Heart Disease Prediction System Using Data Mining Technique.

Devendra Ratnaparkhi, Tushar Mahajan, ,Vishal Jadhav

1 Student, Department of CSE, DPCOE,Pune, Maharashtra,India
2Student, Department of CSE, DPCOE,Pune, Maharashtra,India
3Student, Department of CSE, DPCOE,Pune, Maharashtra,India

---***--

Abstract -. . *The health care environment is found to be rich in information, but poor in extracting knowledge from the information. This is because of the lack of effective analysis tool to discover hidden relationships and trends in them. By applying the data mining techniques, valuable knowledge can be extracted from the health care system. Heart disease is a group of condition affecting the structure and functions of heart and has many root causes. Heart disease is the leading cause of death in the world over past ten years. Research has been made with many hybrid techniques for diagnosing heart disease. We are using Naive Bayes algorithms to increase the accuracy of Heart Disease prediction System. It is implemented as web based application in which user answers the predefined questions. It retrieves hidden data from stored database and compares the user values with trained data set. It can answer complex queries for diagnosing heart disease and thus assist healthcare practitioners to make intelligent clinical decisions which traditional decision support systems cannot. By providing effective treatments, it also helps to reduce treatment costs.*

Key Words: Data Mining, Naïve Bayes, Heart Disease.

1. INTRODUCTION

Data Mining is a non-trivial extraction of implicit, previously unknown and potential useful information about data. In short, it is a process of analyzing data from different perspective and gathering the knowledge from it. The discovered knowledge can be used for different applications for example healthcare industry. Nowadays healthcare industry generates large amount of data about patients, disease diagnosis etc. Data mining provides a set of techniques to discover hidden patterns from data. A major challenge facing Healthcare industry is quality of service. Quality of service implies diagnosing disease correctly & provides effective treatments to patients. Poor diagnosis can lead to disastrous consequences which are unacceptable.

1.1 HEART DISEASE

The heart is important organ or part of our body. Life is itself dependent on efficient working of heart. If operation of heart is not proper, it will affect the other body parts of human such as brain, kidney etc. It is nothing more than a pump, which pumps blood through the body. If circulation of blood in body is inefficient the organs like brain suffer and if heart stops working altogether, death occurs within minutes. Life is completely dependent on efficient working of the heart. The term Heart disease refers to disease of heart & blood vessel system within it.

There are number of factors which increase the risk of Heart disease :
Family history of heart disease
Smoking
Poor diet
High blood pressure , Cholesterol
Obesity
Physical inactivity
Hyper tension

1.2 NAÏVE BAYES

In probability theory, Bayes' law after relates the conditional and marginal probabilities of two random events. It is often used to compute posterior probabilities given observations. For example, a patient may be observed to have certain symptoms. Bayes' theorem can be used to compute the probability that a proposed diagnosis is correct, given that observation. A naive Bayes classifier is a term dealing with a simple probabilistic classification based on applying Bayes' theorem. Depending on the precise nature of the probability model, naive Bayes classifiers can be trained very efficiently in a supervised learning. An advantage of the naive Bayes classifier is that it requires a small amount of training data to estimate the parameters necessary for classification. Bayes Rule:
A conditional probability is the likelihood of some conclusion, C, given some evidence/observation, E, where a dependence relationship exists between C and E. This probability is denoted as P$(C |E)$ where
P(C | E) =(P(C | E) P(C))/P(E).

2.LITERATURE SURVEY.

Numerous studies have been done that have focus on diagnosis of heart disease. They have applied different data mining techniques for diagnosis & achieved different probabilities for different methods.

According to Dr.Durairajan, persia and sivagowry ,Assistant Professor Trichy University,Tamilnadu ,In his paper "Emperical study on data mining techniques for the analysis and prediction of heart disease"

Comparisons among different algorithm models given as,

Algorithm used	Accuracy	Time Taken
Neural Network	45.67%	1000 ms
Decision Tree	52%	719 ms
Naïve Bayes	53.00%	609 ms

> According to Ms .Chaitrali S .Dangare in his International Journal Paper on "Improved Study of Heart Disease Prediction System using Data Mining Classification Techniques",the comparisons among different model is as follows.

Algorithm used	Accuracy	Time Taken
Neural Network	99%	900 ms
Decision Tree	98%	850 ms
Naïve Bayes	90%	609 ms

Neural networl algorithms are very hard to implement and decision trees also. So Naïve Bayes is Good Choice For Prediction system.

Decision Support in Heart Disease Prediction System using Naive Bayes(2011)

This research has developed a Decision Support in Heart Disease Prediction System (DSHDPS) using data mining modeling technique, namely, Naïve Bayes.

Using medical profiles such as age, sex, blood pressure and blood sugar it can predict the likelihood of patients getting a heart disease.

Diagnosis of Heart Disease Using Data Mining Algorithm(2014)

The principle of this study is, hence to extract hidden patterns by applying data mining techniques, which are noteworthy to heart diseases, from a data collected together by an International Cardiovascular Hospital.

3.ADVANTAGES AND FEATURES
> Prepare History Database.
> Apply Training.
> Detect Patient Whether He is Ok Or Having Heart Disease
> Helping for New comer Doctor.
> Less Time Consuming.
> Best Prediction of Disease.
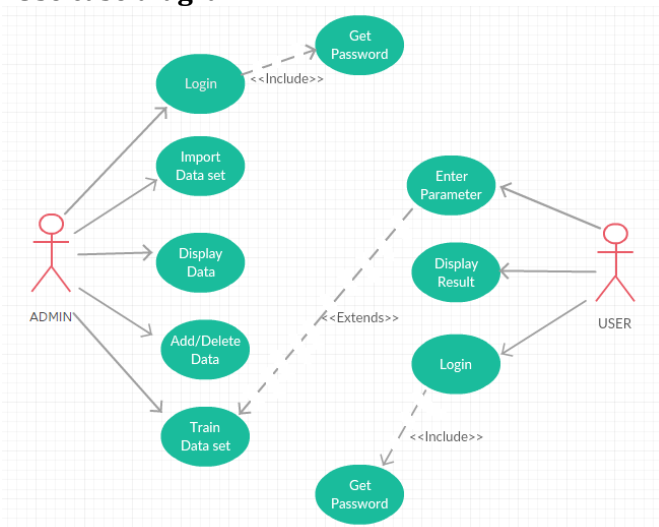
4.TECHNOLOGIES USED

Front end tool: AWT and SWING
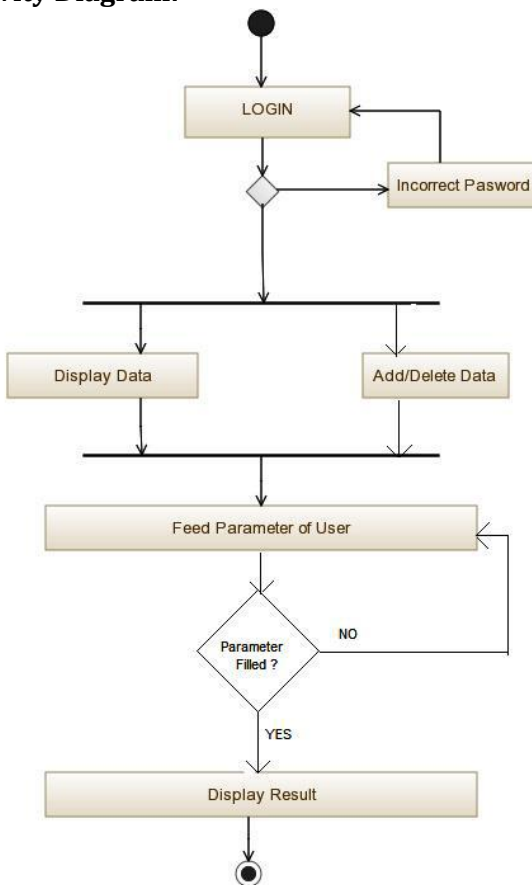
Back end tool: serialization .

Language used: java

5. System Architectures
Use case diagram:

Activity Diagram:

4. CONCLUSIONS

In this paper, we present an heart disease prediction system based on based on naïve bayes algorithm .This system is convenient, effective and gives good prediction of diseases to users. Overall conclusion is that, this is a better prediction system user of customer using data mining technology.

In next phase, we are working on increasing accuracy of naïve bayes algorithm for better prediction of disease.

6.Acknowledgement

This work is completed under the guidance of Prof. Vandana Navale and Prof. Priyanka Kedar. We express our gratitude towards them for their continuous support on this research. We would also like to thank the reviewers for their suggestions to improve this paper.

7.References

[1]G.Subbalakshmi et al. / Indian Journal of Computer Science and Engineering (IJCSE). Decision Support in Heart Disease Prediction System using naïve bayes.

[2]International Journal of Computer Applications (0975 – 888) Volume 47– No.10, June 2012 / Improved Study of Heart Disease Prediction System using Data Mining Classification Techniques .

[3]Volume 4, Issue 1, January 2014 ISSN: 2277 128X International Journal of Advanced Research in Computer Science and Software Engineering /Study of Heart Disease Prediction using Data Mining .