IJARSCT
ISSN (Online) 2581-9429
International Journal of Advanced Research in Science, Communication and Technology (IJARSCT)
Volume 2, Issue 7, May 2022
Impact Factor: 6.252

A Hybrid Intelligent System for the Prediction of Heart Disease Using Machine Learning Algorithms

Prof. Mrs. Abhilasha S. Shinde[1], Nachiket Tambe[2], Parag Patil[3], Sachin Phad[4]
Project Guide, Department of Information Technology[1]
Projecties, Department of Information Technology[2,3,4]
Smt. Kashibai Navale College of Engineering, Pune, Maharashtra, India

Abstract: *In today's modern world cardiovascular disease is the most lethal one. This disease attacks a person instantly that might create unexpected consequences for the human life. So diagnosing patients correctly on time is the most challenging task for the medical fraternity. The heart disease treatment is quite high and not affordable by most of the patients particularly in India. The research scope is to develop an early prediction treatment using data mining technologies. Now a day every hospital keeps the periodical medical reports of cardiovascular patients through some hospital management system to manage their health-care. The data mining techniques namely decision tree and random forest are used to analyze heart attack dataset where classification of more common symptoms related to heart attack is done using c4.5 decision tree algorithm, alongside, random forest is applied to improve the accuracy of the classification result of heart attack prediction. In this system various data mining technologies are applied to make a proactive approach against failures in early predictions diagnosis of the disease. We proposed an automated system for medical diagnosis that would enhance medical care and reduce cost. Our aim is to provide a ubiquitous service that is both feasible, sustainable and which also make people to assess their risk for heart attack at that point of time or later.*

Keywords: Heart attack prediction, ML, Random Forest, Decision Tree, Java, servlet, etc

I. INTRODUCTION

Currently In today's modern world cardiovascular disease is the most lethal one. This disease attacks a person so instantly that it hardly gets any time to get treated with. So diagnosing patients correctly on timely basis is the most challenging task for the medical fraternity. A wrong diagnosis by the hospital leads to earn a bad name and loosing reputation. At the same time treatment of the said disease is quite high and not affordable by most of the patients particularly in India. The motivation of this system is to develop a cost effective treatment using data mining technologies for facilitating data base decision support system. Nowadays, Cardiovascular Disease (Heart Attack) is ubiquitous and one of the major reasons of death worldwide. Early screening of people at risk of having CVD may lead to minimize morbidity and mortality. A simple approach is proposed in this project to predict risk of developing heart attack using Smartphone and data mining technique. In today's modern world cardiovascular disease is the most lethal one. This disease attacks a person so instantly that it hardly gets any time to get treated with. So diagnosing patients correctly on timely basis is the most challenging task for the medical fraternity. A wrong diagnosis by the hospital leads to earn a bad name and loosing reputation. At the same time treatment of the said disease is quite high and not affordable by most of the patients particularly in India. The motivation of this system is to develop a cost effective treatment using data mining technologies for facilitating data base decision support system. Nowadays, Cardiovascular Disease (Heart Attack) is ubiquitous and one of the major reasons of death worldwide. Early screening of people at risk of having CVD may lead to minimize morbidity and mortality. A simple approach is proposed in this project to predict risk of developing heart attack using Smartphone and data mining technique. Project will be developed as a prototype model using python technology. It will run as an android application. System will predict the heart attack using predefined data set. System performance will depend upon dataset provided. WLAN will be used for communication purpose between android and database.

IJARSCT
ISSN (Online) 2581-9429
International Journal of Advanced Research in Science, Communication and Technology (IJARSCT)
Volume 2, Issue 7, May 2022
Impact Factor: 6.252

II. LITERATURE REVIEW

In Paper [1], Abhishek Ta, "Heart Disease Prediction System Using Data Mining Techniques" In this study, the aim was to design a predictive model for heart disease detection using data mining techniques from Transthoracic Echocardiography Report dataset that is capable of enhancing the reliability of heart disease diagnosis using echocardiography.

In Paper [2], Sellappan Palaniappan, Rafiah Awang "Intelligent Heart Disease Prediction System Using Data Mining Techniques" A prototype heart disease prediction system is developed using three data mining classification modeling techniques. The system extracts hidden knowledge from a historical heart disease database. DMX query language and functions are used to build and access the models. The models are trained and validated against a test dataset. Lift Chart and Classification Matrix methods are used to evaluate the effectiveness of the models. All three models are able to extract patterns in response to the predictable state. The most effective model to predict patients with heart disease appears to be Naive Bayes followed by Neural Network and Decision Trees.

In Paper [3], Aditya Methaila, Prince Kansal,"EARLY HEART DISEASE PREDICTION USING DATA MINING TECHNIQUES" In this paper the focus is on using different algorithms and combinations of several target attributes for effective heart attack prediction using data mining. Association classification technique apriori algorithm, was along with a new algorithm MAFIA was used. Straight Apriori based algorithms count all of the 2k subsets of each k-item set they discover, and thus do not scale for long item sets. They use "look a heads" to reduce the number of item sets to be counted. MAFIA is an improvement when the item sets in the database are very long.

In Paper [4], S. Kiruthika Devi, S. Krishnapriya and Dristipona Kalita"Prediction of Heart Disease using Data Mining Techniques" The accuracy of the algorithms used in each technique can be enhanced by hybridizing or combining algorithms a single algorithm which may not be accurate for weakly classified sets of data, and is expected to make quicker and more precise decisions.

In Paper [5], M.Lavanya, Mrs.P.M.Gomathi, "Prediction of Heart Disease using Classification Algorithms" Medical related information are huge in nature and it can be derived from different birthplaces which are not entirely applicable in feature. The research undertook an experience on application of varies data mining algorithm to predict the heart attacks and to compare the based method of prediction. The predictive accuracy determined by J48,REPTREE, naive bayes, neural networks, CART. The overall objective is to study the various data mining techniques available to predict the heart

III. SYSTEM ANALYSIS AND PROPOSED SYSTEM

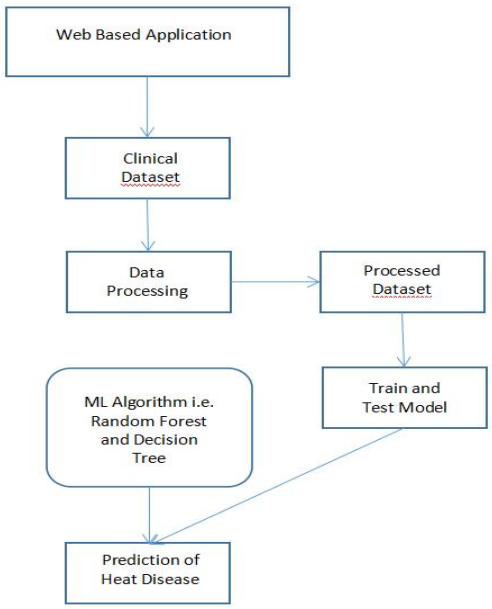

Fig. 1 System Architecture

IJARSCT
ISSN (Online) 2581-9429
International Journal of Advanced Research in Science, Communication and Technology (IJARSCT)
Volume 2, Issue 7, May 2022
Impact Factor: 6.252

The system will go through various operations such as first input image is given , which will go under pre processing, segmentation, feature extraction, classification. In pre processing the unwanted part is removed, segmentation will divide the area of interest into number of parts, feature extraction phase will extract the data and will store it for comparison purpose. After classification phase and comparison finally the heart attack will be predicted. After reviewing various research papers we found that this is the best system architecture for our project. Just the difference between existing system architecture and our proposed system architecture is we used some clinical data instead of data warehouse because data warehouse will berequire only when this project will be widely used in medical industry. We simply created easy to use web based application for prediction f heart disease. The proposed system inspired after studying similar systems and the team customized to make it more efficient. The section also explains the working of the proposed architecture in a brief. The performance of the proposed architecture is also explained in the section where it can be considered to be more efficient than that of the existing one.

3.1 Algorithm-Random Forest

Random forest could be a supervised learning algorithmic rule that is employed for each classification yet as regression. however but, it's primarily used for classification issues. As we all know that a forest is created from trees and additional trees suggests that additional sturdy for standard time. Similarly, random forest algorithmic rule creates call trees on knowledge samples so gets the prediction from every of them and at last selects the most effective resolution by suggests that of choice. it's associate degree ensemble methodology that is best than one call tree as a result of it reduces the over-fitting by averaging the result.

Working of Random Forest Algorithm

We can perceive the operating of Random Forest algorithmic rule with the assistance of following steps-

- First, begin with the choice of random samples from a given data-set.
- Next, this algorithmic rule can construct a choice tree for each sample. Then it'll get the prediction result from each call tree.
- In this step, choice are going to be performed for each foreseen result.
- At last, choose the foremost voted prediction result because the final prediction result. The following diagram can illustrate its operating.

Precondition: A training set S := $(x_1, y_1), \ldots, (x_n, y_n)$, features F, and number of trees in forest B.

```
1  function RANDOM FOREST(S,F)
2     H←0
3     for iε1,....,B do
4           S(i) ←A bootstrap sample from S
5           Hi ←RANDOMIZEDTREELEARNS(S(i),F)
6           H←HU{hi}
7     end for
8     return H
9  end function
10 function RANDOMIZEDTREELEARNS(S, F)
11    At each node:
12          f←very small subset of F
13          Split on best feature in f
14          return the learned tree
15 end function
```

3.2 Algorithm-Decision Tree

Step-1: Begin the tree with the foundation node, says S, that contains the entire data set.

Step-2: Realize the most effective attribute within the data set mistreatment Attribute choice live (ASM).

IJARSCT
ISSN (Online) 2581-9429
International Journal of Advanced Research in Science, Communication and Technology (IJARSCT)
Volume 2, Issue 7, May 2022
Impact Factor: 6.252

Step-3: Divide the S into subsets that contains potential values for the most effective attributes. Step-4: Generate the choice tree node, that contains the most effective attribute.

Step-5: Recursively build new call trees mistreatment the subsets of the data set created in step -3. Continue this method till a stage is reached wherever you can not any classify the nodes and referred to as the ultimate node as a leaf node

INPUT: *S,* ***Where S****=set of classified instances*
OUTPUT: *Decision Tree*
Require: *S≠Ø, num_attributes>0*

```
1   procedure: BUILDTREE
2       repeat
3           marGain←0
4           splitA←null
5           e←Entropy(Attributes)
6           for all  Attitributes a in S do
7               gain←InformationGain(a, e)
8               if gain>maxGain then
9                   maxGain←gain
10                  splitA←a
11              end if
12          end for
13          Partition(S, split A)
14      until all partitions processed
15  end procedure
```

IV. RESULTS

Here is the screenshot of Login Page from where admin have to login using their username and password which he/she set. In admin login there are many sections which handled by only admin. Also our heart attack prediction system's main part is in admin login i.e. prediction part. Admin have to handle patient's health record and diagnose them via our already built system.

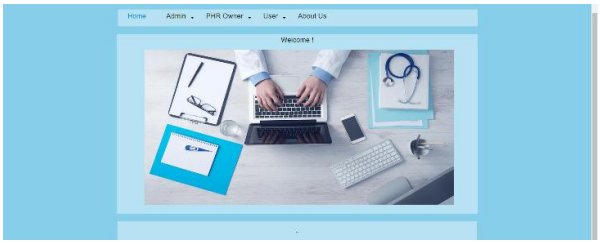

Fig. 2 Home Page of System

This is the screenshot of registration page of PHR owner that is Patient Health Record owner which is nothing but patient. For creating health record patient have to register himself via creating username, password, entering email id and mobile name, self name then after registering admin have to activate that particular patient for login and for health record creation.

Fig. 3 Registration Page of System

IJARSCT
ISSN (Online) 2581-9429
International Journal of Advanced Research in Science, Communication and Technology (IJARSCT)
Volume 2, Issue 7, May 2022
Impact Factor: 6.252

This is the screenshot of Patient's Health Record. We used 13 attributes which is generally common but most important attributes in view of heart attack failure or heart disease. This 13 attributes are age, sex, chest pain type, blood pressure, cholesterol, fasting blood pressure, ecg type, maximum heart rate achieved, exercise induced angina, exercise relative to rest, heart rate slope, cardiac stress and blood disorder.

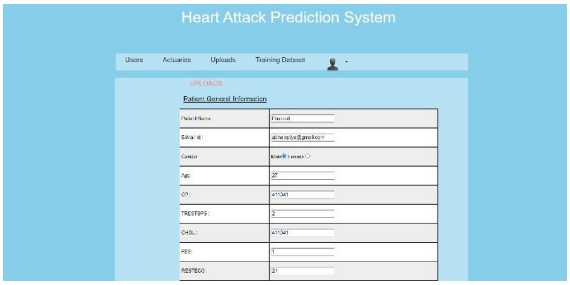

Fig. 4 Patient's Health Record Page of System

This is the screenshot of our system which showing that how patient diagnose and how message is display on screen that patient is having negative/positive chances of heart attack.

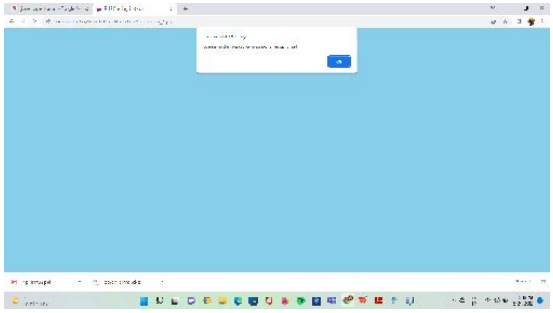

Fig. 5 Prediction Page of System

V. CONCLUSION

The symptoms of heart disease are familiar for everyone. But no one can predict when it may happen and also the prediction will not give any instance status i.e. normal or abnormal. Our research intension is to notify the patients that neither normal nor abnormal at its early stage. The application of data mining algorithms shows it better performance results among themselves. This app will be supposed to predict the heart attack from the chest pain at an early stage and guide the person to take treatment early such as to get their ECG done as early as possible and get evaluated from a doctor to make diagnoses of heart attack.

FUTURE SCOPE

In future genetic algorithm will be used in order to reduce the actual data size to get the optimal data subset of attribute sufficient for heart disease prediction. Prediction of the heart disease will be evaluated according to the result produced from it.

LIMITATIONS

The system has less training data set as compared to huge medical industry and huge parameters. It have only 13 attributes to detection of heart disease. The health record data is in the form of numerical data only.

REFERENCES

[1]. Taneja, Orient. J. Comp. Sci. & Technol., Vol. 6(4), 457-466 (2013)

[2]. S. Palaniappan and R. Awang, "Intelligent heart disease prediction system using data mining techniques," 2008 IEEE/ACS International Conference on Computer Systems and Applications, 2008, pp. 108-115, doi: 10.1109/AICCSA.2008.4493524.

IJARSCT
ISSN (Online) 2581-9429
International Journal of Advanced Research in Science, Communication and Technology (IJARSCT)
Volume 2, Issue 7, May 2022
Impact Factor: 6.252

[3]. B. V. Baiju and R. J. Remy Janet, "A survey on heart disease Diagnosis and Prediction using Na¨ıve Bayes in Data Mining", IJCET, Vol 5, No.2, April 2015.

[4]. Dr.I.Lakshmi, "Prediction Analysis on Heart Disease using HNB and NB Techniques" SSRG International Journal of Computer Science and Engineering 5.10 (2018): 11-15.

[5]. TY - BOOK AU - Ngare, Kennedy PY - 2019/03/08 SP - T1 - Heart Disease Prediction System ER -

[6]. Sharmila S et al., "Analysis of Heart Disease Prediction using Data Mining Techniques", International Journal of Advanced Networking & Applications (IJANA), Volume: 08, Issue: 05,Pages: 93-95 (2017).

[7]. R. Atallah and A. Al-Mousa, "Heart disease detection using machine learning majority voting ensemble method," in Proceedings of the 2019 2nd International Conference on New Trends In Computing Sciences (ICTCS), pp. 1–6, Amman, Jordan, (October 2019)

[8]. P. Venkatesan and N. R. Yamuna. "Treatment response classification in randomized clinical trials: a decision tree approach", Indian Journal of Science and Technology, 6.1 (2013): 3912-3917.