An Effective Prediction of Heart Diseases Using Machine Learning Techniques

O. Obulesu, Koulampeta Yakshitha, Sanjana Vuyyuru, Tumkur Meenakshi, and Pujala Vaishnavi(✉)

Department of Computer Science and Engineering, G. Narayanamma Institute of Technology and Science, Hyderabad, Telangana, India
obulesh194@gnits.ac.in, pujalavaishnavi@gmail.com

Abstract. The heart plays a crucial role in the survival of living organisms by serving as the central pump that circulates blood throughout the body. Heart disease is often considered one of the most life-threatening conditions in humans due to its significant impact on health and mortality rates. The healthcare sector accumulates massive volumes of healthcare data, but regrettably, much of this valuable information often remains untapped, preventing the industry from uncovering hidden insights that could enhance decision-making processes. Hidden patterns and relationships are often left undiscovered and underutilized. The use of advanced machine learning techniques offers a promising solution to address the complexities involved in predicting heart diseases. It represents a crucial challenge within the realm of clinical data analysis, as accurately identifying and forecasting heart-related conditions is of paramount importance for improving healthcare outcomes. Early prediction of heart attack may save many lives hence preventing these has become more than necessary. Machine learning (ML) has the potential to provide highly efficient solutions for decision- making and precise predictive capabilities. In the field of machine learning, there are several classification models such as Logistic Regression, K- nearest neighbors (K-NN), and Support Vector Machine (SVM) that have proven effective in achieving the goal of predicting and diagnosing heart diseases. Decision Tree Classifier, in particular, serves as a valuable decision support system for detecting and forecasting heart diseases and heart attacks in individuals by utilizing risk factors associated with heart disease. Datasets containing medical parameters play a pivotal role in this process, as they are processed through various machine learning algorithms. These algorithms help uncover correlations among the different attributes present in the dataset using standard machine learning techniques. The overarching aim of this project is to employ machine learning techniques for the prediction and diagnosis of heart diseases, contributing to better healthcare outcomes.

Keywords: Random forest · Heart disease · Logistic regression · Decision tree · Machine learning

A. Kumar et al. (Eds.): ICDSMLA 2023, LNEE 1274, pp. 1044–1058, 2025.
https://doi.org/10.1007/978-981-97-8043-3_162

1 Introduction

Heart disease is often considered one of the most life-threatening conditions in humans due to its significant impact on health and mortality rates. The escalating incidence of cardiovascular ailments, characterized by their substantial mortality rates, is imposing a significant risk and strain on healthcare systems across the world. Cardiovascular diseases are generally more prevalent in men, especially during middle or old age, although they can also affect children. Cardiovascular diseases, which include heart—related conditions, have traditionally been a major contributor to the global mortality rate. These diseases encompass a range of health issues like coronary artery disease [1], heart failure, and irregular heart rhythms, all of which can lead to severe health problems and, in some instances, fatalities. The World Health Organization (WHO) and various health agencies emphasize the critical importance of preventive measures to mitigate the global public health burden of heart disease, which accounts for approximately 17.9 million annual deaths worldwide and is more prevalent in Asia. According to the European Cardiology Society (ESC), 26 million adults worldwide have received a heart disease diagnosis, with 3.6 million new cases identified annually. Alarmingly, about half of all patients diagnosed with heart disease succumb to the condition within just 1–2 years, and heart disease treatment consumes roughly 3% of the total healthcare budget.

To predict heart disease multiple tests are required. Insufficient knowledge or expertise among healthcare professionals can lead to incorrect forecasts or diagnoses in the medical field. It is difficult to identify heart diseases because of several contributory risk factors such as diabetes, high blood pressure, high cholesterol and many more factors. A variety of unhealthy habits can elevate the chances of developing heart disease. These include factors such as high cholesterol, obesity, raised triglyceride levels, and hypertension. Early diagnosis can be difficult. Accurate assessment of the risk of cardiac failure is crucial for preventing severe heart attacks and enhancing patient safety. Machine learning algorithms have proven effective in disease identification when trained on relevant data. Introducing machine learning and artificial intelligence [2] into healthcare research allows for the creation of robust prediction models using extensive databases.

Researchers often include different sets of risk factors or features when building predictors for heart disease, reflecting the complexity and diversity of this important area In the development of heart disease prediction models, various features such as age, sex, chest pain (cp), fasting blood sugar (FBS, which is linked to diabetes), resting electrocardiographic results (Restecg), exercise-induced angina (exang), ST depression induced by exercise relative to rest (oldpeak), slope, number of major vessels colored by fluoroscopy (ca), and heart status (thal) are considered. However, researchers are currently grappling with the challenge of effectively combining these features with suitable machine learning techniques to create accurate heart disease prediction models.

Background Study

The central objective is the development of systems capable of learning from past experiences and deriving meaningful predictions. In the context of predicting the aforementioned diseases, the prediction model employs categorization algorithms. The exploration of interconnected concepts, including machine learning and its techniques, is presented with concise definitions. This framework also encompasses the vital steps of data preprocessing and the utilization of assessment metrics to gauge model performance. Additionally, crucial details about the dataset under use are integrated into the model. The ultimate aim is to establish a robust predictive system that enhances early disease detection and supports informed healthcare decisions for the specified medical conditions.

1.1 Problem Statement

Heart diseases pose a significant health challenge, impacting a substantial portion of the population. Despite advancements in healthcare, the accurate prediction of heart diseases remains a crucial concern. The existing healthcare infrastructure is often burdened with limited resources, making it imperative to identify individuals at high risk early on. Machine learning techniques offer the potential to revolutionize this process by predicting heart diseases with greater accuracy and efficiency.

The problem at hand is the development of an effective predictive model that utilizes machine learning techniques to identify individuals who are at risk of developing heart diseases in the future. The challenge lies in harnessing the power of data-driven insights to improve early detection, enabling timely interventions and tailored preventive measures. The ultimate goal is to reduce the burden on healthcare systems and enhance patient outcomes by leveraging the predictive capabilities of machine learning in the domain of heart disease prediction.

1.2 Advantages and Drawbacks Advantages

The proposed solution leverages the integration of computerized patient records with clinical decision support systems, specifically using a dataset. This approach offers a range of benefits, including enhanced patient safety, reduced variations in clinical practices, decreased occurrences of medical errors, and ultimately, improved patient outcomes. The strategy's potential lies in the utilization of data analysis and modeling tools, particularly machine learning, such as Random Forest [3]. These tools have the capability to transform the healthcare landscape into a knowledge-rich domain, significantly enhancing the quality of healthcare options. In the modern healthcare industry, the utilization of machine learning and Random Forest techniques for prediction and decision-making has become increasingly crucial. The vast volume of medical data within the healthcare domain necessitates advanced analytical tools to harness its potential fully. Consequently, the integration of machine learning, especially through methods like Random Forest, plays a pivotal role in business intelligence for disease detection and prognosis. These approaches are instrumental in maximizing the value derived from medical data mining, thereby advancing the quality of healthcare decisions and ultimately benefitting patients and healthcare providers alike.

Drawbacks

The realm of medical diagnosis is considered vital yet intricate, necessitating precision and efficiency. Often, clinical judgments lean heavily on a physician's expertise and instincts, at times overlooking the treasure trove of data-driven insights concealed within databases. This approach not only impacts the quality of patient care but also unintentionally introduces biases, inaccuracies, and substantial healthcare expenses.

The establishment of a knowledge-enriched environment, empowered by machine learning, bears the potential to notably enhance the quality of medical decision-making. This advancement holds the promise of raising the standards of therapeutic judgments, representing a transformative stride in the evolution of healthcare practices.

1.3 Proposed System

The proposed system serves as a valuable decision support tool, offering assistance to physicians in the diagnostic process. Its primary objective is to predict the likelihood of heart disease in patients. The core approach involves the application of classification techniques to categorize the dataset into two distinct groups: “yes” and “no”. These classification techniques are implemented through various machine learning algorithms, including Decision Tree Classification, Random Forest, Grid Search CV, and Naïve Bayes Classification models.

These models play a pivotal role in elevating the accuracy of the classification process. They are designed to handle both classification and prediction tasks, all of which are executed using the Python programming language. This comprehensive approach harnesses the power of these machine learning models to enhance the accuracy and efficiency of heart disease prediction, providing valuable support to healthcare professionals in their diagnostic endeavors.

2 Literature Survey

This section delves into the extensive research conducted by various experts to employ computational methods in the realm of heart disease prediction. These studies have widely acknowledged the efficacy of such computer-based approaches. Historically, researchers have explored diverse methodologies to diagnose and anticipate heart ailments. However, their focus has often centered on highlighting the merits of specific computational techniques, rather than striving for their optimization through the utilization of ideal algorithms. In select instances, some investigators have ventured into the realm of hybrid optimization techniques, aiming to enhance the precision of heart disease prediction.

In a research study conducted by Latha and Jeeva [11], they utilized ensemble techniques to enhance the accuracy of predicting heart disease. Specifically, they employed bagging and boosting techniques to improve the performance of weak classifiers. To create a hybrid model, they incorporated several classifiers, including Naïve Bayes, Bayes Net, C 4.5, Multilayer Perceptron, PART, and Random Forest (RF). The ensemble model made decisions by majority voting among these classifiers, resulting in an 85.48%. In a more recent study, machine learning and traditional techniques, such as Random Forest (RF), Support Vector Machine (SVM), and other learning models, were tested using the UCI Heart Disease dataset [4] svm. By employing a voting-based modelthat combined multiple classifiers, they were able to further enhance accuracy. The research demonstrated a notable improvement of 2.1% in accuracy for classifiers that initially showed signs of being less accurate or “anemic.”

In the study conducted by Amin [2] and their team, they investigated crucial risk factors using a range of machine learning models, such as k-Nearest Neighbors (k-NN), Decision Trees (DT), Naive Bayes (NB), Logistic Regression (LR), Support Vector Machines (SVM), Neural Networks, and a hybrid model combining Naive Bayes and Logistic Regression through a voting mechanism. They conducted a comparative analysis of these models. The research findings demonstrated that the hybrid model, when used in conjunction with specific selected attributes, achieved an impressive accuracy rate of 87.41%. In the technique proposed by Saqlain and their collaborators, they employed the Mean Fisher Score Feature Selection Algorithm (MFSFSA) alongside the Support Vector Machine (SVM) classification model. The combination of MFSFSA and SVM [5] yielded notable results, including an accuracy rate of 81.19%, a sensitivity rate of 72.92%, and a specificity rate of 88.68%. These outcomes suggest the effectiveness of their approach in identifying and classifying pertinent features in their study.

3 Proposed Methodology

The proposed methodology employs Random Forest as a predictive model for identifying individuals with heart disease. The Random Forest is constructed in two stages: first, a set of N decision trees is created to form a diverse forest, and second, predictions are generated for each individual tree. All trees are constructed in a similar manner. To assess its predictive performance on new, unseen data, Random Forest employs a technique known as the out-of-bag error. This error is computed by evaluating each tree in the forest on data points that were not included in its bootstrap sample. Random Forest is versatile and can handle both classification and regression tasks effectively. It is particularly well-suited for datasets with a large number of features. In essence, Random Forest leverages the collective strength of multiple decision trees to provide accurate predictions, and it also provides valuable insights into the importance of different features in the dataset.

The proposed methodology employs Random Forest as a predictive model for identifying individuals with heart disease. The Random Forest is constructed in two stages: first, a set of N decision trees is created to form a diverse forest, and second, predictions are generated for each individual tree. All trees are constructed in a similar manner. To assess its predictive performance on new, unseen data, Random Forest employs a technique known as the out-of-bag error. This error is computed by evaluating each tree in the forest on data points that were not included in its bootstrap sample [6]. Random Forest is versatile and can handle both classification and regression tasks effectively. It is particularly well-suited for datasets with a large number of features. In essence, Random Forest leverages the collective strength of multiple decision trees to provide accurate predictions, and it also provides valuable insights into the importance of different features in the dataset (Fig. 1).

Objectives

The objectives for achieving effective prediction of heart diseases using machine learning encompass several facets, including:

(a) Early Detection: Develop a model that can detect signs of heart disease at an early stage, enabling timely intervention and treatment.

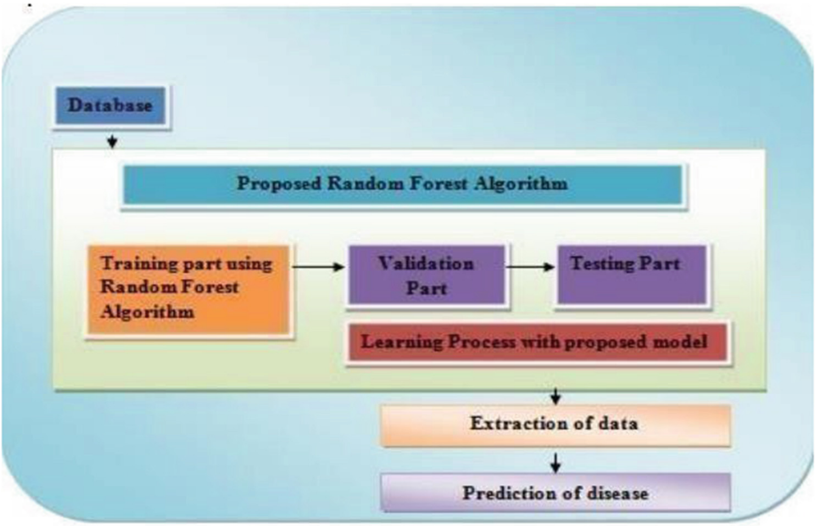


Fig. 1. Architecture of the proposed system

(b) Treatment Guidance: Provide personalized recommendations and guidance to healthcare professionals for tailoring treatment plans for individuals at risk of heart disease.
(c) Research Insights: Generate valuable insights and knowledge from the analysis of medical data, contributing to the broader understanding of heart disease causes and risk factors.

To construct a prediction model for heart disease using the Random Forest algorithm, the following architectural framework is employed:

Data Collection and Preprocessing: The data is preprocessed, addressing missing values, encoding categorical variables, and potentially scaling numerical features to ensure data quality and compatibility.
Dataset Splitting: The dataset is split into two segments: a training set and a testing/validation set. The training set is utilized to train the Random Forest model, while the testing set serves to assess its performance.
Random Forest Model Training: The next phase involves training a Random Forest model on the training dataset. Random Forest is an ensemble learning technique that amalgamates multiple decision trees for prediction.
Model Evaluation and Prediction: In the final stage, the trained Random Forest model can be employed to make predictions on new, unseen data by providing the relevant features as input.

4 Implementation

4.1 Dataset Description

"Framingham Heart Study" dataset is a well-known and widely used dataset in cardiovascular research and epidemiology. The Framingham Heart Study is a long-term, ongoing, community-based study that began in 1948 in Framingham, Massachusetts, USA.

- **Age, Float**
- **sex–Category**
  - 0 = female
  - 1 = male
- **cp**, chest pain,**Category.**
  - 0 = typical angina,
  - 1 = atypical angina,
  - 2 = non-anginal pain and 3 = asymptomatic
- **fbs**, fasting blood sugar, **Category**
  - 0 =>=120mg/dl
  - 1 = <120mg/dl
- **restbp**, resting blood pressure (in mmHg), **Float**
- **chol**, serum cholesterol in mg/dl, **Float**
- **Restecg**, resting electrocardio graphic results, **Category**
  - 0 = normal
  - 1 = having ST-T wave abnormality
  - 2 = showing probable or definite left ventricular hypertrophy
- **thalach**, maximum heart rate achieved, **Float**
- **exang**, exercise induced angina, **Category**
  - 0 = no
  - 1 = yes
- **oldpeak**, induced by exercise relative to rest. **Float**
- **slope**, the slope of the peak exercise ST segment,**Category**
  - 0 = positive ST segment
  - 1 = flat
  - 2 = negative ST segment

- **ca**, number of major vessels (0–3) colored by fluoroscopy, **Float**
- **Thal, Thalium**
  - 1 = normal (no cold spots)
  - 2 = fixed defect (cold spots during rest and exercise)
  - 3 = fireversible defect (when cold spots only appear during exercise (Table 4.1)

Table 4.1 Sample dataset

Age	Sex	cp	trestbps	chol	fbs	restecg	thalach	exang	oldpeak	slope	ca	thal	target
63	1	3	145	233	1	0	150	0	2.3	0	0	1	1
37	1	2	130	250	0	1	187	0	3.5	0	0	2	1
41	0	1	130	204	0	0	172	0	1.4	2	0	2	1
56	1	1	120	236	0	1	178	0	0.8	2	0	2	1
57	0	0	120	354	0	1	163	1	0.6	2	0	2	1
57	1	0	140	192	0	1	148	0	0.4	1	0	1	1
56	0	1	140	294	0	0	153	0	1.3	1	0	2	1

4.2 Data Preprocessing

Data preprocessing is a fundamental phase in machine learning that involves transforming raw and often messy data into a structured and clean dataset suitable for training machine learning models. It plays a critical role in ensuring that the data is prepared for analysis and modeling. Let's delve deeper into the key steps involved in data preprocessing:

(a) Importing Required Libraries: Data preprocessing in Python relies on specific libraries that offer a range of functions and tools. These libraries are essential for performing various tasks in data preparation. NumPy, for instance, is a fundamental library that supports mathematical operations, making it invaluable for data manipulation. Matplotlib, on the other hand, enables data visualization, allowing data analysts to gain insights and make data-driven decisions. Pandas is a versatile library used for dataset management, offering powerful data structures and data analysis tools. Importing these libraries is the first step to equip oneself with the necessary tools for data preprocessing.

(b) Importing the Dataset: The dataset, often stored in a structured format like CSV, is loaded into the Python environment using Pandas. This step is pivotal, as it brings the data into the analysis pipeline, making it accessible for further processing.

This step lays the foundation for model development by separating the input data from the target variable.

(c) Handling Missing Data: Missing data is a common challenge in real-world datasets and can disrupt the performance of machine learning models. Two primary approaches are widely used:

1. Removing rows or columns with null or missing values: This method is effective but can lead to data loss, potentially reducing the dataset's informativeness.
2. Calculating and replacing missing values with the mean of the respective column or row: This approach is particularly useful for numeric data attributes, such as age or salary. It helps retain data integrity while filling in the gaps.

(d) Encoding Categorical Data: Machine learning models require numeric data. To deal with categorical variables, encoding is necessary to convert them into numerical representations for seamless model building. To accommodate these categorical variables, data preprocessing involves encoding them into numerical representations. Techniques like one-hot encoding or label encoding are commonly used to ensure the model can work effectively with these data. The Scikit-learn library, along with its Imputer class, is frequently employed to handle missing values and streamline the data preprocessing workflow. Data preprocessing is a crucial aspect of the machine learning pipeline, as the quality of the data directly impacts the performance and reliability of the final predictive models.

4.3 Splitting the Data

Training a model with one dataset and testing it with another can lead to difficulties in understanding correlations. To ensure a model performs well with both training and test data, we create two subsets:

(a) Training Set: This subset is used to train the machine learning model, and we already know the expected output.

(b) Test Set: This subset is employed to test the machine learning model's performance. The model uses this set to predict the output.To create these training and test sets, we establish four sets:

X_train: The training portion of the feature matrix.
X_test: The testing portion of the feature matrix.
Y_train: The corresponding dependent variables for the X_train set.
Y_test: The corresponding dependent variables for the X_test set.

We utilize the test_train_split function, specifyingthe arrays (X and Y) and the test_size parameter to determine the data split ratio. This process ensures that the model can effectively perform with both the training and test datasets.

4.4 Model Training

A training model in machine learning is a crucial component that plays a fundamental role in the development of intelligent algorithms. This model is essentially a carefully curated dataset that consists of pairs of input data and corresponding output data. The input data represents the various variables or features that influence the output, while the output data signifies the desired or expected results. For instance, in a medical diagnosis scenario, the input data might encompass patient characteristics, such as age, medical history, and symptoms, while the output data indicates whether the patient has a particular disease or not.

The training process involves feeding this dataset to the machine learning algorithm, allowing it to learn and understand the intricate relationships and patterns that exist within the data. As the algorithm processes the input data, it starts making predictions based on its current understanding. These predictions are then meticulously compared to the known output data. Crucially, the iterative nature of this process, known as "model fitting," comes into play. During each iteration, the algorithm fine-tunes its internal parameters and decision-making processes with a clear objective: to minimize the disparities between its predictions and the actual output data. This continual adjustment and refinement are at the heart of model training [9], with each iteration making the model more precise and accurate in its predictions.

4.5 Model Evaluation

In the domain of machine learning, model evaluation is an essential and intricate process that plays a fundamental role in the development and refinement of predictive models. The goal of this process is to thoroughly and systematically assess the model's performance and its ability to make accurate predictions for future scenarios. Model evaluation is pivotal because it serves several crucial purposes. Firstly, it provides a comprehensive understanding of how well the model performs in practical applications. By employing a range of evaluation metrics, we gain insights into different aspects of the model's performance, allowing us to identify both its strengths and weaknesses. Several metrics are commonly used to evaluate machine learning models, each shedding light on specific dimensions of performance. Several metrics, including Accuracy, Precision, Recall, F1 score, Area under Curve [7], Confusion Matrix, and Mean Square Error [8], are employed to assess the model's effectiveness (Fig. 2).

5 Results and Discussion

5.1 Input and Output

Figure 3 presents multivariate data, likely comprising various patient attributes, used as input features for predicting th e presence or absence of heart disease. Figure 4 depicts the results of an experimental task, likely evaluating the accuracy and performance of a predictive model in diagnosing heart disease based on the data from Fig. 4. Together, these figures illustrate the process of using patient data to make informed predictions and assess the model's diagnostic capabilities.

The Figs. 5 and 6 illustrates a system where, using user-specific information, it offers a determination of whether an individual has a heart disease or not. In the event of a heart disease diagnosis, the system provides access to relevant doctor details for further medical consultation. Conversely, if there is no indication of heart disease, the system offers a personalized diet chart as a health-related resource (Figs. 7, 8, 9).

6 Conclusion and Future Enhancements

Random Forest stands out as a highly promising algorithm for predictive analysis, showcasing its exceptional accuracy in predicting outcomes, especially when applied to extensive datasets. In the realm of healthcare, it exhibits significant potential, particularly in

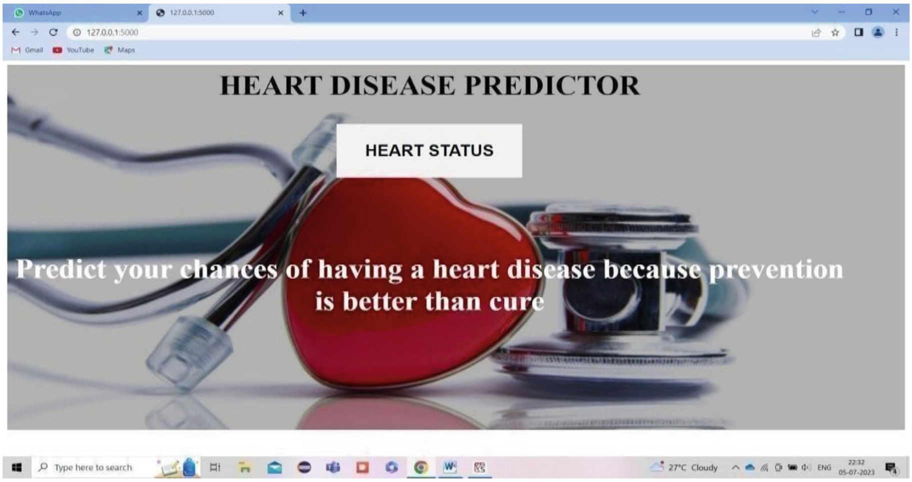


Fig. 2. Input page

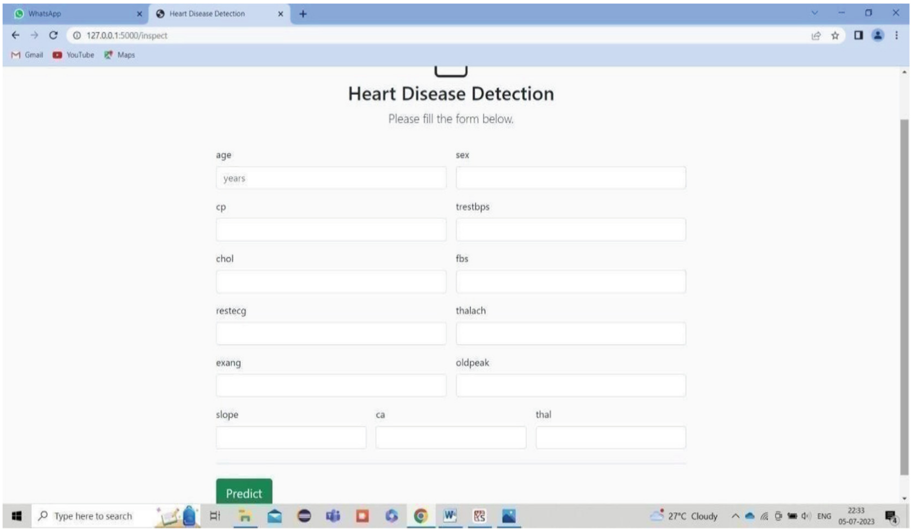


Fig. 3. Input form

early predictions of conditions like heart diseases. For this system to reach its full potential, it is essential to recognize that it currently operates with a limited dataset. Machine learning algorithms tend to improve their accuracy with exposure to larger, more diverse datasets. Therefore, expanding the dataset used for training is a critical step to significantly enhance the accuracy of heart disease predictions. Additionally, integrating the system with electronic systems that offer real-time inputs is a pivotal advancement. This integration would not only enhance its value but also provide patients with immediate results, a feature that can be crucial in critical healthcare scenarios. The completion of

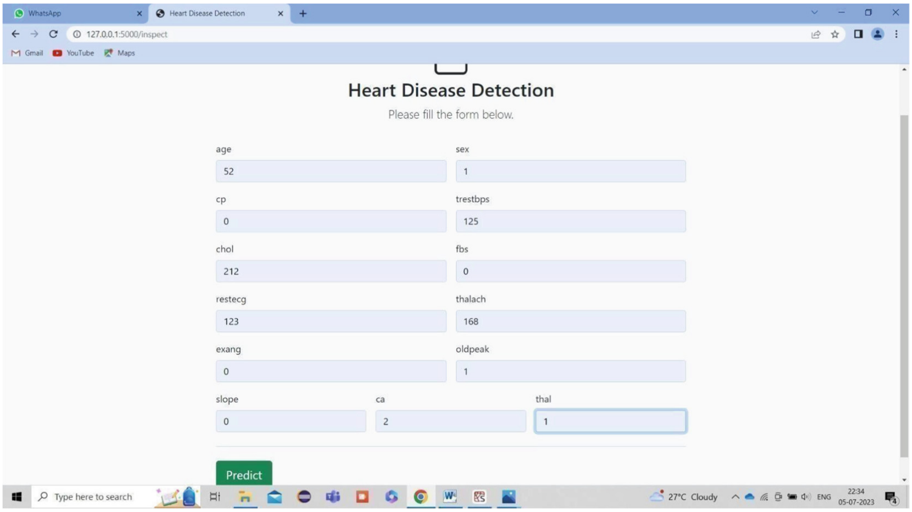


Fig. 4. Filled input form

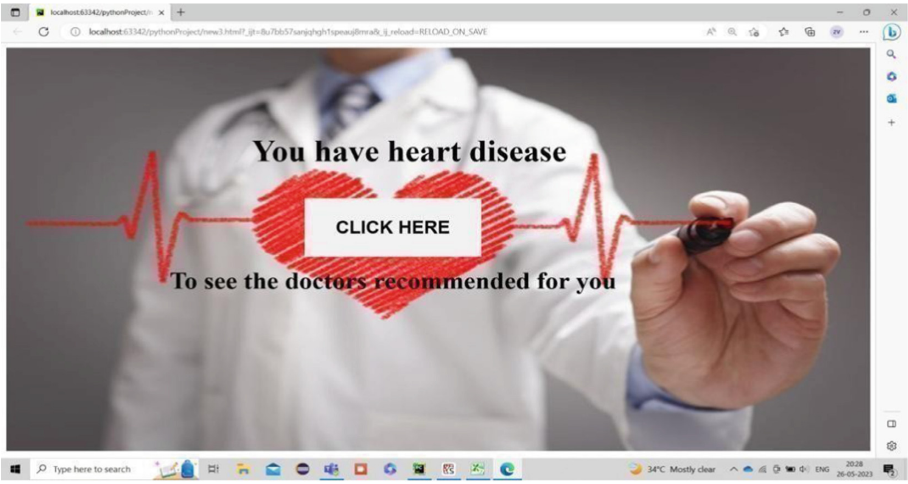


Fig. 5. You have heart disease

the analysis component marks a significant milestone in the system's development, and seamless integration with real-time data input systems will further elevate its utility.

Moreover, exploring a variety of algorithm combinations on the expanded dataset is a strategy worth pursuing. By testing multiple algorithm combinations on these datasets, we can aim for even better predictive results. This iterative approach can refine the accuracy and efficiency of heart disease predictions, ultimately benefiting patient care and contributing to medical research. In conclusion, Random Forest's potential for predicting heart diseases is substantial, and its performance can be further optimized through

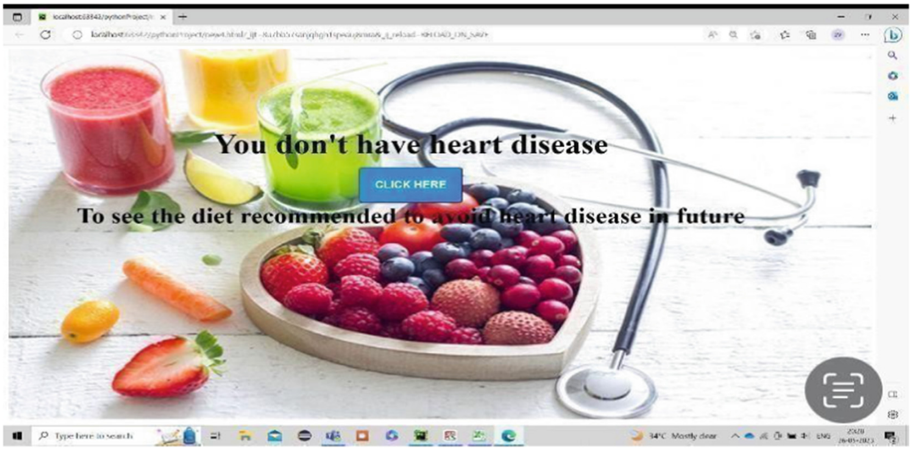
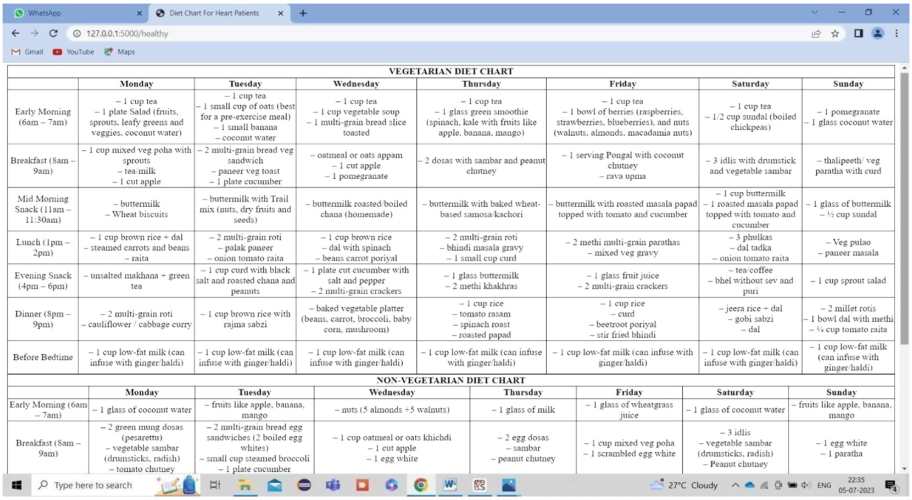


Fig. 6. You don't have heart disease

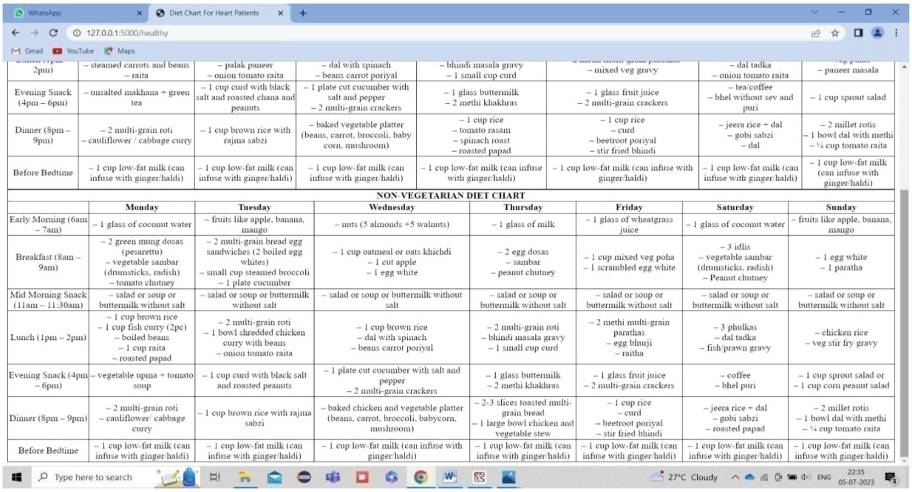


VEGETARIAN DIET CHART	Monday	Tuesday	Wednesday	Thursday	Friday	Saturday	Sunday
Early Morning (6am – 7am)	– 1 cup tea – 1 plate Salad (fruits, sprouts, leafy greens and veggies, coconut water)	– 1 cup tea – 1 small cup of oats (best for a pre-exercise meal) – 1 small banana – coconut water	– 1 cup tea – 1 cup vegetable soup – 1 multi-grain bread slice toasted	– 1 cup tea – 1 glass green smoothie (spinach, kale with fruits like apple, banana, mango)	– 1 cup tea – 1 bowl of berries (raspberries, strawberries, blueberries), and nuts (walnuts, almonds, macadamia nuts)	– 1 cup tea – 1/2 cup sundal (boiled chickpeas)	– 1 pomegranate – 1 glass coconut water
Breakfast (8am – 9am)	– 1 cup mixed veg poha with sprouts – tea/milk – 1 cut apple	– 2 multi-grain bread veg sandwich – paneer veg toast – 1 plate cucumber	– oatmeal or oats appam – 1 cut apple – 1 pomegranate	– 2 dosas with sambar and peanut chutney	– 1 serving Pongal with coconut chutney – rava upma	– 3 idlis with drumstick and vegetable sambar	– thalipeeth/ veg paratha with curd
Mid Morning Snack (11am – 11:30am)	– buttermilk – Wheat biscuits	– buttermilk with Trail mix (nuts, dry fruits and seeds)	– buttermilk roasted/boiled chana (homemade)	– buttermilk with baked wheat-based samosa/kachori	– buttermilk with roasted masala papad topped with tomato and cucumber	– 1 cup buttermilk – 1 roasted masala papad topped with tomato and cucumber	– 1 glass of buttermilk – ½ cup sundal
Lunch (1pm – 2pm)	– 1 cup brown rice + dal – steamed carrots and beans – raita	– 2 multi-grain roti – palak paneer – onion tomato raita	– 1 cup brown rice – dal with spinach – beans carrot poriyal	– 2 multi-grain roti – bhindi masala gravy – 1 small cup curd	– 2 methi multi-grain parathas – mixed veg gravy	– 3 phulkas – dal tadka – onion tomato raita	– Veg pulao – paneer masala
Evening Snack (4pm – 6pm)	– unsalted makhana + green tea	– 1 cup curd with black salt and roasted chana and peanuts	– 1 plate cut cucumber with salt and pepper – 2 multi-grain crackers	– 1 glass buttermilk – 2 methi khakhras	– 1 glass fruit juice – 2 multi-grain crackers	– tea/coffee – bhel without sev and puri	– 1 cup sprout salad
Dinner (8pm – 9pm)	– 2 multi-grain roti – cauliflower / cabbage curry	– 1 cup brown rice with rajma sabzi	– baked vegetable platter (beans, carrot, broccoli, baby corn, mushroom)	– 1 cup rice – tomato rasam – spinach roast – roasted papad	– 1 cup rice – curd – beetroot poriyal – stir fried bhindi	– jeera rice + dal – gobi sabzi – dal	– 2 millet rotis – 1 bowl dal with methi – ¼ cup tomato raita
Before Bedtime	– 1 cup low-fat milk (can infuse with ginger/haldi)	– 1 cup low-fat milk (can infuse with ginger/haldi)	– 1 cup low-fat milk (can infuse with ginger/haldi)	– 1 cup low-fat milk (can infuse with ginger/haldi)	– 1 cup low-fat milk (can infuse with ginger/haldi)	– 1 cup low-fat milk (can infuse with ginger/haldi)	– 1 cup low-fat milk (can infuse with ginger/haldi)

NON-VEGETARIAN DIET CHART	Monday	Tuesday	Wednesday	Thursday	Friday	Saturday	Sunday
Early Morning (6am – 7am)	– 1 glass of coconut water	– fruits like apple, banana, mango	– nuts (5 almonds +5 walnuts)	– 1 glass of milk	– 1 glass of wheatgrass juice	– 1 glass of coconut water	– fruits like apple, banana, mango
Breakfast (8am – 9am)	– 2 green mung dosas (pesarettu) – vegetable sambar (drumsticks, radish) – tomato chutney	– 2 multi-grain bread egg sandwiches (2 boiled egg whites) – small cup steamed broccoli – 1 plate cucumber	– 1 cup oatmeal or oats khichdi – 1 cut apple – 1 egg white	– 2 egg dosas – sambar – peanut chutney	– 1 cup mixed veg poha – 1 scrambled egg white	– 3 idlis – vegetable sambar (drumsticks, radish) – Peanut chutney	– 1 egg white – 1 paratha

Fig. 7. Diet chart 1

expanding the dataset for training, integration with real-time data input systems, and the exploration of various algorithm combinations. With these advancements, the system can offer more accurate and timely predictions, ultimately enhancing patient care and contributing to ongoing research in cardiology.

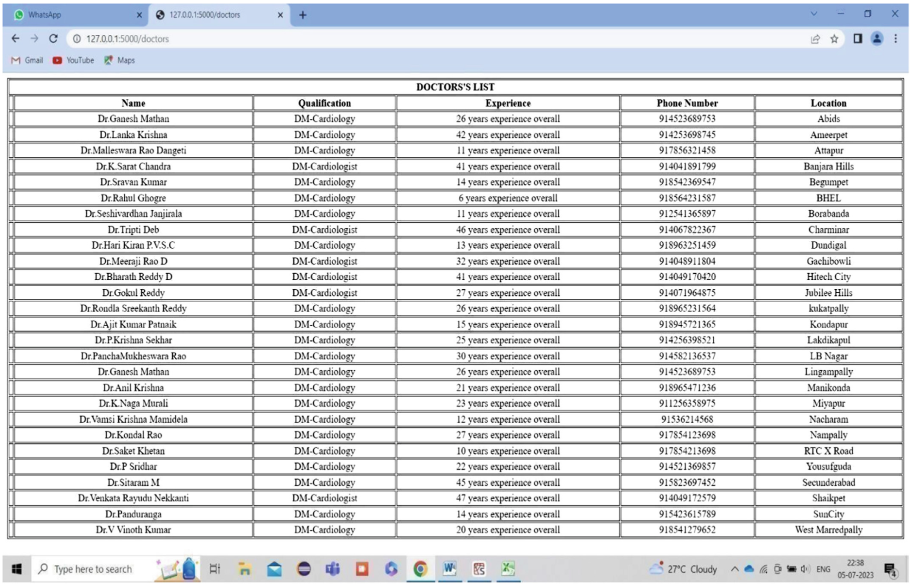


2pm)	– steamed carrots and beans – raita	– palak paneer – onion tomato raita	– dal with spinach – beans carrot poriyal	– bhindi masala gravy – 1 small cup curd	– mixed veg gravy	– dal tadka – onion tomato raita	– paneer masala
Evening Snack (4pm – 6pm)	– unsalted makhana + green tea	– 1 cup curd with black salt and roasted chana and peanuts	– 1 plate cut cucumber with salt and pepper – 2 multi-grain crackers	– 1 glass buttermilk – 2 methi khakhras	– 1 glass fruit juice – 2 multi-grain crackers	– tea/coffee – bhel without sev and puri	– 1 cup sprout salad
Dinner (8pm – 9pm)	– 2 multi-grain roti – cauliflower / cabbage curry	– 1 cup brown rice with rajma sabzi	– baked vegetable platter (beans, carrot, broccoli, baby corn, mushroom)	– 1 cup rice – tomato rasam – spinach roast – roasted papad	– 1 cup rice – curd – beetroot poriyal – stir fried bhindi	– jeera rice + dal – gobi sabzi – dal	– 2 millet rotis – 1 bowl dal with methi – ¼ cup tomato raita
Before Bedtime	– 1 cup low-fat milk (can infuse with ginger/haldi)	– 1 cup low-fat milk (can infuse with ginger/haldi)	– 1 cup low-fat milk (can infuse with ginger/haldi)	– 1 cup low-fat milk (can infuse with ginger/haldi)	– 1 cup low-fat milk (can infuse with ginger/haldi)	– 1 cup low-fat milk (can infuse with ginger/haldi)	– 1 cup low-fat milk (can infuse with ginger/haldi)

NON-VEGETARIAN DIET CHART

	Monday	Tuesday	Wednesday	Thursday	Friday	Saturday	Sunday
Early Morning (6am – 7am)	– 1 glass of coconut water	– fruits like apple, banana, mango	– nuts (5 almonds +5 walnuts)	– 1 glass of milk	– 1 glass of wheatgrass juice	– 1 glass of coconut water	– fruits like apple, banana, mango
Breakfast (8am – 9am)	– 2 green mung dosas (pesarettu) – vegetable sambar (drumsticks, radish) – tomato chutney	– 2 multi-grain bread egg sandwiches (2 boiled egg whites) – small cup steamed broccoli – 1 plate cucumber	– 1 cup oatmeal or oats khichdi – 1 cut apple – 1 egg white	– 2 egg dosas – sambar – peanut chutney	– 1 cup mixed veg poha – 1 scrambled egg white	– 3 idlis – vegetable sambar (drumsticks, radish) – Peanut chutney	– 1 egg white – 1 paratha
Mid Morning Snack (11am – 11:30am)	– salad or soup or buttermilk without salt	– salad or soup or buttermilk without salt	– salad or soup or buttermilk without salt	– salad or soup or buttermilk without salt	– salad or soup or buttermilk without salt	– salad or soup or buttermilk without salt	– salad or soup or buttermilk without salt
Lunch (1pm – 2pm)	– 1 cup brown rice – 1 cup fish curry (2pc) – boiled beans – 1 cup raita – roasted papad	– 2 multi-grain roti – 1 bowl shredded chicken curry with beans – onion tomato raita	– 1 cup brown rice – dal with spinach – beans carrot poriyal	– 2 multi-grain roti – bhindi masala gravy – 1 small cup curd	– 2 methi multi-grain parathas – egg bhurji – raitha	– 3 phulkas – dal tadka – fish/prawn gravy	– chicken rice – veg stir fry gravy
Evening Snack (4pm – 6pm)	– vegetable upma + tomato soup	– 1 cup curd with black salt and roasted peanuts	– 1 plate cut cucumber with salt and pepper – 2 multi-grain crackers	– 1 glass buttermilk – 2 methi khakhras	– 1 glass fruit juice – 2 multi-grain crackers	– coffee – bhel puri	– 1 cup sprout salad or – 1 cup corn peanut salad
Dinner (8pm – 9pm)	– 2 multi-grain roti – cauliflower/ cabbage curry	– 1 cup brown rice with rajma sabzi	– baked chicken and vegetable platter (beans, carrot, broccoli, babycorn, mushroom)	– 2-3 slices toasted multi-grain bread – 1 large bowl chicken and vegetable stew	– 1 cup rice – curd – beetroot poriyal – stir fried bhindi	– jeera rice + dal – gobi sabzi – roasted papad	– 2 millet rotis – 1 bowl dal with methi – ¼ cup tomato raita
Before Bedtime	– 1 cup low-fat milk (can infuse with ginger/haldi)	– 1 cup low-fat milk (can infuse with ginger/haldi)	– 1 cup low-fat milk (can infuse with ginger/haldi)	– 1 cup low-fat milk (can infuse with ginger/haldi)	– 1 cup low-fat milk (can infuse with ginger/haldi)	– 1 cup low-fat milk (can infuse with ginger/haldi)	– 1 cup low-fat milk (can infuse with ginger/haldi)

Fig. 8. Diet chart 2



DOCTORS'S LIST

Name	Qualification	Experience	Phone Number	Location
Dr.Ganesh Mathan	DM-Cardiology	26 years experience overall	914523689753	Abids
Dr.Lanka Krishna	DM-Cardiology	42 years experience overall	914253698745	Ameerpet
Dr.Malleswara Rao Dangeti	DM-Cardiology	11 years experience overall	917856321458	Attapur
Dr.K.Sarat Chandra	DM-Cardiologist	41 years experience overall	914041891799	Banjara Hills
Dr.Sravan Kumar	DM-Cardiology	14 years experience overall	918542369547	Begumpet
Dr.Rahul Ghogre	DM-Cardiology	6 years experience overall	918564231587	BHEL
Dr.Seshivardhan Janjirala	DM-Cardiology	11 years experience overall	912541365897	Borabanda
Dr.Tripti Deb	DM-Cardiologist	46 years experience overall	914067822367	Charminar
Dr.Hari Kiran P.V.S.C	DM-Cardiology	13 years experience overall	918963251459	Dundigal
Dr.Meeraji Rao D	DM-Cardiologist	32 years experience overall	914048911804	Gachibowli
Dr.Bharath Reddy D	DM-Cardiologist	41 years experience overall	914049170420	Hitech City
Dr.Gokul Reddy	DM-Cardiologist	27 years experience overall	914071964875	Jubilee Hills
Dr.Rondla Sreekanth Reddy	DM-Cardiology	26 years experience overall	918965231564	kukatpally
Dr.Ajit Kumar Patnaik	DM-Cardiology	15 years experience overall	918945721365	Kondapur
Dr.P.Krishna Sekhar	DM-Cardiology	25 years experience overall	914256398521	Lakdikapul
Dr.PanchaMukheswara Rao	DM-Cardiology	30 years experience overall	914582136537	LB Nagar
Dr.Ganesh Mathan	DM-Cardiology	26 years experience overall	914523689753	Lingampally
Dr.Anil Krishna	DM-Cardiology	21 years experience overall	918965471236	Manikonda
Dr.K.Naga Murali	DM-Cardiology	23 years experience overall	911256358975	Miyapur
Dr.Vamsi Krishna Mamidela	DM-Cardiology	12 years experience overall	91536214568	Nacharam
Dr.Kondal Rao	DM-Cardiology	27 years experience overall	917854123698	Nampally
Dr.Saket Khetan	DM-Cardiology	10 years experience overall	917854213698	RTC X Road
Dr.P Sridhar	DM-Cardiology	22 years experience overall	914521369857	Yousufguda
Dr.Sitaram M	DM-Cardiology	45 years experience overall	915823697452	Secunderabad
Dr.Venkata Rayudu Nekkanti	DM-Cardiologist	47 years experience overall	914049172579	Shaikpet
Dr.Panduranga	DM-Cardiology	14 years experience overall	915423615789	SunCity
Dr.V Vinoth Kumar	DM-Cardiology	20 years experience overall	918541279652	West Marredpally

Fig. 9. Doctors list

References

1. Ashraf et al (2020) Prediction of heart disease using ensemble approach. Int J Pharmaceutical Res 12(4):56–66
2. Amin et al (2019) Heart disease prediction system using machine learning models. In Proceedings of the 5th International Conference for Convergent in Technology (I2CT), pp 1–4

3. Javeed A, Zhou S, Yongjian L, Qasim I, Noor A, Nour R (2019) An intelligent learning system based on random search algorithm and optimized random forest model for improved heart disease detection. IEEE Access 7:180235–180243
4. Gupta A, Kumar R, Arora HS, Raman B (2020) MIFH: a machine intelligence framework for heart disease diagnosis. IEEE Access 8:14659–21467
5. Yang C, An B, Yin S (2018) Heart-disease diagnosis via support vector machine-based approaches. In Proceedings of IEEE International Conference on Systems, Man and Cybernetics (SMC), pp 3153–3158
6. Zhao J, Yi L, Wang D (2021) Research on heart disease prediction algorithm based on optimized random forest. J Qingdao Univ Sci Technol pp 491–500
7. Patel J, Upadhyay T, Patel S (2016) Heart disease prediction using machine learning and data mining technique. Int J Comput Sci Commun September 2015-March 2016, pp 129–137
8. Srinivas K, Rani B, Govardhan D (2010) Applications of data mining techniques in healthcare and prediction of heart attacks. Int J Comput Sci Eng 2:250–255
9. Polaraju K, Durga Prasad D (2017) Prediction of heart disease using multiple linear regression model. Int J Eng Dev Res Dev. ISSN: 2321-9939
10. Gomathi K, Shanmuga Priyaa D (2016) Multi disease prediction using data mining techniques. Int J Syst Softw Eng, pp 12–14
11. Latha, Jeeva, Sun R (2018), A hybrid intelligent system framework for the prediction of heart disease using random forest algorithm. Mobile Inf Syst, 1–21
12. Yahaya L, Oye ND, Garba EJ (2020) Acomprehensive review on heart disease prediction using data mining and machine learning techniques. Amer J Artif Intell