushfiqur Rahman, Anupam Kumar Bairagi, Walid El-Shafai, Rajkumar Singh Rathore, and Deepak Gupta. "Efficient prediction of coronary artery disease using machine learning algorithms with feature selection techniques." *Computers and Electrical Engineering* 115 (2024): 109130.
22. Khan, Arsalan, Moiz Qureshi, Muhammad Daniyal, and Kassim Tawiah. "A novel study on machine learning algorithm-based cardiovascular disease prediction." *Health & Social Care in the Community* 2023, no. 1 (2023): 1406060.
23. Bhatt, Chintan M., Parth Patel, Tarang Ghetia, and Pier Luigi Mazzeo. "Effective heart disease prediction using machine learning techniques." *Algorithms* 16, no. 2 (2023): 88.
24. Hassan, Ch Anwar Ul, Jawaid Iqbal, Rizwana Irfan, Saddam Hussain, Abeer D. Algarni, Syed Sabir Hussain Bukhari, Nazik Alturki, and Syed Sajid Ullah. "Effectively predicting the presence of coronary heart disease using machine learning classifiers." *Sensors* 22, no. 19 (2022): 7227.
25. Kaur, Bavneet, and Gaganpreet Kaur. "Heart disease prediction using modified machine learning algorithm." In *International Conference on Innovative Computing and Communications: Proceedings of ICICC 2022, Volume 1*, pp. 189-201. Singapore: Springer Nature Singapore, 2022.

26. Garg, Apurv, Bhartendu Sharma, and Rijwan Khan. "Heart disease prediction using machine learning techniques." In *IOP Conference series: materials science and engineering*, vol. 1022, no. 1, p. 012046. IOP Publishing, 2021.
27. Bharti, Rohit, Aditya Khamparia, Mohammad Shabaz, Gaurav Dhiman, Sagar Pande, and Parneet Singh. "Prediction of heart disease using a combination of machine learning and deep learning." *Computational intelligence and neuroscience* 2021, no. 1 (2021): 8387680.
28. Reddy, Karna Vishnu Vardhana, Irraivan Elamvazuthi, Azrina Abd Aziz, Sivajothi Paramasivam, Hui Na Chua, and S. Pranavanand. "Heart disease risk prediction using machine learning classifiers with attribute evaluators." *Applied Sciences* 11, no. 18 (2021): 8352.
29. Padmaja, B., Chintala Srinidhi, Kotha Sindhu, Kalali Vanaja, N. M. Deepika, and E. Krishna Rao Patro. "Early and accurate prediction of heart disease using machine learning model." *Turkish Journal of Computer and Mathematics Education* 12, no. 6 (2021): 4516-4528.
30. Rani, Pooja, Rajneesh Kumar, Nada MO Sid Ahmed, and Anurag Jain. "A decision support system for heart disease prediction based upon machine learning." *Journal of Reliable Intelligent Environments* 7, no. 3 (2021): 263-275.
31. Yaswanth, Raparthi, and Y. Md Riyazuddin. "Heart disease prediction using machine learning techniques." *International Journal of Innovative Technology and Exploring Engineering* 9, no. 5 (2020): 1456-1460.
32. Obasi, Thankgod, and M. Omair Shafiq. "Towards comparing and using Machine Learning techniques for detecting and predicting Heart Attack and Diseases." In *2019 IEEE international conference on big data (big data)*, pp. 2393-2402. IEEE, 2019.
33. Mohan, Senthilkumar, Chandrasegar Thirumalai, and Gautam Srivastava. "Effective heart disease prediction using hybrid machine learning techniques." *IEEE access* 7 (2019): 81542-81554.
34. UCI Machine Learning Repository: Heart Disease Data Set. http://archive.ics.uci.edu/ml/datasets/Heart+DiseaseStatlog Database.
35. Dhanka, Sanjay, and Surita Maini. "Random forest for heart disease detection: a classification approach." In *2021 IEEE 2nd International Conference on Electrical Power and Energy Systems (ICEPES)*, pp. 1-3. IEEE, 2021.

36. Vijayashree, J., and H. Parveen Sultana. "A machine learning framework for feature selection in heart disease classification using improved particle swarm optimization with support vector machine classifier." *Programming and Computer Software* 44, no. 6 (2018): 388-397.