Prediction of Cardiac Disease using Machine Learning

Y Sai Sravani [1] | Ch. Srinivasa Rao[2]

[1]PG Scholar, Department of Computer Science, SVKP &Dr K S Raju Arts & Science College(A), Penugonda, W.G.Dt., A.P, India
[2]Associate Professor in Computer science, SVKP &Dr K S Raju Arts & Science College(A), Penugonda, W.G.Dt., A.P, India

To Cite this Article

Y Sai Sravani and Ch. Srinivasa Rao. Prediction of Cardiac Disease using Machine Learning. International Journal for Modern Trends in Science and Technology 2022, 8(09), pp. 190-195. https://doi.org/10.46501/IJMTST0809040

ABSTRACT

Heart assault disorder is one of the main reasons of the demise worldwide. In today's common modern life, deaths because of the coronary heart disorder had grown to be one in all fundamental issues, that kind of one man or woman misplaced his or her lifestyles consistent with minute because of coronary heart infection. Predicting the prevalence of disorder at early levels is a main challenge nowadays. Machine getting to know while applied in fitness care is able to early and correct detection of disorder. In this work, the bobbing up conditions of Heart Disease infection is calculated. Datasets used have attributes of medical parameters. The datasets are been processed in python the use of ML Algorithm i.e., Random Forest Algorithm. This approach makes use of the beyond antique affected person information for getting prediction of recent one at early levels stopping the lack of lives. In this work, dependable coronary heart disorder prediction gadget is applied the use of strong Machine Learning algorithm which is the Random Forest algorithm. Which read the patients record dataset in the form of CSV file. After having access to dataset the operation is finished and powerful coronary heart assault degree is produced. Advantages of proposed gadget are high overall performance and accuracy price and it's far very bendy and excessive charges of fulfilment are achieved.

1. INTRODUCTION

Heart sickness outcomes the functioning of the coronary heart. World Health Organization had made a survey and made an end that 10 million human beings are affected with coronary heart sickness and misplaced their lives. The hassle that the healthcare enterprise faces in today's existence is early prediction of sickness after someone is affected. Records or fact of clinical records is very huge and the facts in actual international is probably incomplete and inconsistent. In beyond predicting the sickness correctly and remedy to sufferers won't be feasible for each affected person at early ranges below those circumstances[2].

Many scientists attempted to construct a version that's able to predicting the coronary heart sickness withinside the early stage, but they're now no longer capable of construct a great version. Every proposed gadget has risks in its personal way. In the present gadget, Shenet al. had initially, proposed a gadget that's primarily based totally on self-applied -questionnaire. In this gadget the person want to input all of the signs which he's struggling from, primarily based totally on that the end result is predicted. This study is primarily based totally at the evaluation facts amassed in SAQ.

Chen et al. got here up with the concept to expect coronary heart sickness. He used the approach of Vector Quantization that's certainly considered one among the synthetic intelligence strategies for category and prediction purpose. Training of neural networks is executed the use of lower back propagation to assess the prediction gadget. In the checking out section about 80% accuracy is achieved on checking out set. Practical use of

facts collected from preceding facts is time consuming. Low accuracy rate.

So to conquer this we are imposing Random woodland set of rules if you want to attain correct outcomes in much less time. Machine getting to know is given a primary precedence in cutting-edge existence in many packages and in healthcare sector. Prediction is certainly considered one among region wherein machine getting to know performs an essential role, our subject matter is to expect coronary heart sickness through processing affected person's dataset and a facts of sufferers i.e., person of whom we want to expect the probabilities of incidence of a coronary heart sickness.

2. LITERATURE SURVEY

Heart conditions may result in senility, severe disorder, and a short lifespan. Additionally, it might be fatal. Therefore, implying heart disease is currently the most distressing. This essay focuses on several machine learning techniques that help identify and recognise a wide range of heart conditions. Hidden Markov Models, Support Vector Machines, Feature Selection, Computational Intelligent Classifiers, Prediction Systems, Data Mining Techniques, and Genetic Algorithm are just a few of the diverse machine learning technologies discussed here. We were able to choose the best strategy by carefully examining each one. This ultimately enables us to present an EnsembleModel that precisely classifies various heart illnesses by utilising applicable machine learning techniques. Utilizing cutting-edge technologies, the proposed technique has been evaluated. The accuracy, receiver operating characteristics, root mean square error, and precision of the proposed technique are significantly better than those of the K-Nearest Neighbour, Artificial Neural Networks, and Support Vector Machines algorithms, with an accuracy of 94.21%, ROC of 0.981, RMSE of.2568, and precision of 0.953. The suggested Ensemble Model and the implemented algorithms have both undergone analysis and evaluation taking into account the receiver operator characteristics.

The human body's most important and vital organ is the heart. The proper operation and efficient working of our hearts are absolutely necessary for life. It is a significant cause of death in the modern world. One of the most important health problems facing people today is heart disease. It is reportedly the main cause of death across the globe. Medical specialists frequently find it challenging to predict a cardiac illness early on. There are many valuable hidden facts and information in the health sector today that might be used to make predictions, particularly in the realm of medicine. Data mining is a process or strategy used to examine huge datasets and then produce substantial and practical outcomes using exceptional AI-based tools. This article aims to anticipate cardiovascular or heart illness using three of these AI-based methods: Decision Tree, Naive Bayes, and Neural Network. All of these techniques will be assessed based on many special & parameters with improvements for greater accuracy. The accuracy based on various parameters of each approach will then be compared. Then, the most reliable method is used to determine whether a man or woman will develop coronary heart disease. Medical professionals can utilise this method to predict diseases early so that prompt treatment can be taken by the patient. The detection of heart disease, which affects millions of people's lives, is one of the most significant uses of machine learning systems. Numerous independent factors, including age, sex, serum cholesterol, blood sugar, and others, are present in patients with heart disease and can be utilised to diagnose them extremely successfully. These characteristics are modelled in this study using the Extreme LearningMachine (ELM) method. A warning system for patients of the potential existence of heart disease can take the place of pricey medical examinations under the suggested approach. The Cleveland Clinic Foundation's genuine data, which includes information on 300 patients, is used to construct the system. According to simulation data, this architecture can accurately predict cardiac illness to a degree of 80%.

3. PROBLEM STATEMENT

The current systems' advanced coronary heart disease prediction machine uses a variety of algorithms, yet there are certain drawbacks. Using techniques for machine learning, each machine learning system will pay attention to these tactics,

A. The database or dataset we utilise for pre-processing contains NaN values, or "not a number," values. It is necessary to convert these non-numerical values to numerical values because the programme we use cannot

handle them. The method used is to replace NaN values with the column mean.

B. The data inside the database is divided into two types, namely training and testing sets. Data from the training set make up 80% of the data in the testing set.

C. Various machine learning algorithms, such as KNN, Adaptive boost, Decision Tree, and K-mean, are capable of classifying the selected data from the database under the training set.

1. Decision Tree Decision trees come in a variety of forms. Their primary difference is that they used to excel in the class feature. Tree root is the name of the function that exploits the benefit of records to reduce entropy in an entropy system. First, it is anticipated that information gained from all attributes contained inside the dataset will identify a tree root. The characteristic that makes use of the records advantage can then be determined.

2. KNN, one of the most straightforward and effective categorization techniques. . At the time of quality check, a few dependable regular controls of possibility densities are difficult to recognize due to the fact the consumer isn't always aware about them. So this KNN classification technique is applied to calculate such kind of calculations. With the assist of education datasets the area of K- nearest neighbour is predicted. Euclidean distance is used to discover how near the education dataset is from goal. Find the k-nearest pals and assign them to organization of rows that is examined. Repeat the step for the rows outstanding within the target set. In this application the highest value of K may be decided on, after that the software program software automatically builds a comparable parallel version at the values of K up to the maximum defined. KNN algorithm with aid of WEKA tool concludes that training dataset, input and output variables must derive in. The exceptional best value of K is used to construct parallel models on all of the values of K up to maximum known value. In this application, the largest value of K may be chosen, and then the software automatically creates a parallel version that is equivalent up to the highest value specified. The training dataset, input variables, and output variables must all originate from, according to the KNN method and WEKA tool. On all values of K up to

the highest known value, parallel models are built using the exceptional best value of K.
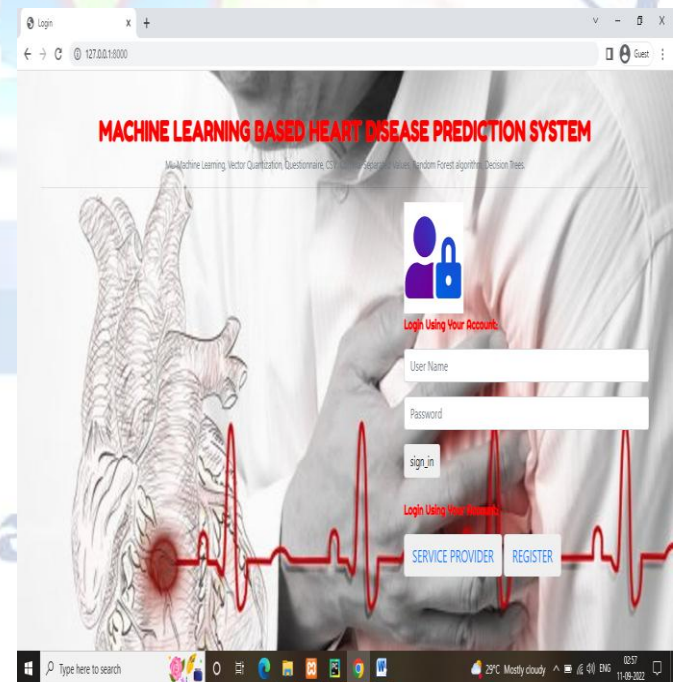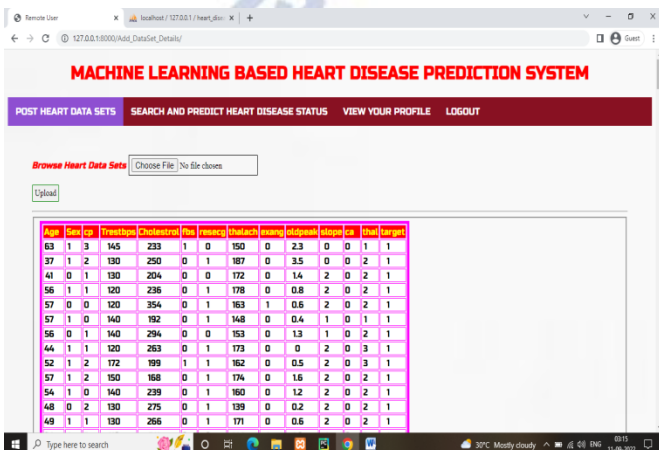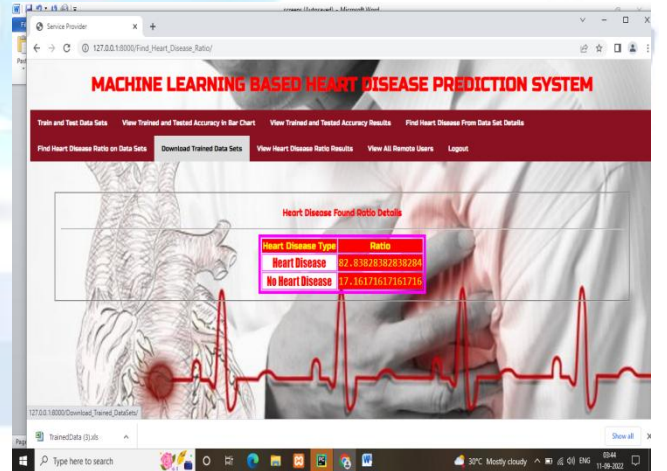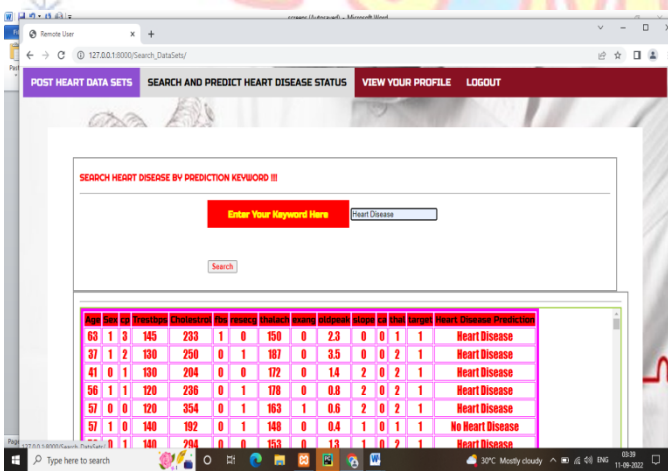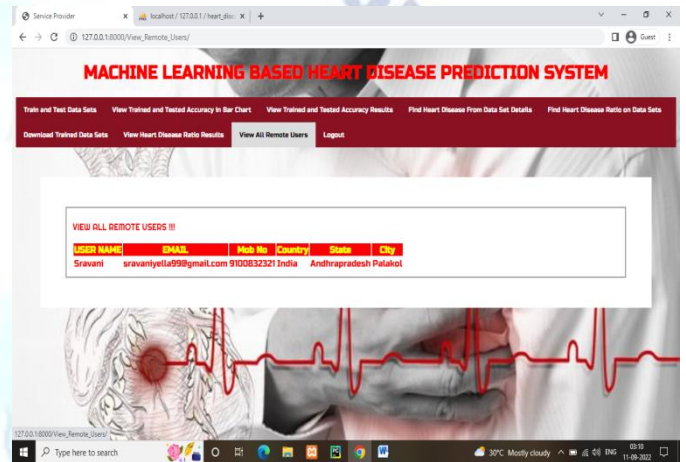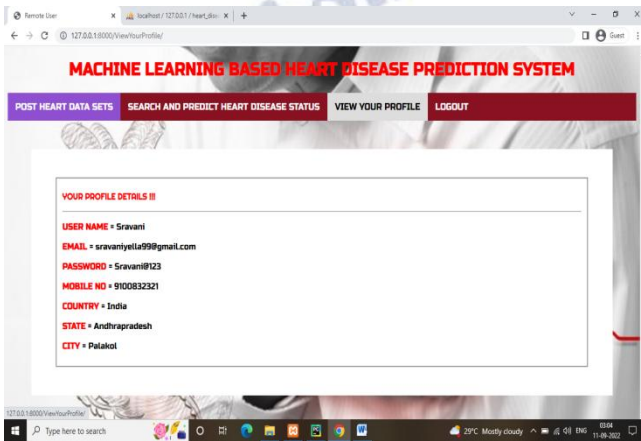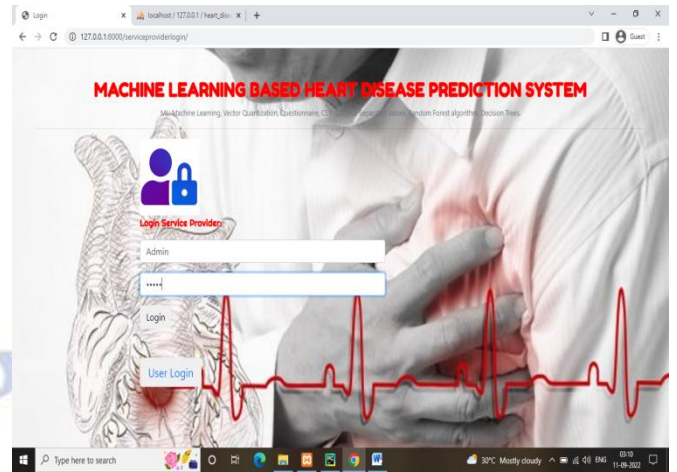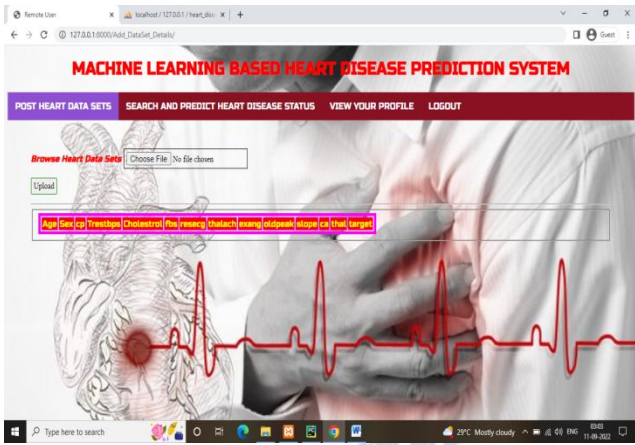
Disadvantages:

The existing system used the method of Vector Quantization which is one of the artificial intelligence strategies for classification and prediction purpose and is very much less performance.
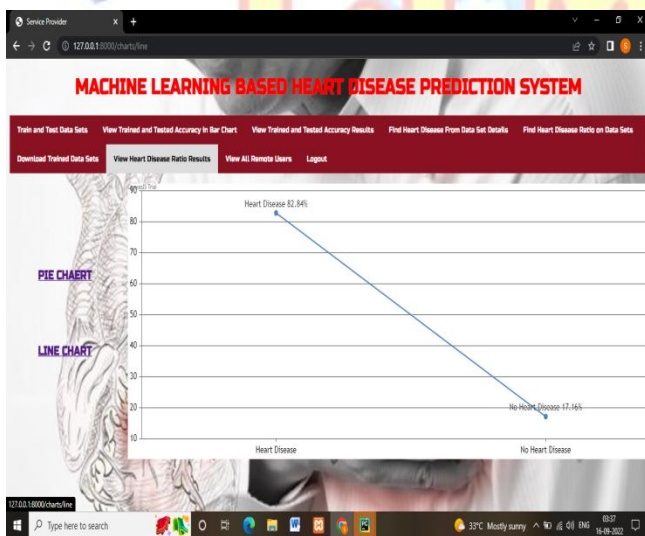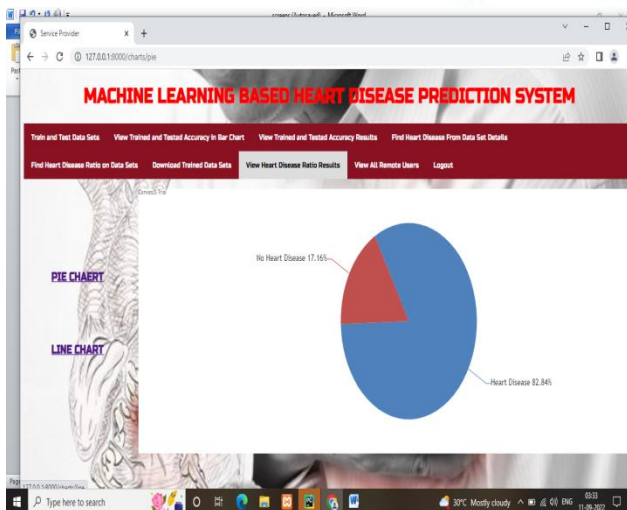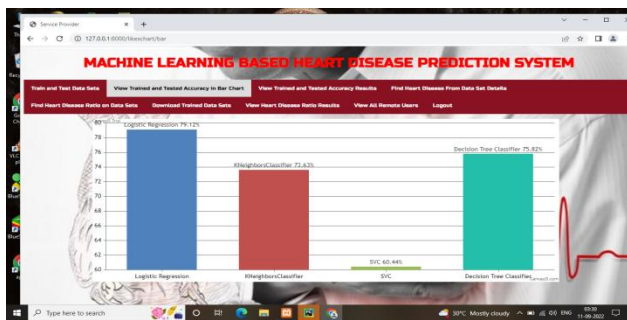
4. ARCHITECTURE
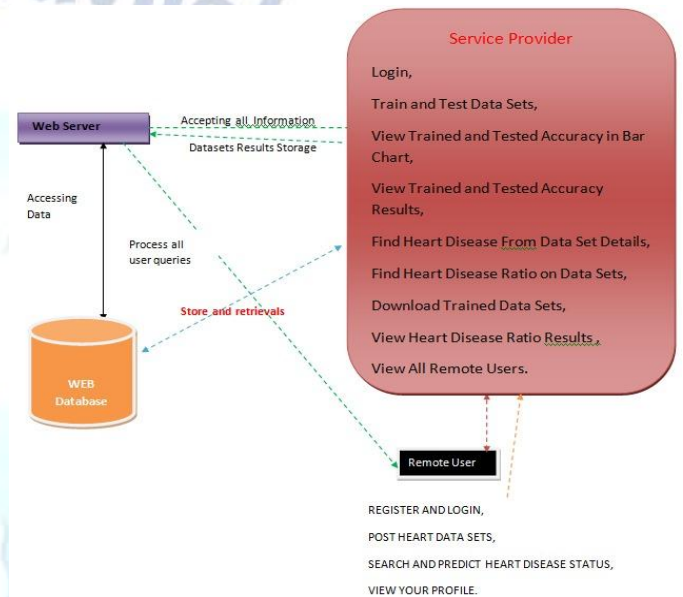
5. RESULTS

6. CONCLUSION

Random Forest algorithm is an efficient algorithm which is an ensemble learning approach for regression and classification techniques. The algorithm constructs N of Decision trees and outputs the magnificence that is the average of all decision trees output. So accuracy of prediction at early stages is accomplished effectively. Processing of healthcare data i.e., data associated with coronary heart will assist in early detection of coronary heart disorder or peculiar situation of coronary heart which ends up in saving of long time deaths. Heart

disorder prediction is a main task in the present contemporary life. With this utility if the patient/user is away from attain of health practitioner, he/she will be able to make use of the utility in prediction of disorder simply with the aid of using getting into the report values and can continue similarly whether or not to seek advice from a health practitioner or not.

Conflict of interest statement

Authors declare that they do not have any conflict of interest.

REFERENCES

[1] KaanUyar and Ahmetİlhan, "Diagnosis of heart disease using genetic algorithm based trained recurrent fuzzy neural networks" in B.V ICTASC, Elsevier,pp

[2] AshishChhabbi, LakhanAhuja, SahilAhir and Y. K. Sharma, "Heart Disease Prediction Using Data Mining Techniques", © IJRAT Special Issue National Conference "NCPC-2016", pp. 104-106, 19 March 2016.

[3] Berry JD, Lloyd-Jones DM, Garside DB, et al. Framingham risk score and prediction of coronary heart disease death in International journal of computer applications, published in 2006.

[4] Q. K. Al-Shayea, "Artificial neural networks in medical diagnosis," International Journal of Computer Science Issues,published in 2011. Disease Prediction System using Data Mining Techniques", ©

[5] Salam Ismaeel, Ali Miri et al., "Using the Extreme Learning Machine (ELM) technique for heart disease diagnosis", IEEE Canada International Humanitarian Technology

[6] TahiraMahboob, RidaIrfan and BazelahGhaffar et al."Evaluating ensemble prediction of coronary heart disease using receiver operating characteristics" ©2017 IEEE

[7] AmmarAsjad Raja, Irfan-ul-Haq ,MadihaGuftarTamim Ahmed Khan "Intelligence syncope Disease Prediction Framework using DM-techniques" FTC 2016 –Future Technologies Conference 2016.

[8] M.A. Jabbar, B.L.Deekshatulu, and Priti Chandra, " Intelligent heart disease prediction system using random forest and evolutionary approach", Journal of Network and Innovative Computing, Vol. 4, pp.174-184, 2016.

[9] N. Friedman, D. Geiger, and M. Goldszmidt, "Bayesian network classifiers," Machine Learning,(1997).

[10] AyonDey, Jyoti Singh, N. Singh "Analysis of supervised machine learning algorithms for heart disease prediction" International journal of computer applications, published in 2006.

[11] Q. K. Al-Shayea, "Artificial neural networks in medical diagnosis," International Journal of Computer Science Issues,published in 2011.

Authors Biography

Y Sai Sravani currently pursuing MCA in SVKP & Dr.K.S Raju Arts & Science College affiliated to Adikavi Nannaya University, Rajamahendravaram. Her research interests include Machine Learning , Big Data.

Ch.Srinivasa Rao Research Scholar in the Department of Computer Science & Engineering at Aacharya Nagarjuna University, Guntur, AP, India. He is working as Associate Professor in SVKP & Dr .KS Raju Arts & Science College (A), Penugonda, AP. He received Master's Degree in Computer Science & Engineering (M.Tech,CSE) from Jawaharlal Nehru Technological University, Kakinada, India. He qualified in UGC NET and AP SET. His research Interests include Data Mining, Data Science and Machine Learning.