International Conference on Innovative Data Communication Technologies and Application (ICIDCA-2023)
IEEE Xplore Part Number: CFP23CR5-ART; ISBN: 979-8-3503-9720-8

Machine Learning based Mobile App for Heart Disease Prediction

S. Dwijesh Reddy1
Information Technology Student
VR Siddhartha Engineering College
Vijayawada, India
dwijeshreddy27@gmail.com

S. Lohitha2
Information Technology Student
VR Siddhartha Engineering College
Vijayawada, India
singamanenilohitha@gmail.com

Fathimabi Shaik3 Information Technology Assistant Professor
VR Siddhartha Engineering College
Vijayawada, India
fathimabi@vrsiddhartha.ac.in

2023 International Conference on Innovative Data Communication Technologies and Application (ICIDCA) | 979-8-3503-9720-8/23/$31.00 ©2023 IEEE | DOI: 10.1109/ICIDCA56705.2023.10099714

Abstract: The world's leading cause of death is heart disease. A variety of modern technologies are utilized to treat cardiac disease The most common problem in medical centers around the world is that many medical personnel lack equal knowledge and courage to treat their patients, so they develop their own opinions, which leads in bad outcomes and, in some cases, death. Predictions of cardiac illness are employed to overcome these issues. This study has used various criteria to predict cardiac disease. These characteristics are Age, Gender, Cerebral Palsy (CP), Blood Pressure (BP), Fasting blood sugar test (FBS), and so on. The major goal of the research is to create a mobile app that reduces the cost of medical tests while also avoiding human bias. The outcome of the research is to forecast cardiac disease. The research made advantage of the built-in dataset and used PHR data to make predictions. Machine Learning is being used to build the model. This study utilizes a variety of machine learning algorithms, including Logistic Regression, ANN Multi-Layer Perceptron (MLP), and Random Forest (RF). Random Forest (RF) outperforms the other two algorithms in terms of accuracy. As a result, this study employs random forest to forecast heart healthand builds the mobile app with MIT App Inventor and stores the data in the Firebase database. The App could be to maintain personal health records and share our info with doctors. It will forecast heart health when you enter the criteria.

Keywords: PHR, MIT App Inventor, Firebase database, Logistic Regression, ANNMulti layer Perceptron, Random Forest.

1. Introduction

According to the World Health Organization, heart disease kills 12 million people each year. The worldwide incidence of coronary heart disease has now been rapidly increasing in recent years. Many research has been conducted in an attempt to determine the most important risk factors for developing heart disease and to quantify the riskiness. [Early identification of heart disease is crucial in making lifestyle changes decisions in high- risk patients, avoiding consequences]. The purpose of this research is to use machine-learning algorithms to forecast future heart disease by reviewing patient data and categorizing whether they have heart disease or not.

Machine learning has been demonstrated to be effective in assisting with decision making and prediction from the massive amounts of data generated by the health care industry. This study uses a machine-learning system to examine patient data and classify whether people have heart disease or not in order to predict future heart disease. In this aspect, machine learning approaches can be quite beneficial. [Heart disease remains the major cause of death globally, killing more people than any other cause each year. Heart disease killed an estimated 17.3 million individuals in 2008, responsible for 30% of all deaths worldwide]. Coronary heart disease was responsible for an estimated 7.3 million of these deaths, whilestroke was responsible for 6.2 million. Latest clinical research has revealed risk factors that may cause the onset of heart disease, but more research is required before this information can be used to reduce the occurrence of heart problems. Diabetes, hypertension, and high cholesterol levels have all been established as significant risk factors for cardiovascular disease. Foodpreferences, lack of physical activity, cigarettes, alcohol intake, and overweight are all linked to major risk factors for heart disease. According to research, lowering these risk factors for cardiovascular disease can significantly assist in avoiding heart disease. Numerous studies and researches have been conducted on the dangers of heart disease. The use of data from population studies to forecast cardiovascular disease related to blood pressure, smoking habits, cholesterol and blood pressure levels, and diabetes has benefited in the prediction of heart disease.

CVDs kill more people than any other cause each year, with an estimated 12 million people dying from heart disease each year. In the United States, one person dies from heart disease every 34 seconds. Heart attacks are sometimes devastating events that occur as a result of a blockage of blood flow to the heart or brain. People at risk of heart disease may have high blood pressure, glucose, and cholesterol levels, as well as be stressed. Basic health facilities can readily measure all of these factors at home. Many diseases, disabilities, and deaths are caused by cardiovascular disease (CVD). Disease diagnosis is an important and challenging task in medicine. The automation of this task is quite beneficial. Unfortunately, not all physicians are [subject specialists], and certain areas have a scarcity of resources. Data mining can be used to uncover hidden patterns and knowledge that can help with decision making. This helps healthcare professionals make accurate decisions and provide quality services to the public. The approach supplied by the health care organization to professionals who do not have more information and skills is also very crucial. The capacity to draw precise conclusions as needed is one of the key constraints of existing approaches. The study made use of the built-in data set and the PHR data to make predictions.

Despite the possibility that cardiovascular disease manifests itself in a number of ways, there is a common set of important risk factors that impact whether a person is at risk of a heart attack or not. This technique is well suited to predicting cardiac disease since it collects data from multiple sources, categorizes it under relevant topics.

We use a variety of machine learning algorithms, including Logistic Regression, ANN Multi-Layer Perceptron, and Random Forest. Random forest has more accuracy. As a result, we are employing the random forestto forecast cardiac disease. We're building the mobile app with MIT App Inventor and storing the data in the Firebase database. We can use the App to maintain personal health records and share our info with doctors.It will forecast heart health when you enter the criteria.

2. Preliminaries

The basic terms and concepts used in this paper are explained in this section.

2.1. PHR

Personal Health Records serve as a repository for storing and maintaining medical records. It is a default page that is later modified to meet our needs. This is where they can keep their personal records.

2.2. NGROK

ngrok is a reverse proxy that builds a secure tunnel between a public endpoint and a web service operating locally. ngrok records and analyses every tunnel traffic for subsequent review and playback. ngrok is a tool in a tech stack's local host Tool. In a nutshell, it gives a publicly available web URL to any locally hosted application, such as a Spring boot or Nodejs- based application or a web hook for a chat application, and so on.

2.3. Fire Base Database

The Firebase Real Time Database is a database that's hosted in the pall. Data is kept in JSON format and is synchronized in real time to all connected guests. Whenyou produce cross-platform apps with our Apple, Android, and JavaScript SDKs, all of your guests partake a single real-time Database case and are automatically streamlined with the most recent data.

2.4. MIT App Inventor

MIT App Inventor is a web- based app development tool for Android. [This is generally appertained to as visual programming, which indicates that the stoner can execute programming tasks without having to enter any computer law]. MIT's MobileLearning Lab laboriously manages and develops App Inventor (the design was firstly erected by Google).

2.5. Machine Learning

Machine learning is an area of artificial intelligence (AI) and computational intelligence that focuses on the use of data and algorithms to mimic how humans learn, gradually improving its accuracy.

Machine learning is a critical component in the expanding field of data [wisdom]. Algorithms are trained using statistical techniques to form groups or forecasts and to identify critical perceptivity in data mining systems. These perceptions, in turn, affect decision-making acrossoperations and organizations, immaculately influencing critical growth indicators. It will forecast heart health using Machine Learning methods.

3. Literature Study

Yu- Xuan Wang etal. examined a variety of operations that demonstrated the utility of data mining and machine learning in numerous areas [1]. Since the maturity of the work was done on system characteristics that changed on a regular base, a study on the association designs of colorful bias was proposed. The structure's functionality was altered exercising the proposed system design fashion. A fresh fashion for designing an operating system was offered in this work. The proposed system made use of data mining and machine learning ways. The whole data acquired from the structure was anatomized after entering a response from a data miner. Grounded onthe results of the testing, it was determined that the proposed strategy produced effective results.

ZhiqiangGe etal presented a exploration paper on former data mining and analytics operations. These approaches were used in business for a variety of purposes. For data mining and analytics exploration, eight unsupervised and ten supervised learning algorithms were used [2]. The semi- supervised learning styles were given an operation status in this study. Unsupervised and supervised machine learning wayswere used in roughly 90- 95 of assiduity procedures. lately, semi-supervised machine learning has been proposed. therefore, it was demonstrated that data mining and judges played an important part in the development of new machine learning algorithms in assiduity- related operations.

BayuAdhi Tama et al. (2016) proposed a research into the chronic disease known as diabetes. Thiscomplaint was thought to be quite widespread and caused major repercussions. According to a check conducted by the International Diabetes Federation IDF)[3,] around 285 million people worldwide suffer from diabetes. These figures may rise in the near future as a result of the implementation of a suitable approach for the complete minimization and prevention of this disease. Type 2 diabetes was the most common type. The key issue was the identification of TTD as the vaccination of all its effects was a difficult task. As a result, data mining was used because it handled the best issues and aided in the identification of information from data available.

P. Suresh Kumar etal. developed a new approach for addressing multitudinous issues in clustering and bracket approaches in data mining infrastructures. This approach was used to determine the type of diabetes.

The information gathered was used to calculate each case's fitness position. This disease leads to a number of consequences [4].In this study, all attained information about 650 cases was used for exploration purposes, and its goods were also honored. The bracket system used this clustered data package as input. This categorization fashion further divided information into several classes, similar as mild, moderate, and severe diabetes threat situations in cases.

Han Wu et al. (2018) suggested a new model for predicting diabetes mellitus type 2 based on data mining procedures (T2DM). The main goal of this study was to improve the correctness of the [vaticination] model. The other end of this investigation was the development of an adaptive data suite model. The proposed model consisted of two corridors, which were based on the sequence of preprocessing methods [5]. These two sections were coupled as the logistic retrogression algorithm and the augmented K-means algorithm. The Pima Indians Diabetes Dataset and the Waikato Environment, coupled with the Knowledge Analysis toolbox, were used to compare the issues of new and existing methods. The relative issues demonstrated that the proposed model demonstrated superior perfection in comparison to other methods while also providing adequate dataset quality. The suggested model's performance was estimated byrunning it on many other diabetic datasets. It was discovered that both ways performed well.

Roehrs et al. [6] attempted to define the taxonomy and open questions in order to investigate the most recent research on PHRs. Furthermore, this study tried to clearly identify data kinds, standards, profiles, goals, methodologies, functionalities, and architecture with reference toPHRs. Regarding coverage and localization,they gave a comparison of various architecture types and PHR implementation styles. Last but not least, the difficulties and problems relating to security, privacy, and trust are described and are expected to be the subject of future studies.

Sumithra et al. [7] developed an innovative method for restricting access to PHRs, and the RSA algorithm is utilized to encrypt the PHR files for each patient. For the safe exchange of individual health records in the cloud, a novel structure was put forth. To fully actualize the patient-centric idea, partially trustworthy cloud servers are taken into consideration and discussed. Encrypting their PHR data will provide patients total control over their own privacy and allow for fine-grained access. In comparison to other research,the framework greatly decreases the complexity of key management while improving privacy guarantees.

It accomplishes this by solving the specific problems raised by numerous PHR users and owners. In order to provide access from both personal users and a variety of users from the public domains with a range of professional positions, qualifications, and affiliations, thePHR data is encrypted using RSA.

4. Architecture Diagram

Step 1: First the user will register into the mobile app for hearthealth.
Step 2: After registration he/she will login into the mobile app for heart health.
Step 3: After login in that screen, we havePredict option, Previous Reports and Logout, About us options.
Step 4: When we select the predict, it will redirect intoanother screen. In that screen he/she has to enter the parameters.
Step 5: Based on the parameters it will predict the heartdisease and store the values in the database.
Step 6: When we select previous reports, It will redirectinto another screen. In that screen he/she can add their personal health records.
Step 7: And they can delete the records which are not important.
Step 8: And also they can share their reports to doctors or family members through whatsapp, mail, etc..
Step 9: When the user clicks the logout the screen will redirectinto the login screen.

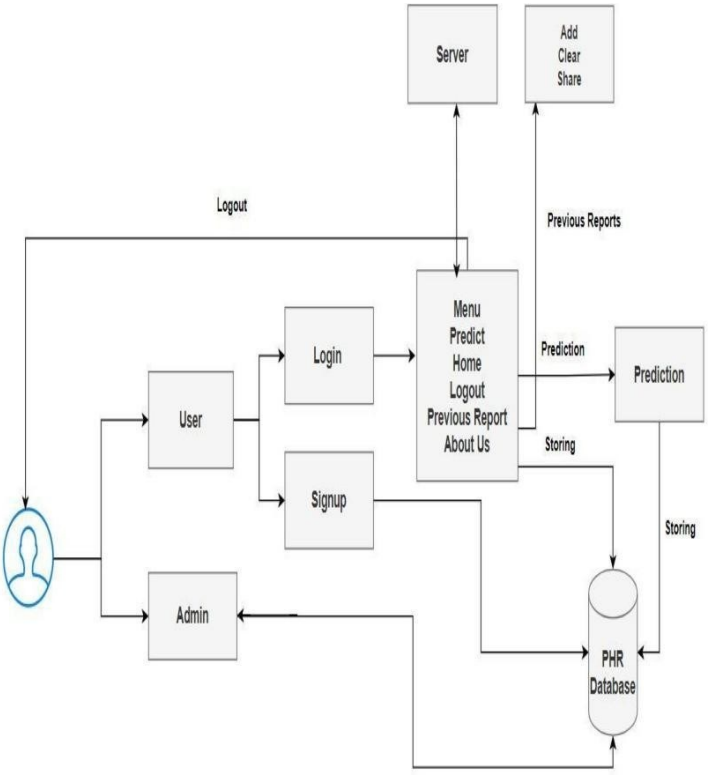

Figure 1. System Architecture

5. Design Methodology

Random Forest (RF) is a machine learning algorithm that is supervised. It is, as the name implies, a forest of randomly produced decision trees. It basically employs a method bagging, in which different learning models are blended to improve overall outcomes. It generates numerous decision trees and combines them to create a refined output when performing the bagging procedure. It is one of the best machine learning algorithms available. It employs a random subset of characteristics obtained by splitting a node to acquire the best feature that contributes the most to the model's construction. The result is even improved by assigning random threshold values to each feature. To score the characteristics, the Random Forest method is also used.

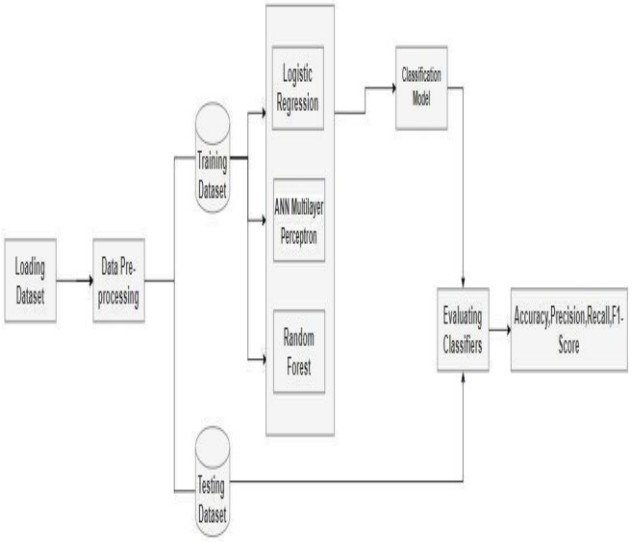

Figure 2.Architecture Diagram

6. Implementation and Results

Register/Signin : This is the registration and signin page of our application.

Figure 3.Register

Log in : This is the Login page of our application

Figure4.Login

Home Page: After successful login of the user home page will be displayed with the health analysis

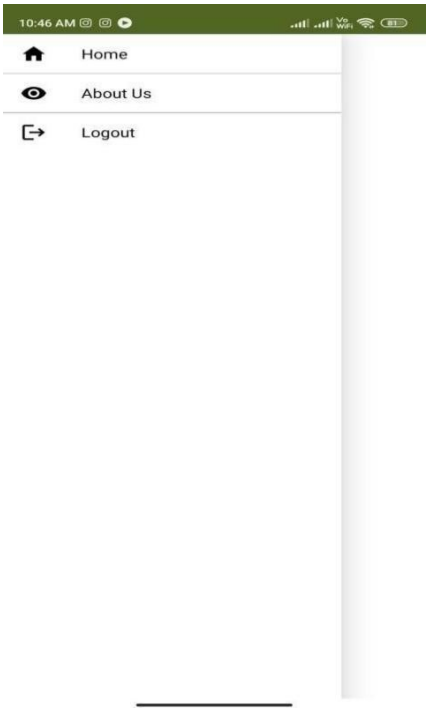

Figure 5.Home Page

Database: The username and password are stored in one database and the prediction parameters are stored in another database.

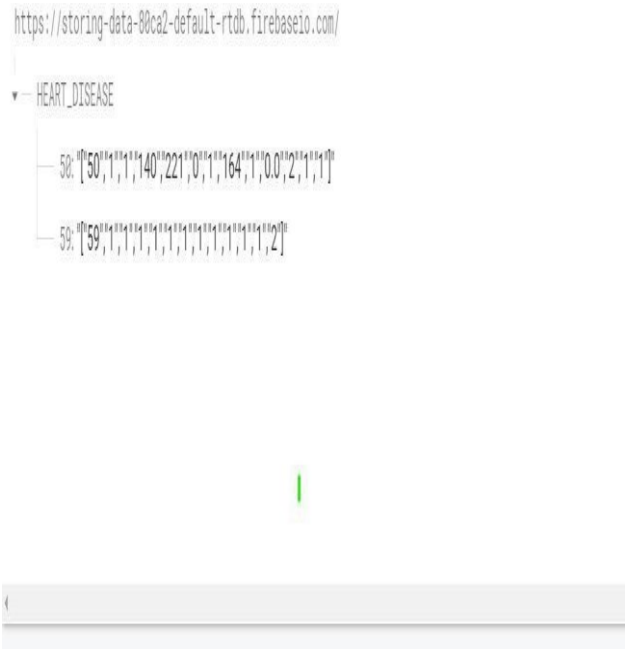

Figure 7.Adding Records

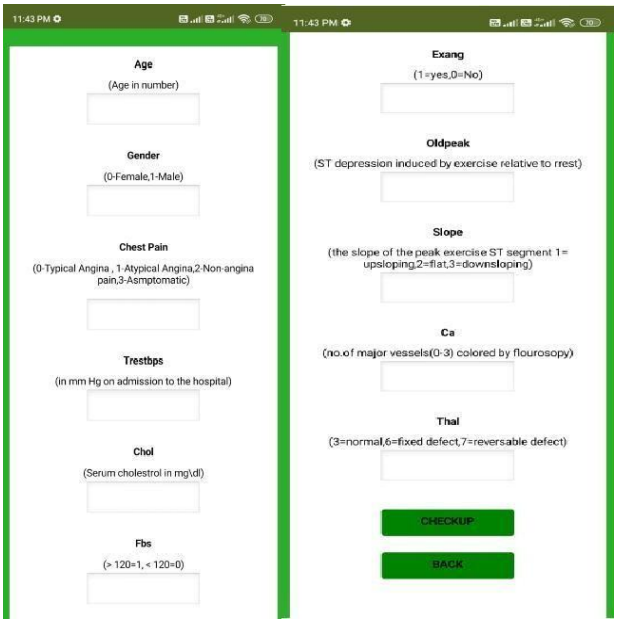

Figure 8.Prediction Screen

Prediction page: In this page the user will give the parameters and the parameters are stored in the database. Based on the parameters the model will predictwhether the person has heart disease or not.

Figure 9.Result

The final result is the prediction of the heart disease based on the given parameters.

7. Conclusion

Heart disease is a leading cause of death in India and around the world; using advanced tool like machine learning to detect heart disease early will have a significant impact on society. Early detection of heart disease can help high-risk individuals make lifestyle changes and, as a result, avoid problems, which can be a significant medical milestone. In this case, we used Random Forest, Logistic Regression, and ANN Multilayer Perceptron. Furthermore, this research focuses on personal health record systems, which are critical to EHR adoption (Electronic Health Record). PHR encourages the use of EHR because the doctor does not have to create the record from scratch. It also allows each user to maintain their own health records.

8. References

[1] Yu-Xuan Wang, QiHui Sun, Ting-Ying Chien, PoChun Huang, "Using Data Mining and Machine Learning Techniques for System Design Space Exploration and Automatized Optimization", Proceedings of the 2017 IEEE International Conference on Applied System Innovation, vol. 15, pp. 1079-1082, 2017.

[2] ZhiqiangGe, Zhihuan Song, Steven X. Ding, Biao Huang, "Data Mining and Analytics in the Process Industry: The Role of Machine Learning", 2017 IEEE. Translations and content mining are permitted for academic research only, vol. 5, pp. 20590-20616, 2017.

[3]P. Suresh Kumar and V. Umatejaswi, " Diagnosing Diabetes using Data Mining Techniques", International Journal of Scientific and Research Publications, Volume 7, Issue 6, June 2017.

[4] A. Roehrs, C. A. Da Costa, R. da Rosa Righi, & K. S. F. De Oliveira (2017). Personal health records: a thorough evaluation of the literature. 19(1), e5876; Journal of medical Internet research.

[5] Sumithra, R., Priya, R., and Jayalakshmi, V. (2020). Using cloud computing technology, create a personal health record. Research & Development in Indian Journal of Public Health, 11 (3).

[6]An overview of the Tesseract OCR engine by Ray Smith Document analysis and recognition's ninth international conference (ICDAR 2007). IEEE, 2007. Vol.

[7]Xiong, C., Yuan, S., and Mu, B. (2013). App Engine-Based Personal Health Records System. 2012 International Conference on Information Engineering .

[8]Latha, K., Ambika, M. (2015). A patient-centered framework for an intelligence-based recommendation system for healthcare. international conference on cutting-edge engineering andtechnologies, CAASR.

[9]Machine learning based decision support systems (DSS) for heart disease diagnosis: a review. Online: 25 March 2017 DOI: 10.1007/s10462-01

[10]Prerana T H M, Shivaprakash N C et al "Prediction of Heart Disease Using Machine Learning Algorithms- Naïve Bayes,Introduction to PAC Algorithm, Comparison of Algorithms and HDPS",Vol 3, PP: 90-99 ©IJSE, 2015

[11]Prediction System for heart disease using Naïve Bayes *Shadab Adam Pattekari and Asma Parveen Department of Computer Science and Engineering Khaja Banda Nawaz College of Engineering

[12]Comak E, Arslan A (2012) A biomedical decision support system using LS-SVM classifier with an efficient and new parameter regularization procedure for diagnosis of heart valve diseases. J Med Syst 36:549–556

[13]Han Wu, Shengqi Yang, Zhangqin Huang, Jian He, Xiaoyi Wang, "Type 2 diabetes mellitus prediction model based on data mining", ScienceDirect, Vol. 11,
issue 3, pp. 12-23, 2018.

[14] JahinMajumdar, Anwesha Mal, Shruti Gupta, "Heuristic Model to Improve Feature Selection Based on Machine Learning in Data Mining", 2016 6th International Conference - Cloud Systemand Big Data Engineering (Confluence), vol. 3, pp. 73-77, 2016.

[15]Ahmed Fawzi Otoom , Emad E. Abdallah , Yousef Kilani , Ahmed Kefaye and Mohammad Ashour(2015)Effective Diagnosis and Monitoring of Heart Disease ISSN: 1738-9984 IJSEIA