International Journal of Engineering Research & Management Technology ISSN: 2348-4039
July-2016 Volume 3, Issue-4
Email: editor@ijermt.org www.ijermt.org

PREDICTION FOR HEART DISEASE PROBLEM BASED ON MOST SUITABLE RECOMMENDATION

Abhale Babasaheb Annasaheb
M. Tech
Department CSE
Lord Krishna College of Technology Indore
M.P

Vijay Kumar Verma
Assistant Professor
Department CSE
Lord Krishna College of Technology Indore
M.P

ABSTRACT: *Knowledge discovery and data mining have found numerous applications in business and scientific domain. Valuable knowledge can be discovered from application of data mining techniques in healthcare system. The healthcare industry collects huge amounts of healthcare data which, unfortunately, are not "mined" to discover hidden information Diagnosis of heart disease is a significant and tedious task in medicine. The term Heart disease encompasses the various diseases that affect the heart. The exposure of heart disease from various factors or symptom is an important issue. Aim of this paper is to propose a model based on most suitable and appropriate value for early detection and correct diagnosis of the disease which will help the doctor in saving the life of the patient*

KEYWORD: - *Prediction, Classification Diagnosis, Heart Attack, Symptoms*

I INTRODUCTION

Heart Diseases problem is biggest cause of deaths in the last a few two decades. Development in computer technology and machine learning techniques are able to design and develop good quality software to assist doctors in making efficient decision of various diseases at the early stage. The diagnosis of each and every disease depends on clinical and pathological tested data. Heart disease prediction system can assist medical professionals in predicting heart disease status based on the clinical data of patients. Data mining with intelligent algorithms can be used to tackle the problem of prediction in medical dataset involving multiple inputs. Now a day's Artificial neural network has been used for complex and difficult tasks. The neural network is trained from the historical data with the hope that it will discover hidden dependencies and that it will be able to use them for predicting.

There Are Number Of Factors Which Increases Risk Of Heart Disease. Some Of Them Are Listed Below

- Family history of heart disease
- Smoking
- Cholesterol
- High blood pressure
- Obesity
- BMI

Predication should be done to reduce risk of Heart disease. Diagnosis is usually based on signs, symptoms and physical examination of a patient. Almost all the doctors are predicting heart disease by learning and experience. The diagnosis of disease is a difficult and tedious task in medical field. Predicting Heart disease from various factors or symptoms is a multi-layered issue which may lead to false presumptions and unpredictable effects

II. ARCHITECTURE FOR HEART DISEASE DIAGNOSIS

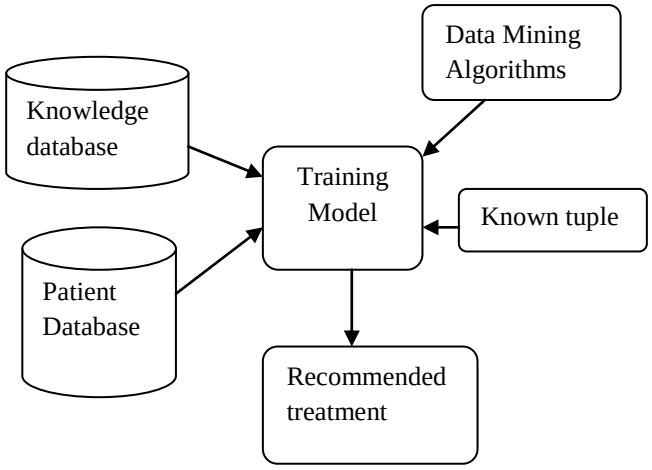

Figure1 Architecture for heart disease diagnosis

III. LITERATURE REVIEW

In 2011 Mr. K. Ramesh proposed "Decision Support in Heart Disease Prediction System using Naive Bayes". The proposed works developed a Decision Support in Heart Disease Prediction System (DSHDPS) using data mining modeling technique by using 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. Proposed work implemented as web based questionnaire application. It can serve a training tool to train nurses and medical students to diagnose patients with heart disease [1].

In 2011 Mai Shouman Tim Turner and Rob Stocker proposed "Using Decision Tree for Diagnosing Heart Disease Patients". In proposed work they investigate and applying a range of techniques to different types of Decision Trees seeking better performance in heart disease diagnosis. A widely used benchmark data set is used in proposed work. They evaluate the performance of the alternative Decision Trees the sensitivity, specificity, and accuracy. The proposed model outperforms J4.8 Decision Tree and Bagging algorithm in the diagnosis of heart disease patients [2].

In 2012 Chaitrali S. Dangare and Sulabha S. Apte proposed "Improved Study of Heart Disease Prediction System using Data Mining Classification Techniques". They analysed prediction systems for Heart disease using more number of input attributes. The system uses medical terms such as sex, blood pressure, cholesterol like 13 attributes to predict the likelihood of patient getting a Heart disease. They added two more attributes i.e. obesity and smoking. The data mining classification techniques, namely Decision Trees, Naive Bayes, and Neural Networks are analyzed on Heart disease database. The performance of these techniques is compared, based on accuracy. They showed that out of these three classification models Neural Networks predicts Heart disease with highest accuracy [3].

In 2013 Vikas Chaurasia proposed "Early Prediction of Heart Diseases Using Data Mining Techniques". Their main objective is to take advantage of available technological advancements and develop prediction models for heart disease survivability. They used three popular data mining algorithms CART (Classification and Regression Tree), ID3 (Iterative Dichotomized 3) and decision table (DT) extracted from a decision tree or rule-based classifier to develop the prediction models using a large dataset [4].

In 2013 Shamsher Bahadur et al. proposed "Predict the Diagnosis of Heart Disease Patients Using Classification Mining Techniques". They focused to predict the diagnosis of heart disease with reduced number of attributes. There fourteen attributes involved in predicting heart disease. But they used only six attributes and apply Genetic algorithm. Subsequently three classifiers like Naive Bayes, Classification by Clustering and Decision Tree are used to predict the diagnosis of heart disease after the reduction of number of attributes [5].

In 2014 N S Nithya and K Duraiswamy proposed "Gain ratio based fuzzy weighted association rule mining classifier for medical diagnostic interface". They proposed an information gain based fuzzy association rule mining algorithm for extracting both association rules and membership functions of medical data to reduce the rules. They used a ranking based weight value to identify the potential attribute. A large number of distinct values, the computation of information gain value is not feasible. By experimental results they showed that there is a marginal improvement in the attribute selection process and also improvement in the classifier accuracy [6].

In 2015 Ankur Makwana and Jaymin Patel proposed "Decision Support System for Heart Disease Prediction using Data Mining Classification Techniques". In the proposed work they design a predictive model for heart disease detection using Machine Learning and Data Mining techniques. They show fuzzy model was found to have the capacity of evaluating the connections between the input of predicted patients and predicted outcomes of patients results. This methodology gives a decent answer for manage instability in health framework variables and instability in the admission of a patient [7].

In 2016 Gunsai Pooja and Mrs. Lolita Singh " A Review on Data Mining for Heart Disease Prediction". They provide a survey of current techniques of knowledge discovery in databases using data mining techniques. The system uses medical terms such as sex, blood pressure, cholesterol etc. like attributes to predict the likelihood of patient getting a Heart disease. The performance of these techniques is compared, based on accuracy. The main objective of this research is to develop a prototype Intelligent Heart Disease Prediction System (IHDPS) using data mining modeling technique, namely, Clustering. it can discover and extract hidden knowledge (patterns and relationships) associated with heart disease from a historical heart disease database. [8].

In 2016 K. Manimekalai proposed "Prediction of Heart Diseases using Data Mining Techniques". The objective of proposed work is to find the best method of prediction to predict the heart disease and. Their main objective is to provide a study of Heart Diseases using various Data Mining Techniques. When the data mining technique is used separately the accuracy is low. To improve the accuracy value, data mining techniques should be combined together. SVM Classifier with genetic Algorithm contains 95% high accuracy than other technique. In future the techniques are hybrid, the accuracy will high. It will definitely help the patients as well as the medical practitioners to predict the heart disease [9].

In 2016 S. B. Bhalerao and Dr. B. L. Gunjal proposed "Survey of Heart Disease Prediction Based on Data Mining Algorithms". They showed that data mining provides the methodology and technology to transform data into useful information for decision making. Use of data mining algorithms will result in quick prediction of disease with high accuracy. They proposed a survey work on different papers in which single or hybrid combination of data mining algorithms are used for the prediction of heart disease. [10].

IV. PROPOSED APPROACH

Consider A Simple Example

Table 1 Training tuples

RID	Age	Smoking	Hypertension	(BMI)	Heart disease
1	42	Yes	Yes	40	Yes
2	62	Yes	No	23	No
3	55	No	Yes	40	Yes
4	62	Yes	Yes	50	Yes
5	45	No	Yes	25	No
6	43	Yes	Yes	42	Yes
7	63	Yes	No	23	No
8	52	No	Yes	44	Yes
9	61	Yes	Yes	48	No
10	44	No	Yes	28	No

In this database age has discrete value so that we first divide age into a rage
Youth if the age <40
Middle_age if the age>40 and age< 58
Senior if the age>58
Similarly we made a rage for Body Mass Index (BMI)
Under weight: BMI is less than 18.5
Normal weight: BMI is 18.5 to 24.9
Overweight: BMI is more than 25

Now we assign fitness value to every attribute value. These values are an assumption based on the medical domain knowledge

Table 2 attribute with recommended values

RID	Attribute	Rage	Recommended value
1	Age	Less than 40	0.1
2		Between 41 to 58	0.3
3		More than 58	0.5
4	Smoking	Yes	0.6
5		No	0.4
6	Hypertension	Yes	0.5
7		No	0.4

8	BMI	Less than 25	0.1
9		Between 26 and 30	0.3
10		More than 40	0.5

On the basis of this fitness value we calculate the fitness value of each record.

Table 2 attribute with recommended and average values

RID	Age	Smoking	Hypertension	(BMI)	Average value
1	0.3	0.6	0.5	0.5	0.475
2	0.5	0.6	0.4	0.1	0.400
3	0.3	0.4	0.5	0.5	0.425
4	0.5	0.6	0.5	0.5	0.525
5	0.3	0.4	0.5	0.1	0.325
6	0.3	0.6	0.5	0.5	0.475
7	0.5	0.6	0.4	0.1	0.400
8	0.3	0.4	0.5	0.5	0.425
9	0.5	0.6	0.5	0.5	0.525
10	0.3	0.4	0.5	0.3	0.375

For example we want to classify the tuple as per the fitness value
Consider the following a rule
Smoking=Yes and Hypertension=Yes Than Heart Disease =Yes

$$\frac{\text{(Sum record having Smoking=Yes, Hypertension=Yes Then Heart Disease =Yes)}}{\text{Sum record having Smoking=Yes and Hypertension=Yes}}$$

$$\frac{0.475+0.525+0.475}{0.475+0.525+0.475+0.525}$$

$$\frac{1.475}{2}$$

$$=0.7375$$

V. EXPERIMENTAL ANALYSIS

We used VB dot net 2013 and SQL server 2010 R2 for experimental analysis. We have taken 5 attribute and 100 records of different patient with corresponding attribute and tested the proposed method. We are different parameter for our Experimental analysis one of them is number of records are correctly classified. We compare the proposed method with Bayesian Classification.

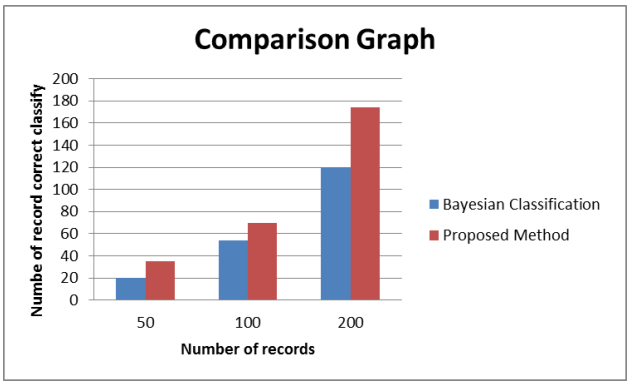

Figure 2. Comparison Graph

VI. CONCLUSION AND FUTURE WORKS

There are several techniques are available to predict heart disease problem like Decision trees, Bayesian classifiers, classification by back propagation, support vector machines, nearest-neighbor classifiers and case-based reasoning classifiers These techniques are compared on basis of Sensitivity, Specificity, Accuracy, Error Rate, True Positive Rate and False Positive Rate. The proposed method reduces number of attribute and reduces complex calculation. In future we also used fuzzy data set to include more desecrate value for the attribute

REFERENCE

1. Mrs. G. Subbalakshmi "Decision Support in Heart Disease Prediction System using Naive Bayes" Indian Journal of Computer Science and Engineering ISSN: 0976-5166 Vol. 2 No. 2 Apr-May 2011
2. Mai Shouman, Tim Turner, Rob Stocker "Using Decision Tree for Diagnosing Heart Disease Patients" Proceedings of the 9-th Australasian Data Mining Conference (AusDM'11), Ballarat, Australia 2011, Australian Computer Society
3. Chaitrali S. Dangare and Sulabha S. Apte Improved Study of Heart Disease Prediction System using Data Mining Classification Techniques International Journal of Computer Applications (0975 – 888) Volume 47– No.10, June 2012
4. Vikas Chaurasia "Early Prediction of Heart Diseases Using Data Mining Techniques 2013. The Authors. Published under Caribbean Journal of Science and Technology ISSN 0799-3757
5. Shamsher Bahadur Patel 1, Pramod Kumar Yadav2 Predict the Diagnosis of Heart Disease Patients Using Classification Mining Techniques "IOSR Journal of Agriculture and Veterinary Science (IOSR-JAVS) e-ISSN: 2319-2380, p-ISSN: 2319-2372. Volume 4, Issue 2 (Jul. - Aug. 2013),
6. N S NITHYA and K DURAISWAMY Gain ratio based fuzzy weighted association rule mining classifier for medical diagnostic interface Sadhana Vol. 39, Part 1, February 2014, Indian Academy of Sciences
7. Ankur Makwana "Decision Support System for Heart Disease Prediction using Data Mining Classification Techniques" International Journal of Computer Applications (0975 8887) Volume 117 - No. 22, May 2015
8. Gunsai Pooja Dineshgar "A Review on Data Mining For Heart Disease Prediction ISSN: 2278 – 909X International Journal of Advanced Research in Electronics and Communication Engineering (IJARECE) Volume 5, Issue 2, February 2016
9. K. Manimekalai "Prediction of Heart Diseases using Data Mining Techniques" International Journal of Innovative Research in Computer and Communication Engineering (An ISO 3297: 2007 Certified Organization) Vol. 4, Issue 2, February 2016
10. S. B. Bhalerao, Dr. B. L .Gunjal "Survey Of Heart Disease Prediction Based On Data Mining Algorithms" Vol-2 Issue-2 2016 IJARIIE-ISSN (O)-2395-4396