International Journal of Science and Research (IJSR)
ISSN (Online): 2319-7064
Index Copernicus Value (2013): 6.14 | Impact Factor (2013): 4.438

Future Prediction For Heart Attack Problem Based on Most Appropriate Attribute value

Amit Kisan Pagare[1], Vijay Kumar Verma[2]

[1] Lord Krishna College of Technology Indore M.P. India, M.Tech IV Sem

[2] Assistant Professor CSE, Loed Krishna College of Technology Indore M.P India,

Abstract: -Superior data mining techniques are developed and used to discover hidden pattern form historical data. New Models are developed from these techniques will be useful for medical practitioners to take successful decision. Diagnosis of heart attack is a significant task in medical science. The term Heart attack includes the various diseases that involve the heart attack problem. The exposure of heart attack problem from different symptoms is an important issue for predicting heart attack problem. In this paper have taken 10 attribute which are responsible for the heart attack problem. We convert the given test data set into binary format with a possible conditions for heart attack. In seconds step we divide the data set and apply most appropriate condition on each attribute Find pair for each attribute which satisfy the condition. We repeat the process for grouping the attribute until no more grouping is possible. At last we find the most common attribute and calculate how much percentage data is accurately classified

Keyword: fitness value, heart disease, attribute, data mining, classifier.

1. Introduction

In our everyday life there are several example exit where we have to analyze the historical data for example a bank loans officer needs analysis of her data in order to learn which loan applicants are"safe"andwhichare "risky" for the bank[1,2]. Similarly for a medical researcher it is necessary to analyze breast cancer data in order to predict specific treatmentsfor a patient. These are some examples where the data analysis task required before taking any decision. Classification is a data analysis process, where a classifier is constructed to predict class, for bank loan exampleprediction class is "yes" or "no" Similarly fora medical researcher prediction class is "treatment A," "treatment B," or "treatment C" for the medical data[3,4].

Classification process can be divided into two parts (1)Learning: Training data are analyzed by a classificationalgorithm.Here, the class label attribute is loan decision, and the learned model or classifier isrepresented in the form of classification rules. (2) Classification: Test data are used to estimatethe accuracy of the classification rules. If the accuracy is considered acceptable, the rules canbe applied to the classification of new data tuples

2. Lecturer Review

In 2010 SunitaSoni and O.P. Vyas proposed "Associative Classifiers for Predictive Analysis in Health Care Data Mining". They used a combined approach that integrates association rule mining and classification rule mining. The integration is done by focusing on mining a special subset of association rules and then classification is being performed using these rules. Given the readability of the associative classifiers, they are especially fit to applications were the model may assist domain experts in their decisions. They also introduce thatcombining the advanced association rule mining with classifiers gives a new type of associative classifiers with small refinement in the definition of support and confidence that satisfies the validation of downward closure property[4,5].

In 2011 Mrs.G.Subbalakshmi and Mr. K. Ramesh Proposed "Decision Support in Heart Disease Prediction System using Naive Bayes ". They proposed 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. They implement the system by using web based questionnaire application. This systemhelps to train nurses and medical students to diagnose patientswith heart disease. Decision Support in Heart Disease Prediction System is developed using Naive Bayesian Classification technique. The system extracts hidden knowledge from a historical heart disease database. This is the most effective model to predict patients with heart disease. This model could answer complex queries, each with its own strength with respect to ease of model interpretation, access to detailed information and accuracy. DSHDPS can be further enhanced and expanded[6,7,15].

In 2011 MaiShouman, Tim Turner, Rob Stocker proposed "Using Decision Tree for Diagnosing Heart Disease Patients "Heart disease is the leading cause of death in the worldover the past 10 years. Researchers have been usingseveral data mining techniques to help health careprofessionals in the diagnosis of heart disease. DecisionTree is one of the successful data mining techniques used.However, most research has applied J4.8 Decision Tree,based on Gain Ratio and binary discretization. Gini Indexand Information Gain are two other successful types ofDecision Trees that are less used in the diagnosis of heartdisease. Also other discretization techniques, votingmethod, and reduced error pruning are known to producemore accurate Decision Trees. This research investigatesapplying a range of techniques to different types ofDecision Trees seeking better performance in heartdisease diagnosis. [7,12]

In 2012 M. Akhiljabbar, Dr.Priti Chandra , Dr. B. L. Deekshatuluc Heart Disease Prediction System using Associative Classificationand Genetic Algorithm Associative classification is a recent and rewarding technique which integrates association rule mining andclassification to a model for prediction and achieves maximum accuracy. Associative classifiers are especially fit toapplications where maximum accuracy is desired to a model for prediction. There are many domains such asmedical where the maximum accuracy of the model is desired. Heart disease is a single largest cause of death in. They proposed an efficient associative classification algorithm using genetic approach for heart disease prediction. The main motivation for using genetic algorithm inthe discovery of high level prediction rules is that the discovered rules are highly comprehensible, having highpredictive accuracy and of high interestingness values.In this paper, we proposed a system for heart disease prediction using data mining techniques. In our feature work we plan to reduce no. of attributes and to determine the attribute which contribute towards the diagnosis of disease using genetic algorithm[8,13]

In 2013 V.Krishnaiah, Dr. G. Narsimha, Dr.N.Subhash Chandra Diagnosis of Lung Cancer Prediction SystemUsing Data Mining Classification Techniques Cancer is the most important cause of death forboth men and women. The early detection of cancer can behelpful in curing the disease completely. So the requirement oftechniques to detect the occurrence of cancer nodule in earlystage is increasing. A disease that is commonly misdiagnosed islung cancer. Earlier diagnosis of Lung Cancer saves enormouslives, failing which may lead to other severe problems causingsudden fatal end. Its cure rate and prediction depends mainlyon the early detection and diagnosis of the disease. One of themost common forms of medical malpractices globally is anerror in diagnosis. Knowledge discovery and data mininghave found numerous applications in business and scientificdomain. Valuable knowledge can be discovered fromapplication of data mining techniques in healthcare system. Inthis study, we briefly examine the potential use of classificationbased data mining techniques such as Rule based, Decisiontree, Naïve Bayes and Artificial Neural Network to massivevolume of healthcare data. The healthcare industry collectshuge amounts of healthcare data which, unfortunately, are not“mined” to discover hidden information. For datapreprocessing and effective decision making One DependencyAugmented Naïve Bayes classifier (ODANB) and naive creedalclassifier 2 (NCC2) are used[9,14].

3. Problem Statement

The heart is very important part of human body. Which pumps blood into the entire 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.Some of the risk factors for heart disease areSmoking, Cholesterol,and Blood pressure. Diabetes, Sedentary life style, Eating Habits, Stressetc[10,11].

4. Proposed Method

In the proposed method we use the concepts of fitness number.In the proposed approach we convert the given data set into binary format as per the given condition for heart attack. We divide the large data set into number of parts. We perform simple intersection calculation for every part of data set and finally we consider only those item set which satisfy the given minimum support fitness valueConsider a simple example

Table 1 Simple patient’s database

S.NO	1	2	3	4	5	6	7	8	9	10
1	63	1	145	233	1	2	150	2	3	6
2	67	1	160	286	0	2	108	1	2	3
3	67	1	129	229	0	2	129	2	2	7
4	37	1	130	250	0	0	187	3	3	3
5	40	0	130	204	0	2	172	1	1	3
6	56	1	126	236	0	0	178	0	1	3
7	62	0	140	268	0	2	160	3	3	3
8	57	0	120	354	0	0	163	0	1	3
9	63	1	130	254	0	2	147	1	2	7
10	53	1	140	203	1	2	155	3	3	7

We have taken 10 attribute which are responsible for the heart attack problem. We convert the given test data set into binary format with a possible conditions forheart attack.

Table 2: Description of attribute for patient’s database

S.NO	Attribute
1	Age
2	Sex
3	Blood pressure
4	Cholesterol
5	Fasting bloodsugar
6	Resting ECG
7	Thalach value
8	Old peak
9	Slope
10	Thal

Possible conditions forheart attack
Age>45, BP>120, Cholesterol range>240, FBS>120,Resting ECG>1, Thalach value>100 Beats/Minute,Old peak>0,Slope>=2,ThalValue>3, As per the given conditions we convert the test data set into binary format. 1 represent for condition true and 0 for condition false.

Table 3: binary conversion as per the given condition

S.NO	1	2	3	4	5	6	7	8	9	10
1	1	1	1	0	0	1	1	1	1	1
2	1	1	1	1	0	1	1	1	1	1
3	1	1	1	0	0	1	1	1	1	1
4	0	1	1	1	0	0	1	1	1	1
5	0	0	1	0	0	1	1	1	0	1
6	1	1	1	0	0	0	1	1	0	1
7	1	0	1	1	0	1	1	1	1	1
8	1	0	0	1	0	0	1	0	0	1
9	1	1	1	1	0	1	1	1	1	1
10	1	1	1	0	0	1	1	1	1	1

5. Proposed Algorithm

Input:
A transaction database D

Appropriate Condition (AC)

Output:
Pair of Attribute satisfy the given conations for heart attack

Method:
1) Scan the database **D** and partition the transactiontable into equal size**.**
2) Find Appropriate Condition (AC) for each attribute .
3) Consider only those attribute which satisfy the given minimum Appropriate Condition remainingattribute.
4) Convert the dataset into binary format 1 forAppropriate Condition and 0 for others.
5) To discover the pair of two attribute use Join and perform logical AND .
6) To determine attribute set,join them and perform logical AND operation .
7) The algorithm iterates to find upto pair of n- attribute item sets
8) From each pair find out pair of n-attribute itemsets. These pair of attribute are said to local attribute which Appropriate Condition .

6. Architecture the Proposed Method

First we convert data set into binary format according toAppropriate Condition conditions. In seconds step we divide the data set into number of parts .Find pair for each attribute which satisfy the condition . We repeat the process for grouping the attribute until no more grouping is possible. At last we find the most common attribute in all parts andcalculate how much percentage data is accurately classified. The working process of proposed model is show the figure 1

7. Experimental Evaluation

We use VB dot net 2010 as front end and SQL server as back end for data base. .All the experiments were performed on a i3 4M Cache, 2.50 GHz Intel PC machine with 2 gigabyte main memory, running Microsoft Windows 7.To evaluate the performance Real life dataset is used. We have implemented three algorithms first one Bayesian classification algorithm, second is weight associated classifiers and third is our proposed method. We have taken 10 attribute which are mainly responsible for heart attack problem. We have taken the data from a pathology laboratory. We perform experiments using 1000, 2,000 and 5000 records.

8. Comparison and Graphs

We have taken number of record and percentage of record classify accurately. In figure 2 show comparison graph. it is clearly show that the percentage of classify correct records by proposed method is more as compared to the previous methods.

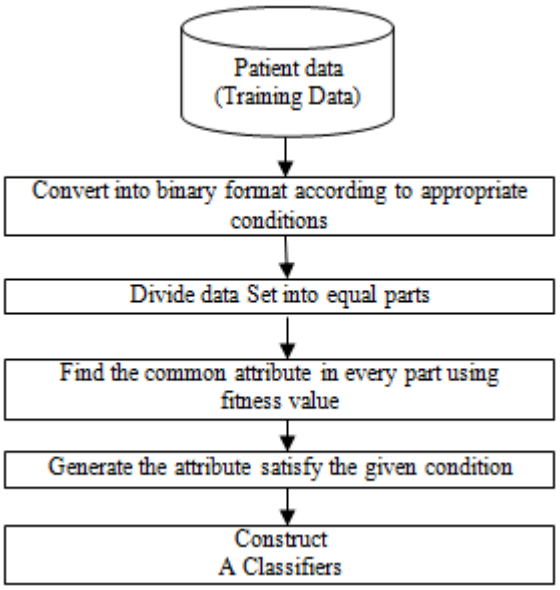

Figure 1: Architecture of proposed system

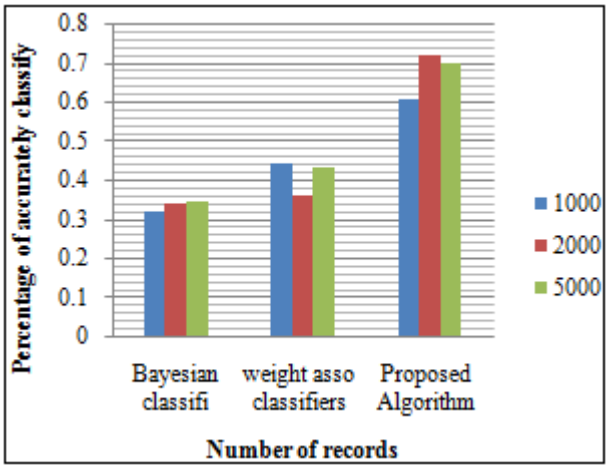

Figure 2: Percentage of records classify correctly

9. Conclusion and Future Work

From the experiment it clear that proposed method is more accurately classify the recodes as compared to pervious method. Proposed method considers all attribute given to heart attack condition.Proposed method is also simple to under stands and calculation is also easy. We have taken only ten attribute which are mainly responsible for heart attack , in future we have consider more then ten attribute which are also responsible for heart attack.

References

[1] S. Sandhiya P. Pavithra and A. Vidhya"Novel Approach for Heart Disease verdict Using Data Mining Technique"International Journal of Modern Engineering Research (IJMER) www.ijmer.com ISSN: 2249-6645
[2] DivyaTomar and SonaliAgarwal A survey on Data Mining approaches for Healthcare International Journal of Bio-Science and Bio-Technology Vol.5, No.5 (2013), pp. 241-266
[3] N.Suneetha CH.V.M.K.Hariand V.Sunil Kumar Modified Gini Index Classification: A Case Study Of Heart Disease Dataset (IJCSE) International Journal on

Computer Science and Engineering Vol. 02, No. 06, 2010, 1959-1965

[4] SunitaSoni O.P.VyasUsing Associative Classifiers for Predictive Analysis in Health Care Data Mining International Journal of Computer Applications (0975 – 8887) Volume 4 No.5, July 2010

[5] Anushya& A. PethalakshmiA Comparative Study of Fuzzy Classifierswith Genetic On Heart Data. International Conference on Advancement in Engineering Studies & Technology, ISBN 978-93-81693-72-8, 15th JULY, 2012, Puducherry

[6] Mariammal.D, Jayanthi.S, Dr. P.S.K.PatraMajor Disease Diagnosis and Treatment Suggestion System Using Data MiningTechniques International Journal of Advanced Research in Computer Science & Technology IJARCST All Rights Reserved 338 Vol. 2 Issue Special 1 Jan-March 2014 ISSN: 2347 - 8446 (Online) ISSN: 2347 - 9817

[7] V. Manikantan& S. LathaPredicting the Analysis of Heart Disease Symptoms Using Medicinal Data Mining Methods International Journal on Advanced Computer Theory and Engineering (IJACTE) ISSN 2319 – 2526, Volume-2, Issue-2, 2013.

[8] Milan Kumari andSunilaGodaraComparative Study of Data Mining Classification Methodsin Cardiovascular Disease Prediction IJCST Vol. 2, Issue 2,June 2011 ISSN 2229 - 4333

[9] NidhiBhatla and KiranJyotiA Novel Approach for Heart Disease Diagnosis using Data Mining and Fuzzy Logic International Journal of Computer Applications (0975 – 8887) Volume 54– No.17, September 2012

[10] V.Jaya Rama krishniahD.V.ChandraSekarandDr. K. Ramchand H RaoPredicting the Heart Attack Symptomsusing Biomedical Data Mining Techniques V Volume 1, No. 3, May 2012 ISSN2278-1080The International Journal of Computer Science & Applications (TIJCSA)

[11] Classification of Data for Heart Disease Prediction System Using MLP International Journal in Multidisciplinary and Academic Research (SSIJMAR) Vol. 2, No. 5, August – September 2013 (ISSN 2278 – 5973)

[12] 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 Copyright © 2011, Australian Computer Society

[13] M. Akhiljabbar, Dr B.L Deekshatulu, Dr.PritiChandra ”Heart Disease Classification Using Nearest Neighbor Classifier With Feature Subset SelectionComputer Science and Telecommunications 2013|No.3(39) ISSN 1512-1232.

[14] Diagnosis of Lung Cancer Prediction System UsingData Mining Classification TechniquesV. Krishnaiah (IJCSIT) International Journal of Computer Science and Information Technologies, Vol. 4 (1) , 2013, 39 – 45

[15] Mrs.G.Subbalakshmi and Mr. K. Ramesh Decision Support in Heart Disease Prediction System usingNaive BayesJournal of Computer Science and Engineering (IJCSE)ISSN : 0976-5166 Vol. 2 No. 2 Apr-May 2011