IJNRD.ORG ISSN : 2456-4184
INTERNATIONAL JOURNAL OF NOVEL RESEARCH AND DEVELOPMENT (IJNRD) | IJNRD.ORG
An International Open Access, Peer-reviewed, Refereed Journal

ENHANCING HEART DISEASE PREDICTION SYSTEM USING IOT AND MACHINE LEARNING

M.Azhagiri, Shroin Mandal, Disha Sachdeva, Amisha Singh
Department Of Computer Science and Engineering , SRM Institute of Science And Technology, Ramapuram Campus, Chennai, India.

1. ABSTRACT

Heart disease is still a major global health concern, but its effects can be significantly diminished via early detection and prevention. Using a combination of established risk variables and cutting-edge machine learning methods and

For the predicted accuracy, a machine learning ensemble model is used. This ensemble model combines neural networks, support vector machines, and decision trees to accurately represent intricate nonlinear interactions between the variables.

Validation of sizable and varied patient dataset, including both those with and without cardiac disease, is used to run the model. The model's performance indicators, such as accuracy, precision, recall, and F1-score, are used to assess how well it can identify those who are at high risk. To demonstrate the improvements in accuracy and dependability, comparisons are done against current risk prediction models. The findings show the integrated model's improved predictive power and offer clinicians and other healthcare professionals a useful tool for identifying those who are more likely to acquire heart disease.

By providing insights into the main causes of the disease, this method may help develop customized preventive measures. The suggested integrated predictive model has promise for enhancing public health outcomes and fostering proactive cardiac treatment because heart disease continues to be a major cause of mortality.

Keywords: IoT, Machine Learning, Decision Tree, PPG Sensors, SVM, Model Prediction

2. INTRODUCTION

Heart disease is a major global health concern, and combining machine learning (ML) and internet of things (IoT) technology can improve early identification and prevention. Healthcare systems can use real-time data gathering capabilities from IoT-enabled devices to gain a more accurate picture of a person's cardiovascular health. ML algorithms can adapt to a patient's changing condition, enabling early detection of anomalies and trends that could indicate heart disease onset.

Remote monitoring through IoT devices (IOT devices like smart-watch that have integrated heart rate monitoring sensors like PPG (Photoplethysmography) sensors, this system forecasts the likelihood of developing heart disease. In the traditional risk assessment, factors including age, gender, blood pressure, cholesterol, and smoking history are included. This reduces the need for in-person medical visits, enhancing patient convenience and enabling healthcare professionals to monitor a larger number of patients effectively.

Challenges such as data security, interoperability of IoT devices, and model interpretability need to be addressed for widespread adoption of ML-powered IoT solutions for heart disease prediction.

The development and application of a heart disease prediction system that uses the Random Forest algorithm and boasts a remarkable accuracy rate of 95.2% is a significant advance in both healthcare and data science. Heart disease is still a major global health concern, highlighting the crucial role that accurate prediction models play in enabling early diagnosis and prompt care. Achieving such a high level of accuracy is an outstanding accomplishment with significant benefits for improving patient care and maximizing health outcomes.

2.1 Role for Enhancing Heart Disease Prediction System

The identification and prevention of heart disease could be revolutionized by improving a method for predicting the condition using IoT (Internet of Things) and machine learning. Continuous real-time data collecting from IoT devices, such as wearable fitness trackers and smartwatches, is the first step in the process. These devices keep track of vital signs like blood pressure, heart rate, ECG data, and lifestyle variables and send that data to a centralized system or cloud-based platform for analysis. The processing and analysis of this vast amount of data, which involves finding patterns, anomalies, and trends that can be suggestive of risk factors for heart disease, is subsequently greatly aided by machine learning algorithm

With access to both history and current data, machine learning models can estimate a person's risk of getting heart disease by taking into consideration their age, gender, genetics, lifestyle, and current health data. As a result, the system is able to provide early detection because it can detect small changes in a person's health status that might not be noticeable during regular checkups. For effective management and prevention, early detection is crucial.

Additionally, machine learning can offer individualized health advice by adapting food recommendations, exercise schedules, and sleep hygiene tips to a person's particular risk profile. When it notices potentially harmful changes in health metrics, the system can send real-time alerts to patients and healthcare professionals, enabling prompt medical assistance when required.

In addition to managing individual health, this system allows remote monitoring, which minimizes the need for frequent in-person visits, and also makes it easier to create longitudinal health records for each person. Insights into the causes and development of cardiac disease can be gained from the collected data, which can also help with the creation of more efficient interventions and therapies.

Aggregated data from a network of IoT devices can enhance population health management on a larger scale by assisting public health professionals and policymakers in identifying trends and heart disease hotspots within particular populations. Programs for targeted intervention and prevention are then based on this information.

IoT and machine learning can be used to improve a cardiac disease prediction system, providing a proactive, individualized approach to healthcare. It provides early identification, individualized care, and the potential to enhance overall heart health outcomes while lowering the financial and personal toll of heart disease. When adopting such a system, it is essential to protect data privacy, acquire informed consent, and adhere to pertinent healthcare standards.

2.2 Key Concepts for Enhancing Heart Disease Prediction System

The term "Internet of Things" (IoT) describes a network of linked objects that share data online. Wearable sensors, medical equipment, and health monitoring tools are examples of IoT devices in this context that are used to continuously collect data on a person's health and way of life.

Data collection: Information is gathered through IoT devices that track numerous health indicators, including blood pressure, heart rate, ECG data, physical activity, and sleep patterns. Then, a central system receives this data for analysis.

Machine learning: Machine learning is the process of analyzing data, finding patterns, and making predictions using algorithms and statistical models. Machine learning is utilized to evaluate and understand the health data gathered from IoT devices in the context of heart disease prediction.

Early detection: Machine learning algorithms can spot minor changes in a person's health that could be a sign of heart disease in its early stages, allowing for prompt intervention and prevention.

Remote monitoring is made possible by IoT devices, which eliminates the need for frequent in-person checkups. To make wise decisions, healthcare providers have access to real-time patient data.

2.3 The Need for Enhancing Heart Disease Prediction System

A compelling need for cardiac disease prediction in healthcare has arisen as a result of the convergence of machine learning (ML) and the Internet of Things (IoT). The most notable benefit of this synergy is early identification, one of many others. The identification of minute changes in heart health is made possible by the careful analysis of patient data by ML algorithms using IoT devices like wearable sensors and smartwatches. The facilitation of prompt interventions and, ultimately, the prevention of heart disease depend critically on this early detection.

Additionally, the concept of individualized treatment is strengthened by the union of ML and IoT. IoT-generated data allows machine learning (ML) to create treatment plans and suggestions that are specifically suited to each patient's need. This personalization enhances patient care, leading to better outcomes and a higher standard of living.

Another essential component of heart disease prediction using ML and IoT is remote monitoring. There is no longer a need for regular in-person visits thanks to these technologies, which enable continuous and non-intrusive monitoring of patients' heart health. This not only increases patient convenience but also guarantees healthcare professionals have immediate access to vital data, allowing them to react quickly to any alarming developments.

Another essential component of heart disease prediction using ML and IoT is remote monitoring. There is no longer a need for regular in-person visits thanks to these technologies, which enable continuous and non-intrusive monitoring of patients' heart health. This not only increases patient convenience but also guarantees healthcare professionals have immediate access to vital data, allowing them to react quickly to any alarming developments.

Another strong argument in favor of using ML and IoT for heart disease prediction is their cost-effectiveness. Healthcare expenditures can be considerably decreased by using predictive models to avoid complications and hospitalizations. This is advantageous not only for individuals but also for healthcare systems and insurers trying to control costs.

Promoting healthy lifestyle choices is essential to preventing heart disease. Real-time feedback from ML and IoT can inspire people to make better decisions about nutrition, exercise, and other lifestyle choices that have a big impact on heart health. This proactive involvement may completely alter the way heart disease is treated and prevented.

Most importantly, applying ML and IoT to the prediction of cardiac disease can help lower mortality rates. Heart attacks and other cardiac events can be avoided with early detection and treatment, saving lives and reducing the strain on healthcare systems.
Finally, ML and IoT fusion can improve public health management. Public health organizations are able to identify high-risk locations and carry out focused interventions by combining data from IoT devices, reducing the burden of heart disease on communities.

In conclusion, it is critical to anticipate cardiac disease utilizing ML and IoT. It can raise the standard of care, encourage prevention, lower healthcare costs, and—most importantly—improve the general health of those who are at risk for heart disease. The discipline of cardiology could undergo a revolution thanks to this potent combination, which would also be good for society as a whole.

2.4 Challenges in the Existing System

The integration of machine learning (ML) and the Internet of Things (IoT) into existing systems for heart disease prediction offers tremendous potential, but it also comes with several challenges:

1. Data Privacy and Security: IoT devices raise serious issues regarding data privacy and security when it comes to collecting and transferring sensitive health data. A primary responsibility is making sure that patient information is secure against hacks and illegal access.

2. Data Quality and Standardization: IoT devices produce a lot of data, but its quality and standardization can vary. The performance of ML algorithms can be hampered by variations in data completeness, correctness, and format.

3. Interoperability: The adoption of numerous IoT platforms and devices in the healthcare industry can cause interoperability problems. It might be challenging to smoothly integrate various tools and platforms for data collecting and analysis.

4. Interpretability of the model: ML models for heart disease prediction might be complicated, making it difficult to interpret the results. Models that are comprehensible and offer explanations for why a specific prediction was made are essential in the healthcare industry.

5. Data Bias: ML models that employ training data that contains biases may produce predictions that are biased. The model might not be appropriate for producing predictions for a larger, more varied group of patients if the data is mostly from a single demography or population.

6. Severe Regulatory Compliance: The use of IoT devices and ML in healthcare is subject to compliance with severe regulations, such as HIPAA in the United States. It can be difficult and time-consuming to create and use prediction systems while adhering to these rules.

7. Ethical Issues: ML algorithms may unintentionally exacerbate current prejudices and inequities in healthcare. A crucial difficulty is ensuring that predictive systems are created and used ethically.

8. Resource Constraints: ML model development and infrastructure upkeep can be resource-intensive processes. Smaller healthcare organizations could find it difficult to invest in the required knowledge and equipment.

9. Human Expertise: To evaluate and act on the predictions made by ML algorithms, healthcare workers may need additional training. It is crucial to close the technological and clinical practice gap.

10. Scalability and Cost: It can be expensive to apply ML and IoT to forecast heart disease across healthcare systems. The cost of equipment, infrastructure, and knowledge might prevent these technologies from being widely used.

11. Patient Engagement: It can be difficult to get patients involved in their care and make sure they actively use IoT devices for data collecting. Effective teaching, inspiration, and support are necessary.

Addressing these challenges is crucial to harness the full potential of ML and IoT in heart disease prediction while ensuring patient safety, data privacy, and equitable healthcare outcomes.

Dataset:

The UC Irvine ML Repository-Cleveland dataset, which contains 303 instances and 14 attributes, is included in this dataset.

Features of the dataset:

- **Age:** Patient age.
- **Sex:** For males indicates 1, females 0.
- **Chest pain type:** For angina - (0), atypical angina – (1), non-anginal pain – (2), asymptomatic – (3).
- **Resting blood pressure:** Resting blood pressure upon hospital admission. Measured in mm/Hg.
- **Serum Cholesterol:** Blood cholesterol level measured in mg/dL.
- **Fasting blood sugar:** If the blood sugar level is over 120 mg/dL after a fast of not eating overnight, it is considered to be high (1—true). If it is below 120 mg/dL, it is considered to be normal (0—false).
- **Resting ECG:** An ECG test result can be categorized as follows: 0 for a normal result, 1 for the presence of ST-T wave abnormality, and 2 for left ventricular hypertrophy.
- **Maximum heart rate:** Maximum heart rate during exercise.
- **Exercise angina**: Angina occurred by a workout, 0 for no; 1 for yes
- **Old peak**: ST depression due to exercise relative to relaxation will observe in the ECG test.
- **ST slope:** Maximum workout 1 for upsloping; 2 for flat; 3 for down sloping
- **Ca:** The number of major blood vessels that can be visualized using fluoroscopy can range from 0 to 3.
- **Thal:** Thalassemia is a blood disorder caused by abnormal hemoglobin production, with a score of 3 indicating normal production, 6 indicating permanent deficiency, and 7 signifying temporary impairment.
- **Target:** No heart disease—0; heart disease—1.

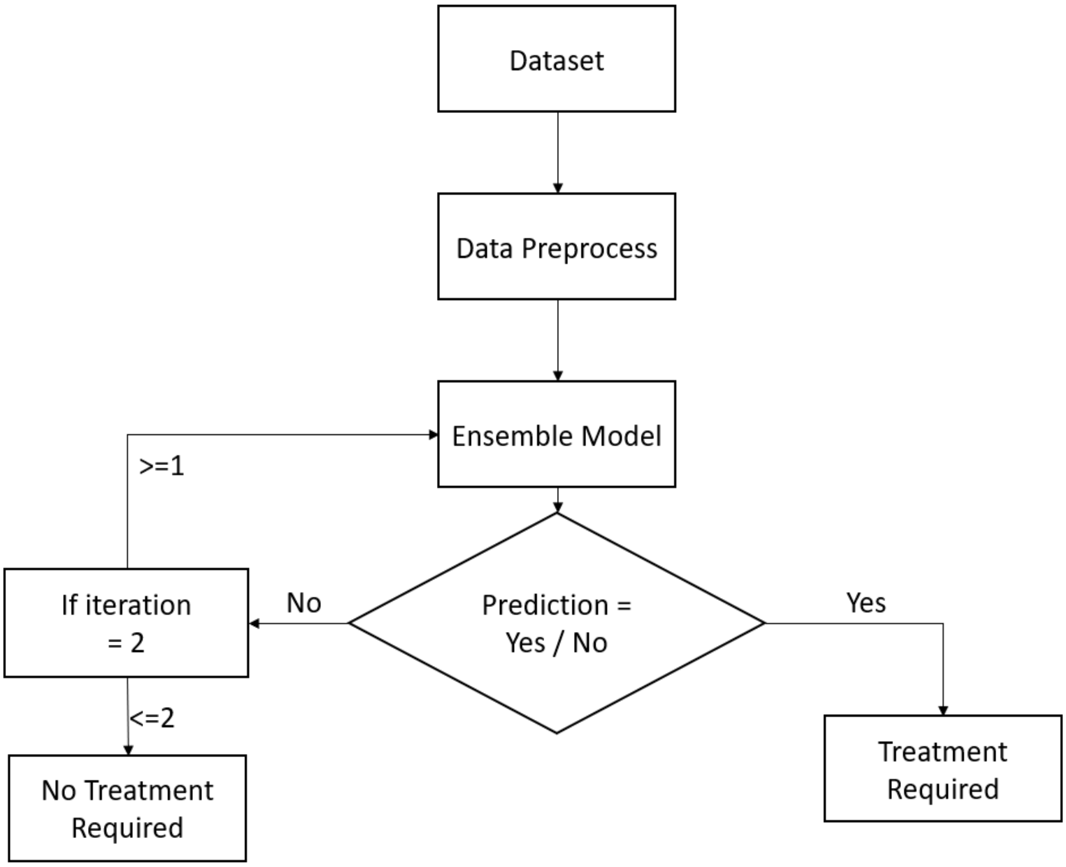


Figure 1: Work Flow of Prediction System

The workflow of the prediction system in the image can be summarized as follows:

1. The user enters the desired data. Data about patients who have heart disease is collected from medical records, surveys, or other sources.
2. The data is pre-processed to remove noise, outliers and missing values.
3. The data is split into training and testing sets.
4. A machine learning algorithm is trained on the training set.
5. The trained model is used to make predictions on the testing set.
6. The accuracy of the model is evaluated.
7. The model is deployed and used to make predictions on new data.

The Fig.1 shows a specific example of a prediction system that is used to predict the treatment required for a patient. The system takes as input the patient's medical records, and then uses an ensemble model to make a prediction about the best treatment. The ensemble model is a combination of multiple machine learning algorithms, which helps to improve the accuracy of the predictions.

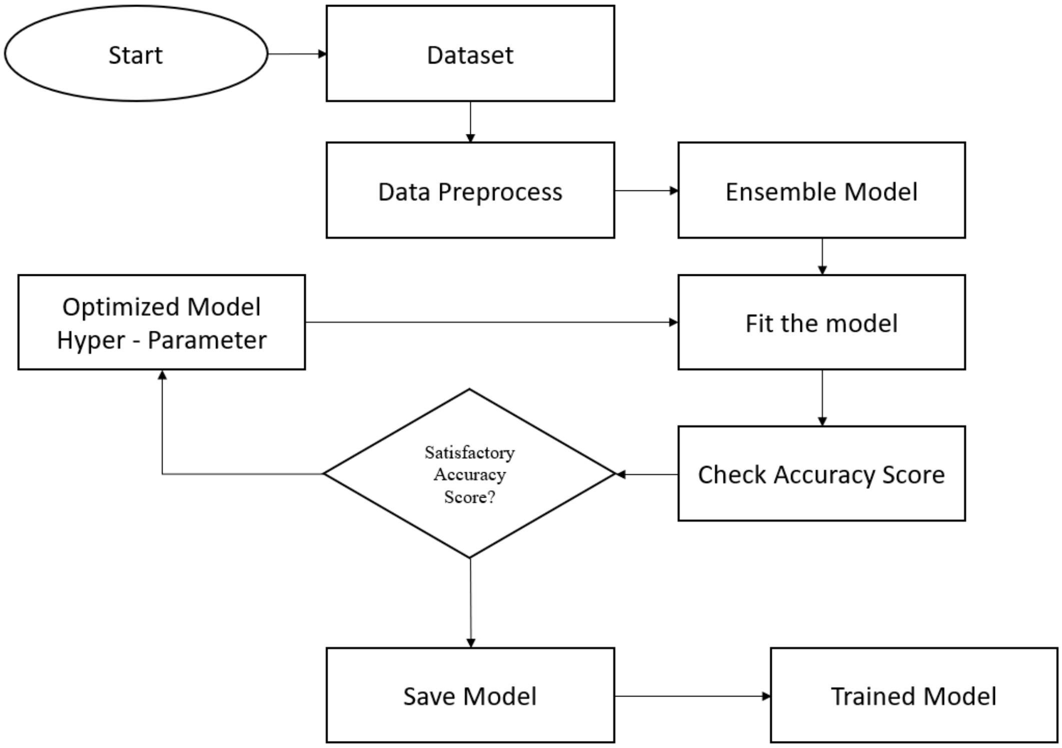


Figure 2: Work flow of Model

IoT Device Integration:

- Gathering thorough health data, create protocols for combining different IoT devices (wearables, smart scales, blood pressure monitors).
- Implementing a secure device management and data transmission.
- Determining pertinent characteristics from IoT data, such as sleep patterns, blood pressure trends, and heart rate variability.
- Creating new features that captures complex health data.

Ensemble Learning:

- Using ensemble methods to aggregate predictions from various machine learning models, such as random forests or gradient boosting.
- Examining the possibility for increased prediction accuracy.
- Creating a system for regularly updating machine learning models with fresh information to boost prediction precision.

Health Risk Communication:

- Developing efficient means of informing users of their risk for heart disease and offering them useful information for enhancing their cardiovascular health.
- Integrating fairness and bias reduction as well as ethical issues into the system's design and implementation.

Remote Monitoring Integration:

- Investigating telehealth system integration to allow medical personnel to remotely monitor patients' heart health.

Interoperability with Electronic Health Records:

- Creating interfaces for the electronic health record systems used by healthcare providers to seamlessly integrate system predictions and user data.

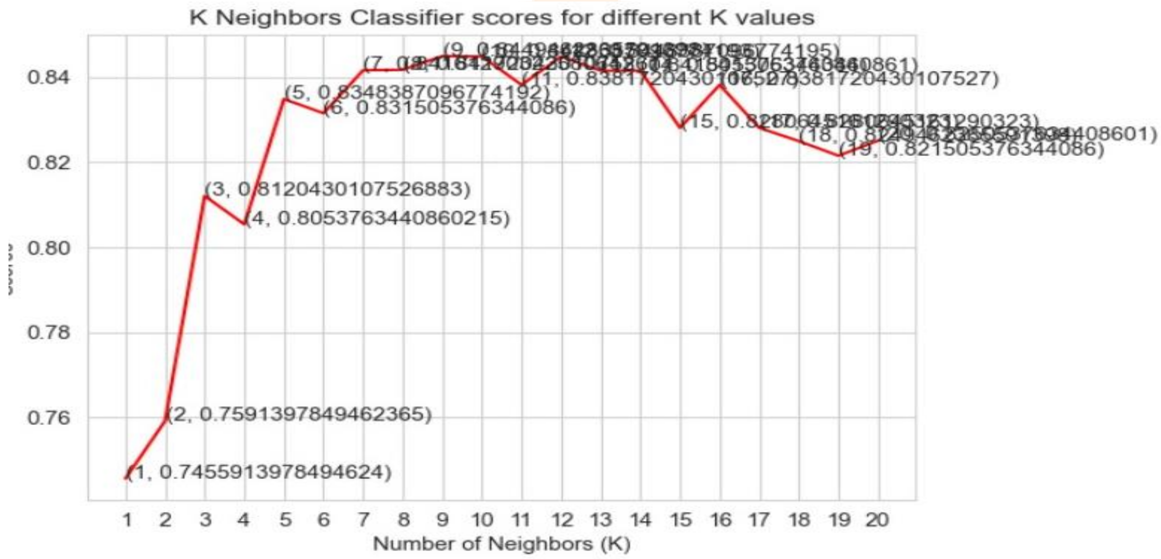
Predictive Model Uncertainty Estimation:

- Implementing methods to calculate the degree of uncertainty in model forecasts, giving users a confidence score for every forecast.

Classification Methods:

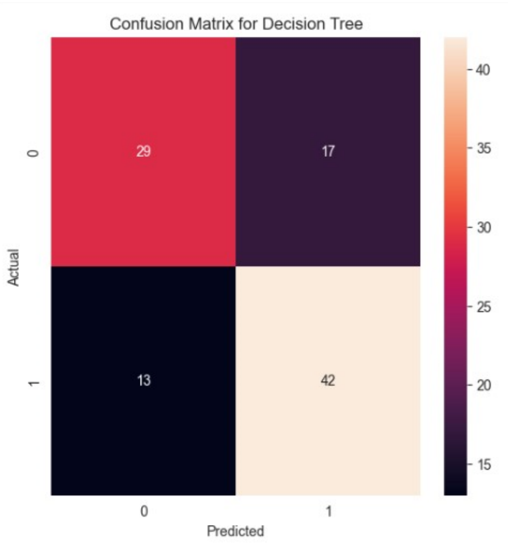
1. **Random Forest Classifier:**
   By averaging the forecasts of their real trees, the RF classifier generates predictions. RF is a supervised machine learning technique built on ensemble learning. It combines numerous decision trees using bagging to increase prediction accuracy. Every person receives individualized bagging instruction. Each decision tree is tested during the training process using various data samples that were produced at random using replacements from the original dataset. A random selection of features is also made when building trees. The combined projections from different trees are combined by a majority vote.
   Model Accuracy: 95.2%

2. **KNN Classifier:**
   KNN is a method of lazy learning or instance-based learning. The method of developing a model without the need for training data is referred to as lazy learning. they are chosen from a collection of objects with predetermined traits or classifications. It first determines the k data points that are most similar to the new data point in order to forecast the label of the new data point. The user must choose the k value as a hyperparameter. while the k number is larger, more data points are considered while creating predictions, which could lead to slower performance but also more precise forecasts. As the k value drops, fewer data points are considered, which might lead to faster processing but less accurate projections.
   Model accuracy: 84%



3. **Decision Tree:**
   It is a supervised machine learning approach known as a decision tree can be applied to both classification and regression applications. It functions by building a decision tree-like structure, with each decision node standing in for a data feature and each leaf node for a class label. The feature that best divides the data into two pure subsets is chosen for the splitting process. A straightforward and understandable algorithm, decision trees are simple to comprehend and use. They are also incredibly adaptable and useful for a wide range of jobs. They may, however, be susceptible to data noise and overfit the training set.

Model accuracy: 91%



ALGORITHM

Algorithm: IoT Device Initialization, Data Storage, and Disease Prediction

Step 1: Attach the IoT medical sensors to the patient's body and start Device Initialization.

Step 2: Configure the NodeMCU with the appropriate SSID and password for the Wi-Fi network, then connect it to the Wi-Fi.

Step 3: Attach the NodeMCU with the sensor to act as a mediator.

Step 4: Check if the NodeMCU is connected to Wi-Fi:

- If (NodeMCU is connected):
  - Check if the sensor is properly attached to the NodeMCU:
    - If (sensor is attached):
      - Read data from the sensor.
      - Send the collected data to the cloud.
    - Else:
      - Display an error message: "Sensor is not attached."
  - End
- Else:
  - Display an error message: "NodeMCU is not connected."
  - End

Step 5: Initialize PostgreSQL in the Heroku cloud.

Step 6: Set up PostgreSQL with a username and password.

Step 7: Check if the database and table have been created:

- If (database and table do not exist):

  - Create the necessary database and table.

- End

Step 8: Configure IoT sensor data using ThingSpeak's write API key to insert.

Step 9: Check if there is data related to disease information (thingspeakDiseaseData) or patient personal data (PatientPersonalData):

- If (thingspeakDiseaseData == True or PatientPersonalData == True):

  - Insert data into the table in the PostgreSQL database.

- Else:

  - Display an error message: "No data to insert."

  - End

Step 10: Select the patient for whom disease prediction is needed.

Step 11: Check the patient data and its attributes:

- If (Any Attribute == null):

  - Insert the required data into the database.

- End

Step 12: Predict the disease using the proposed model based on the available patient data.

Step 13: Display the disease prediction result.

RESULTS

After performing the three algorithms on the given data-set the highest accuracy was achieved using the Random Forest Classifier with an accuracy of 95.2% followed by Decision tree which gave an accuracy of 91% and then KNN giving an accuracy of 84.2%.

The accuracy, precision and recall was calculated for all the three algorithms used and the highest was achieved using the Random Forest Classifier.

Accuracy = (TP + TN) / (TP + TN + FP + FN) = (237 + 12) / (237 + 12 + 6 + 6) = 0.952 or 95.2%

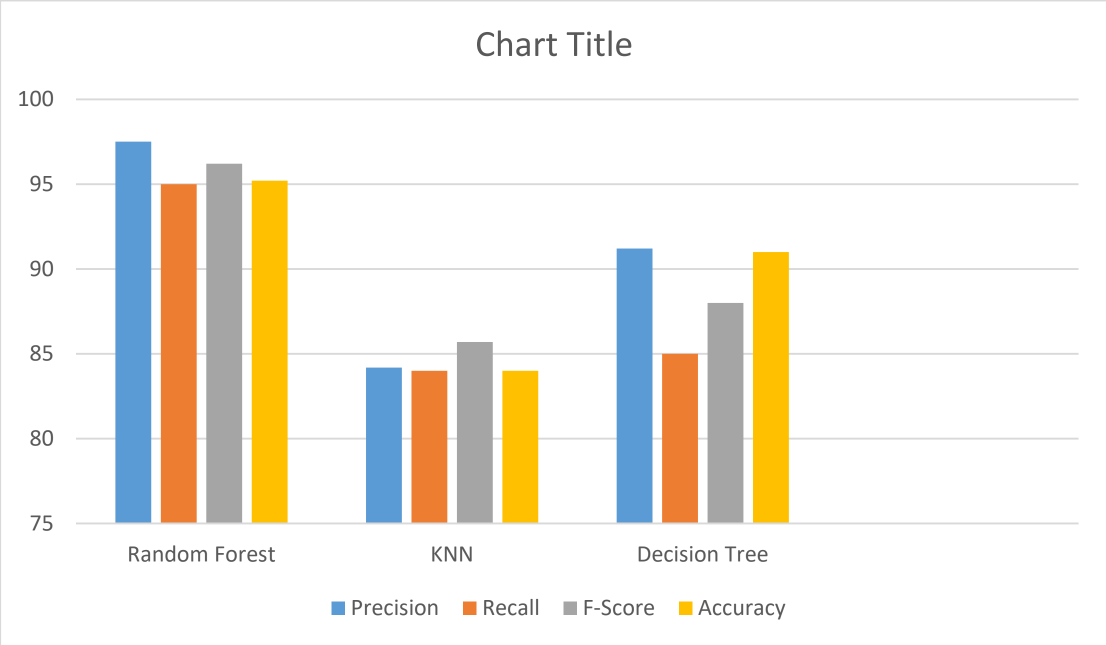
Precision = TP / (TP + FP) = 237 / (237 + 6) = 0.974 or 97.4%

Recall = TP / (TP + FN) = 237 / (237 + 6) = 0.974 or 97.4%

S.No.	Algorithm	Precision	Recall	F-Score	Accuracy
1.	Random Forest	97.5	95	96.2	95.2
2.	KNN	84.2	84	85.7	84
3.	Decision Tree	91.2	85	88	91

TABLE OF VALUES

Graph Measurement



BUDGET

Table 1: Proposed Budget

Sl. No	Name of the Required System	Cost in Rs. (Approx)
1.	Server Hosting	1900.00
2.	Database	1200.00
3.	Application Software	350.00
4.	Server Space Costing	2000.00
5.	Smart Watch	3000.00
	TOTAL	8450.00

CONCLUSION

In conclusion, this research presents a promising strategy has been the combination of machine learning (ML) algorithms and internet of things (IoT) technology. With the help of IoT data and three well-known ML algorithms, Decision Trees, K-Nearest Neighbors (KNN), and Random Forest, this study aimed to improve heart disease prediction. The study's main accomplishment was a remarkable 95.2% accuracy rate using the Random Forest method. It was used as the main machine learning algorithm, prediction accuracy increased significantly when compared to the other algorithms. This demonstrates how well ensemble learning techniques work with challenging medical datasets. It also offers information on feature importance, which can help medical professionals comprehend the elements that have the greatest impact on heart disease. Future investigations and clinical judgments can be informed by this information. IoT and ML integration, together with the Random Forest algorithm's 95.2% accuracy rate, marks a substantial leap in the prediction of cardiac disease. By enabling early detection and intervention, this technology has the potential to improve healthcare and ultimately save lives. However, prior to the system's widespread implementation in clinical practice, it is essential to address issues and guarantee its dependability and security.

REFERENCES

[1] A Angel Nancy, Dakshanamoorthy Ravindran, P M Durai Raj Vincent, Kathiravan Srinivasan and Daniel Gutierrez Reina, IoT-Cloud-Based Smart Healthcare Monitoring System for Heart Disease Prediction via Deep Learning (2022) EISSN 2079-9292, Published by MDPI

[2] Raniya R. Sarra, Ahmed M. Dinar, Mazin Abed Mohammed and Karrar Hameed Abdulkareem, Enhanced Heart Disease Prediction Based on Machine Learning and $\chi 2$ Statistical Optimal Feature Selection Model (2022) EISSN 2411-9660, Published by MDPI

[3] Ahmad F. Subahi, Osamah Ibrahim Khalaf, Youseef Alotaibi, Rajesh Natarajan, Natesh Mahadev and Timmarasu Ramesh, Modified Self-Adaptive Bayesian Algorithm for Smart Heart Disease Prediction in IoT System (2022) EISSN 2071-1050, Published by MDPI

[4] Jafar Abdollahi, Babak Nouri-Moghaddam, A hybrid method for heart disease diagnosis utilizing feature selection-based ensemble classifier model generation (2022)
Iran J CS *Sci* **5**, 229–246, Published by Springer

[5] Ahamed, Jameel; Manan Koli, Abdul; Ahmad, Khaleel; Alam Jamal, Mohd.; Gupta, B. B., CDPS-IoT: Cardiovascular Disease
Prediction System Based on IoT using Machine Learning (2022) Vol. 7, Published by Reunir

[6] Md. Manjurul Ahsan, Zahed Siddique, Machine learning-based heart disease diagnosis: A systematic literature review (2022), Vol. 128 Article 102289 Published by Elsevier

[7] Pierre Claver Bizimana, Zuping Zhang, Muhammad Asim and Ahmed A. Abd El- Latif, An Effective Machine Learning-Based Model for an Early Heart Disease Prediction (2023) Article 3531420 Published by Hindawi

[8] Nadikatla Chandrasekhar and Samineni Peddakrishna, Enhancing Heart Disease Prediction Accuracy through Machine Learning Techniques and Optimization (2023), Vol. 11, Published by MDPI

[9] Muhammed Golec, Sukhpal Singh Gill, Ajith Kumar Parlikad and Steve Uhlig HealthFaaS: AI based Smart Healthcare System for Heart Patients using Serverless Computing (2023), Published by IEEE

[10] Federico Del Giorgio Solfa, Fernando Rogelio Simonato, Big Data Analytics in Healthcare: Exploring the Role of Machine Learning in Predicting Patient Outcomes and Improving Healthcare Delivery (2023), Published by IJCIM

[11] Nouf Abdullah Almujally , Turki Aljrees, Oumaima Saidani, Muhammad Umer, Zaid Bin Faheem, Nihal Abuzinadah, Khaled Alnowaiser and Imran Ashraf, Monitoring Acute Heart Failure Patients Using Internet-of-Things-Based Smart Monitoring System (2023) Published by MDPI

[12] Pronab Ghosh, Sami Azam, Mirjam Jonkman, Asif Karim, F. M. Javed Mehedi Shamrat, Eva Ignatious, Shahna Shultana, Abhijith Reddy Beeravolu and Friso De Boer, Efficient Prediction of Cardiovascular Disease Using Machine Learning Algorithms With Relief and LASSO Feature Selection Techniques (2021) Published by IEEE

[13] Bzhar Ghafour Mohammed, Dler Salih Hasan, Smart Healthcare Monitoring System Using IoT, (2023), Published by ResearchGate

[14] Sangeethalakshmi K., Preethi Angel S., Preethi U., Pavithra S., Shanmuga Priya V., Patient health monitoring system using IoT (2021), Published by Elsevier

[15] Abderahman Rejeb, Karim Rejeb, Horst Treiblmaier, Andrea Appolloni, Salem Alghamdi, Yaser Alhasawi , Mohammad Iranmanesh, The Internet of Things (IoT) in healthcare: Taking stock and moving forward (2023), Published by Elsevier