IJARCCE

ISSN (Online) 2278-1021
ISSN (Print) 2319 5940

International Journal of Advanced Research in Computer and Communication Engineering
Vol. 5, Issue 4, April 2016

Computer Based-Medical Decision Support System for Prediction of Heart Attack Using Data Mining Techniques

Priyanka N[1], Dr.Pushpa RaviKumar[2]

B.E., (M.Tech),CS & Department, Adhichunchanagiri Institute of Technology, Chikkamagaluru, India[1]

B.E., M.Tech., Ph.D., LMISTE CS & Department, Adhichunchanagiri Institute of Technology, Chikkamagaluru, India[2]

Abstract: Medical field is fast growing with adoption of technology tools. Technology has reduced the burden on medical staff with various software which has made the works quick and easy. Medical software when properly planned and utilized has benefits of minimizing the cost and time of the medical professionals. In this regard Data mining is an effective tool that can prove useful to utilize the information in the database in diagnosing a disease or its appropriate treatment. This aids to provide quality care to the patients. In order to achieve this above objective this paper has been designed to develop Computer Based-Medical Decision Support System for Prediction of Heart Attack using Data Mining technique namely Naïve Bayes. It is a web based questionnaire application. It helps to diagnose the heart disease through the answer given by the user. Thus it assists the practitioner in even assessing complex conditions & arriving at accurate medical decision/diagnosis.

Keywords: Naïve Bayes, decision support, data mining, heart disease.

I. INTRODUCTION

Advances in science & technology have made one's life easy & hence there is lack of physical activity. The comfort provided by the sophisticated instruments has brought about changes in the life style. This has resulted in various metabolic diseases like hypertension, diabetes mellitus, heart diseases, obesity etc. Especially the incidence of heart disease is increasing alarmingly high. The World Health Organization (WHO) has estimated that 12 million deaths occur worldwide, every year due to heart diseases. In 2008, 17.3million people died due to Heart Disease. The World Health Statistics 2012 reports enlighten the fact that one in three adults world-wide has raised blood pressure–a condition that causes around half of all deaths from stroke and heart disease. WHO estimated by 2030, almost 23.6 million will die due to Heart disease [5]. These facts are alarming to the healthcare providers & they have to start focusing on developed & more precise ways to improve the medical attention to prevent & effectively treat the condition. Also concern has to be towards providing these services at low cost to reach all the sectors of the society.

In spite of rich manpower in healthcare sector there is an increasing demand for doctors, nurses etc. Adoption of technological tools in health sector can minimize the time & cost for diagnosis.

This paper is aimed to develop an automated system that can make this cumbersome process simpler which requires less manpower & can help to provide health service to patients at economical prices. It is a prototype Computer Based-Medical Decision Support System for Prediction of Heart Attack System using data mining techniques, namely Decision Trees, Naïve Bayes. This system can successfully predict the patients when have the possibility or risks of getting heart disease. Based on the answers provided by the patient it can generate the appropriate solution.

II. LITERATURE SURVEY

In this section, related literature about Data mining, Strategies of data mining and commonly used data mining techniques like Neural Networks and Decision trees will be reviewed and discussed.

A. DATA MINING

Data mining can be defined as a process which extracts potentially useful hidden data from the large collection of database which is usually hidden or previously not known. Data mining uses two strategies: supervised and unsupervised learning. In supervised learning, a training set is used to learn model parameters whereas in unsupervised learning no training set is used (e.g., k means clustering is unsupervised) [6].

The most commonly used data mining techniques are:

B. NEURAL NETWORKS

Neural Networks consists of three layers: input, hidden and output units (variables).Connection between input units and hidden and output units are based on relevance of the assigned value (weight) of that particular input unit. The higher the weight the more important it is. Neural Network algorithms use Linear and Sigmoid transfer functions[8]. Neural Networks are suitable for training

IJARCCE
ISSN (Online) 2278-1021
ISSN (Print) 2319 5940
International Journal of Advanced Research in Computer and Communication Engineering
Vol. 5, Issue 4, April 2016

large amounts of data with few inputs. It is used when other techniques are not satisfactory.

C. DECISION TREES

Decision trees are simple knowledge representation and they classify examples to a finite number of classes, the nodes are labeled with attribute names, the edges are labeled with possible values for this attribute and the leaves labeled with different classes. Tree shaped structures represent set of decisions. These decisions generate rules for the classification of a data set. Decision trees produce rules that are mutually exclusive and collectively exhaustive with respect to the training data base. Specific decision tree methods include classification and regression trees (CART) and chi square automatic interaction and detection (CHAID)

III. PROBLEM STATEMENT

Heart disease can be prevented by predicted the possibility of its occurrence in future by identifying the risk factors. Heart disease is diagnosed through specific signs & symptoms & through physical examination of the patient. These when clubbed with other investigations like ECG etc. can help a doctor in confirming the diagnosis of a disease. The process if diagnosis is not always easy because it needs proper knowledge, updated information & experience of the doctor for proper decision making. If not it may lead to unnecessary bias & wrong interpretations. Hence human intelligence can be suitably synergized with data modeling analysis tools like data mining. These can help in arriving at appropriate diagnosis & taking proper clinical decisions. Here the information provided by the patient will be utilized to make these clinical decisions with an unbiased approach & thereby reducing the time required & cost spent by the patient.

IV. METHODOLOGY

The main objective of this paper is to develop a Computer Based-Medical Decision Support System for Prediction of Heart Attack System using one data mining modeling technique, namely, Naïve Bayes. It is implemented as web based questionnaire application.

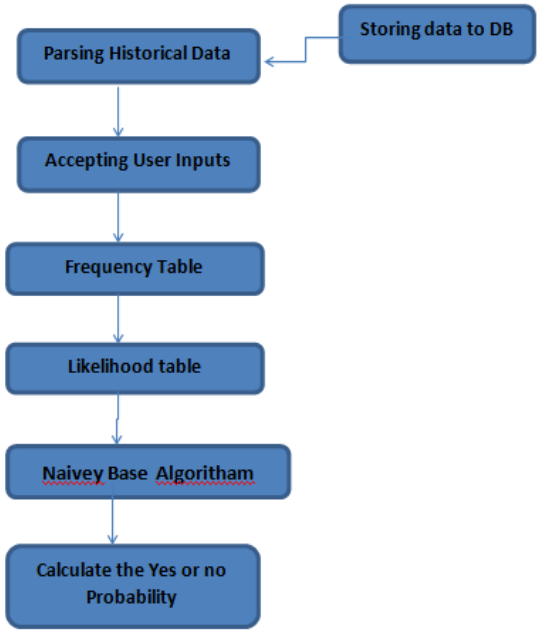
This system ensures quick & easy diagnosis & decision making. Through the answers provided by the user, the system generates result whether the patient is having heart disease or not. It also suggests best possible treatment useful for the condition. It saves time & money to the patient. Thereby ensures quality service to the user. The report of the patient will be generated in 2 ways – Chart & PDF, which can be stored for future reference & documentation.

A. DATA SOURCE

Numerous data about the patient & disease will be collected & stored in Clinical databases. Heart disease is a general term used for all the conditions affecting the heart. It is the leading cause of medical emergencies all over the world. The mortality rate is one in 34 persons every second in United States. Coronary artery disease (CAD), Cardiovascular disease, Cardiomyopathy are the various pathological conditions affecting the heart. Here the heart & blood vessels are affected in various ways, thereby affecting the normal functioning of the heart. Resulting disease leads to various symptoms like chest pain, severe debility, fatigue & in serious conditions death.

The records with medical attributes were obtained from www.archive.ics.uci.edu. The patterns that were conclusive to predict the heart attack was extracted. The records were divided proportionally into 2 datasets training dataset and testing dataset. In order to prevent bias the records were selected randomly.

The attribute “Diagnosis” is assigned with predictable attribute value “1” for patients with heart disease and value “0” for patients with no heart disease. “PatientId” is used as the key; the rest are input attributes. The issues like missing data, inconsistent data & duplicated data have been attended.



Attribute Information:

1) age
2) sex
3) chest pain type (4 values)
4) resting blood pressure
5) serum cholestoral in mg/dl
6) fasting blood sugar > 120 mg/dl
7) resting electrocardiographic results (values 0,1,2)
8) Maximum heart rate achieved
9) exercise induced angina
10) oldpeak = ST depression induced by exercise relative to rest
11) the slope of the peak exercise ST segment
12) number of major vessels (0-3) colored by flourosopy
13) thal: 3 = normal; 6 = fixed defect; 7 = reversable defect

IJARCCE

ISSN (Online) 2278-1021
ISSN (Print) 2319 5940

International Journal of Advanced Research in Computer and Communication Engineering
Vol. 5, Issue 4, April 2016

Attributes type:
1) Real: 1,4,5,8,10,12
2) Ordered:11,
3) Binary: 2,6,9
4) Nominal:7,3,13

Variable to be predicted:
1) Absence (1) or presence (2) of heart disease

B. NAÏVE BAYES ALGORITHM

1. Retrieve all the data from DB associated with the class label, each record associated with the n-dimensional attribute vector X = (x1,x2,x3..)
2. Let DS be a dataset which contain the frequency table associated with the test data
3. Using bayes theorem posterior probability will be calculated P(A|x), from P(A), P(x), and P(x|A). and value which predicated will be independent of the value of other predictors

$$P(A|x) = \frac{P(A|x)P(A)}{P(x)}$$

$$P(A \mid X) = P(x_1|A) \times P(x_2|A) \times \ldots \times P(x_n|A) \times P(A)$$

- P(A|x) is the posterior probability of class (target) given predictor (attribute).
- P(A) is the prior probability of class.
- P(x|A) is the likelihood
- P(x) is the prior probability of predictor

Based on the above equation the probability of each attribute will be calculated and the value with maximum will be the predicated class.

V. RESULTS

Data Set:

Fig 1 Data set of the Heart Disease Prediction System

The data set of the heart attack, which contain the 300records.
Input Process:

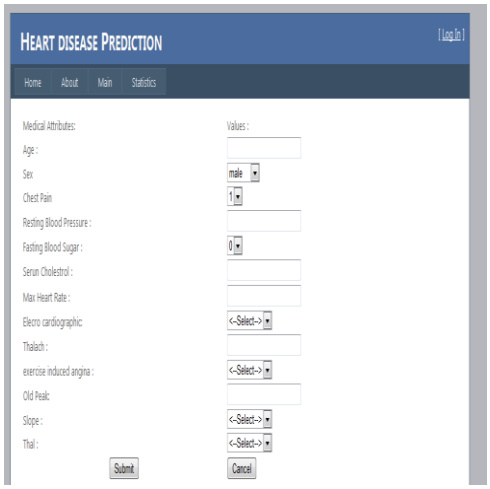


Fig.2 Input Process of the Heart Disease Prediction System

Input process takes the legitimate input from the user.
Prediction Process:

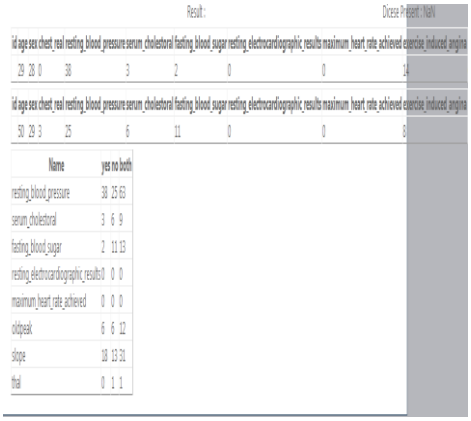


Fig.3 Prediction Process of the Heart Disease Prediction System

This page promote the users to provide medical related information and based on that it will try to find out the posterior probability of attribute by using Naïve Bayes algorithm and also it will show the frequency database of all the attribute.

Statistic Process:

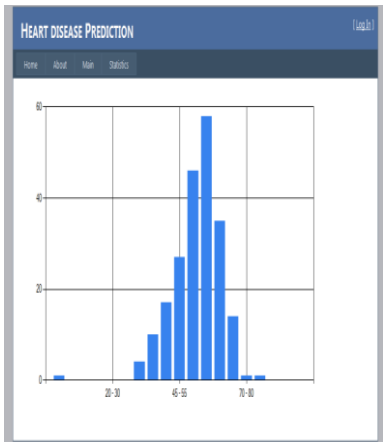


Fig.4 Statistic Process of the Heart Disease Prediction System

DOI 10.17148/IJARCCE.2016.54219

This process will plot the bar chart from the historical data.

VI. CONCLUSION

The paper was planned with an objective to make the clinical decision making in predicting the heart disease easier & quicker. Reliable methods through data mining were adopted to access the information available from the patient. Computer Based-Medical Decision Support System for Prediction of Heart Attack System is developed using Naive Bayesian Classification technique. The hidden knowledge is extracted by the system through heart disease databases.

This system can answer even difficult queries with accurate results. It can not only predict the possibility of heart disease but also can suggest appropriate treatments for the condition. It can generate reports for the hospital & patient use.

AUTHORS ACKNOWLEDGEMENT

I am highly thankful to my hod **Dr.Pushpa RaviKumar** B.E., M.Tech., Ph.D., LMISTE, Professor & Head of the department of CS&E, Adichunchanagiri Institute of Technology, Chikkmagaluru-577102 for her consistent support & guidance all through the paper. I also thank **Dr. Shilpa. S.N.,** Senior consultant, Sriranga Ayureda Chikitsa Mandira, Mysore for her help regarding the usage of medical aspects of heart disease & its aspects. Last but not the least in would like to thank all those who directly or indirectly assisted in the successful completion of the paper.

REFERENCES

[1]. Sellappan Palaniappan, Rafiah Awang Intelligent Heart Disease Prediction System Using Data Mining Techniques, IJCSNS International Journal of Computer Science and Network Security, VOL.8 No.8, August 2008.
[2]. MachineLearningDatabases",http://mlearn.ics.uci.edu/databases/heart-disease/, 2004.
[3]. Chapman, P., Clinton, J., Kerber, R. Khabeza, T., Reinartz, T.,Shearer, C., Wirth, R.: "CRISP-DM 1.0: Step by step data mining guide", SPSS, 1-78, 2000.
[4]. K.Srinivas B.Kavihta Rani Dr. A.Govrdhan Associate Professor,Dept. of CSE Principal andProfessor of CSE 'Applications of Data Mining Techniques in Healthcare and Prediction of HeartAttacks '-(IJCSE) International Journal on Computer Science and Engineering Vol. 02, No. 02, 2010, 250-255
[5]. http://www.worldlifeexpectancy.com/life expectancyresearch
[6]. Obenshain, M.K: "Application of Data Mining Techniques to Healthcare Data", Infection Control and Hospital Epidemiology, 25(8), 690–695, 2004.
[7]. Sellappan, P., Chua, S.L.: "Model-based Healthcare Decision Support System", Proc. Of Int. Conf. on Information Technology in Asia CITA'05, 45-50, Kuching, Sarawak, Malaysia, 2005
[8]. Tang, Z. H., MacLennan, J.: "Data Mining with SQL Server 2005", Indianapolis: Wiley, 2005.
[9]. Obenshain, M.K: "Application of Data Mining Techniques to Healthcare Data", Infection Control and Hospital Epidemiology, 25(8), 690–695, 2004.
[10]. Sellappan, P., Chua, S.L.: "Model-based Healthcare Decision Support System", Proc. Of Int. Conf. on Information Technology in Asia CITA'05, 45-50, Kuching, Sarawak, Malaysia, 2005
[11]. Wu, R., Peters, W., Morgan, M.W.: "The Next Generation Clinical Decision Support: Linking Evidence to Best Practice", Journal Healthcare Information Management. 16(4), 50-55, 2002.
[12]. Sana Shaikh1, Amit Sawant2, Shreerang Paradkar3, Kedar Patil4 ," Electronic Recording System-Heart Disease Prediction System", 2015 International Conference on Technologies for Sustainable Development (ICTSD-2015), Feb. 04 – 06, 2015, Mumbai, India

BIOGRAPHIES

Ms.Priyanka N is a student of Computer Science from, Adhichunchanagiri Institute of Technology, Chikkamagaluru , Presently pursuing M.Tech (CS) from this college. She received B.E from Malnad college of Engineering, affiliated to VTU University, Hassan in the year 2015.

Dr.Pushpa RaviKumar B.E., M.Tech., Ph.D., LMISTE is working as professor and Head, Department of computer science & Engineering, AIT college, Chikkamagaluru. She has 16yr of teaching experience & 5yr of research experience. She has completed her Ph.D from VTU in 2014 as full time research scholar at R.V.College of engineering Bangalore. She has completed her M.Tech from R.V. College of Engineering in the year 2007.She has presented & published many research papers in National & International conferences & journals. Her research interests are Data mining, Neural networks, Social network analysis & Computer networks.

DOI 10.17148/IJARCCE.2016.54219