IJournals International Journal of Software & Hardware Research in Engineering
ISSN No: 2347-4890 Volume 2 Issue 12, December-

ANALYSIS AND DECISION MAKING OF MISDIAGNOSED DISEASE BY NAÏVE BAYESIAN ALGORITHM IN HEALTHCARE

Aakanksha Jaiswal[1]; Mr. Kailash Chandra[2]
PG scholar, CSE department[1]; Assistant professor, CSE Department[2]
Acropolis Technical Campus, Indore, MP[1,2]

ABSTRACT

With the help of this analysis we are going to introduce website, in which we identify the disease name with the symptoms and after that we work on that particular major disease. This analysis is very reliable for the Doctors and also for the patient who is suffering between those diseases. We are going to prove that this is work as a building blocks for a healthcare system and it give us reliable information about the disease with the help of given symptoms.

At first we are identifying the disease probability on the basis on selected symptoms which is done by the doctor, after that we using kinds of filtering techniques to find the disease probability according to requirement of doctor on the basis of survey of symptoms of patient.

General Terms
Diseases, Mathematical model programming, symptoms

Keywords
Data Mining, Common Disease, symptoms

INTRODUCTION

We all know that a little mistake in any test to identify our disease is dangerous for our health, death of many people are common now a days. Every year we heard about many new disease and there causes and about death this ratio is increase day by day and one day it will be more than the birth rate of our country, and if do not take any step against it than it will be become a major problem for our country. Medical errors cause tens of thousands of deaths in U.S. hospitals each year, more than from highway accidents, breast cancer, and AIDS combined [1].

Based on a study of 37 million patient records and average of 195000 people are in the world. We found that people get scared about their health in a wide range and they are also suffering from various kind of common diseases, and then this become a big issue for all of them and then they take a very serious decision about the disease they have been suffered from. This is kept in mind as a serious issue and we have used data mining as a tool to overcome this issue [2].

Every day we know about the new rising diseases in the world and we are unable to take care of all the diseases and to solve them and cure them so that's why we are using data mining technology, it is based on the user oriented approach to novel and hidden patterns in the data.

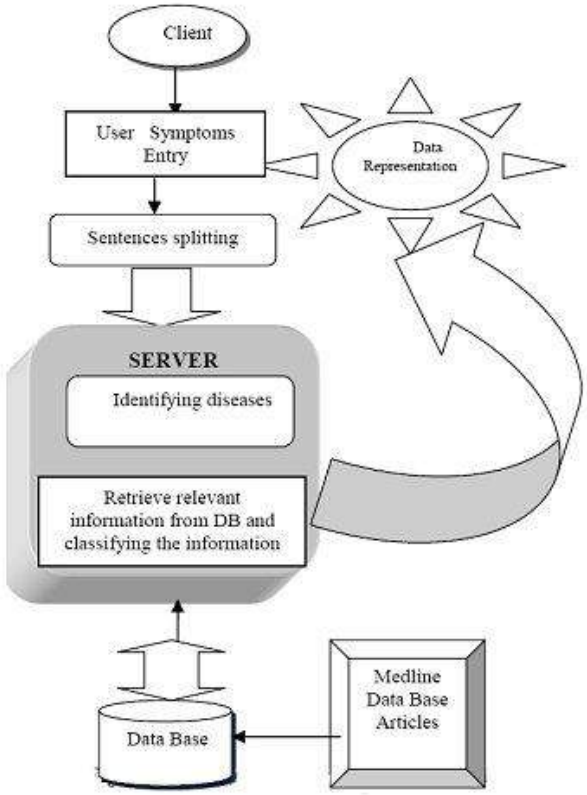
This technology is very efficient for finding appropriate disease on the basis of given symptoms and data. Data Mining has great potential for exploring the meaningful and hidden patterns in the data sets at the medical domain, these methods can be used for the medical and diseases diagnosis. Data mining is more preferable for the huge amount of data, large amount of data always work by the data mining algorithm and that will be used to perform classification, prediction, estimation or other data mining tasks[3].

We all know that health issue is very big issue now a day's that's why healthcare organization update every day a collect a large amount of information about each and every disease. So here we introduce such software which is work on daily basis of disease updating techniques in electronic manner and reduces the manual work that why we work fast and efficient also [4].

We have various ways to cure many diseases today because of medical services. Many hospital and doctors loses their reputation because of misguidance of the diseases, various kind of diseases have the same type of symptoms that's why that kind of problems occurs and doctors are not give a proper treatment to the patient so that's why we are going to introduce a new system which is very helpful for the doctors to know about the

appropriate diseases. In this analysis we are going to feed the detail about every disease in these systems which give us a proper direction [5].

Some of the hospitals are doing fraud in the basis of major disease and give us a costly treatment and because of high cost many poor patient are unable to take that the treatment, so reduce these kind of fraud we introduce this type of software system [6]. Actual work of this system is to *know about the patient problems, symptoms and give them proper treatment by the doctor.*



LITRATURE SERVEY

Donald (1981), Observed that the application of computer to an organization such as making use of data mining in medical department provides individuals with better services, high quality of products and greater efficiency. He went further to say that the use of computer would avoid wastage of time and files. Improved efficiency and productivity reduce cost of goods and services.

According to Rosenberg J. (1997), Computerization is a computer application to any activities formerly done by hand or without the use of computers. It can be drawn from the above that in data mining for medical record system of a hospital, the fundamental devices requires is the computer[7].

Foster O. (1992), Expert System in Health for Developing Countries, Stated that the technology is applied to help find the data of patient been treated in other to cure the present disease. When all the above is considered you say that health care industry is a technology industry. Modern medical practice is based on vast amount of scientific investigation the evaluation of computer in healthcare industry is taking its time. It has advance to a great level in the aspect of data handling, stating that computers are supposed to be used to fill any record about a patient[8].

Ms. M. Thangamani - The conclusions of our study suggest that domain specific knowledge improves the results. Probabilistic models are stable and reliable for tasks performed on short texts in the medical domain. The representation techniques influence the results of the ML algorithms, but more informative representations are the ones that consistently obtain the best results. The source data is from the wed and identifying then classifying the data on the web is a challenge but bringing valuable information in future it has the capability in framework model.

Boris Milovic and Milan Milovic - Data mining has great importance for area of medicine, and it represents comprehensive process that demands thorough understanding of needs of the healthcare organizations. Knowledge gained with the use of techniques of data mining can be used to make successful decisions that will improve success of healthcare organization and health of the patients. Data mining requires appropriate technology and analytical techniques, as well as systems for reporting and tracking which can enable measuring of results. Data mining, once started, represents continuous cycle of knowledge discovery. For organizations, it presents one of the key things that help create a good business strategy. Today, there have been many efforts with the goal of successful application of data mining in the healthcare institutions. Primary potential of this technique lies in the possibility for research of hidden patterns in data sets in healthcare domain.

Arun George Eapen - A comparative study was conducted in this project, for two types of medical databases. Results have shown that most of decision tree based methods implemented have Outperformed the base case we used i.e. WEKA's ZeroR method. An added advantage of decision tree based methods is that it is easier to produce interpretability for the medical practitioners and may help in both the validation of the method and in developing further knowledge of the problem. Also in this chapter the CPU time of some of the experiments were presented.

Dr. Rizwan Beg - With the help of this study we can conclude that in spite of having a large amount of medical data, it lacks in the quality and the completeness of data because of which highly sophisticated data mining techniques are required to build up a efficient decision support system. Even then the overall reliability and generalization capability is still in question. We have to build systems which not only are accurate and reliable but reduce cost of treatment and increase patient care. At the same time we have to build systems which are understandable and which could enhance human decisions. Another major concern which was observed was there insignificant study for proposing treatment plans for patients. Though there is some research but we need a significant amount of study dedicated to this because data mining techniques have shown significant success in prediction and diagnosis of diseases and especially heart diseases, hence we could hopefully use these techniques

L.Sathish Kumar and Mrs.A.Padmapriya - In this paper the problem of constraining and summarizing two algorithms of data mining used in the field of medical prediction are discussed. The focus on using two algorithms and combinations of several target attributes for intelligent and effective common disease diagnosis using data mining. For diagnosis common disease, significantly 12 attributes are listed and with basic data mining technique other approaches e.g. classification and Association Rules, soft computing approaches etc. can also be incorporated. The outcome of diagnosis data mining technique on the same dataset reveals that Decision Tree outperforms and some time Bayesian classification is having similar accuracy as of ID3 but other predictive methods like Neural Networks, Classification based on clustering are not performing well. The second conclusion is that the accuracy of the ID3 and Neural Network. In this paper we proposed the procedure for retrieval of dataset with, relevant fields using ID3 algorithm. Unlike previous works it is based on the individual diagnosis for specific symptoms of the disease. This paper concluded with the individual retrieval of dataset that predicts the diagnosis on the whole.

N.Satyanandam1 and Dr.Ch.Satyanarayana - This paper appraised contemporary art in data mining and machine learning approaches. Assessment illustrates that there is huge progress in utilization of DM and ML for healthcare systems. Along with that it is also clear and noticed that the new dimensions in usage of DM, ML and invention of new approaches and strategies give a greater scope for research in healthcare systems. By seeing the growing fame it extracts sentences from published medical papers that mention diseases and treatments, identifies semantic relations that exist between diseases and treatments, which would predict the usability and maintainability in an efficient manner. So we are sanguine regarding future work in this particular direction [9].

PROPOSED METHODOLOGY

We implement the current concept by *naive Bayesian algorithm. This algorithm says:-*

The probability of any event is the ratio between the value at which an expectation depending on the happening of the event ought to be computed, and the chance of the thing expected upon its happening.

To understand and code for the naive Bayesian algorithm, we will do some math to understand the procedure. The primary equation for the Bayes rule is:

$$P(A|B) = (P(B|A) * P(A)) / P(B)$$

This state mathematically that the posterior probability or probability of future occurrence can be calculated by the product of previous belief P(A) and the likelihood of B if A is true; i.e., P(B|A). P(A|B) is called posterior probability, P(A) is called prior probability, and P(B) is normalization constant. This equation enables us to calculate the probability that A would occur providing that B has happened.

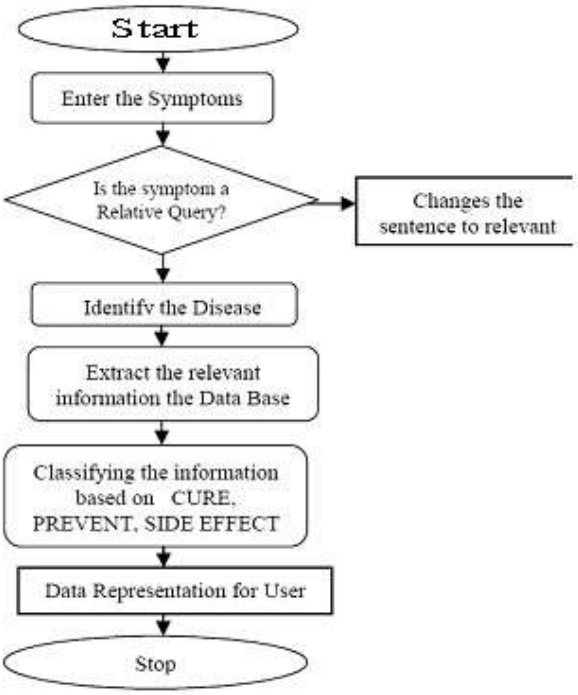


Figure 2 Flow chart for medical disease classification

APPLICATION DOMAIN

1. This analysis work will be carried out on data mining for medical record system.
2. Enhancing databases with functionalities.

CONCLUSION

To analyze all the system and techniques we conclude that in to the medicine's area data mining plays an important role, which is also a demand on healthcare organization. A complete cycle of knowledge discovery in represented by the data mining, in data mining we have abstract knowledge of the disease and give an appropriate solution in the form of disease name.

REFERENCES

[1] S. Novichkova, S. Egorov, and N. Daraselia, "MedScan, A Natural Language Processing Engine for MEDLINE Abstracts,"Bioinformatics, vol. 19, no. 13, pp. 1699-1706, 2003.

[2] D.Senthil Kumar, DECISION SUPPORT SYSTEM FOR MEDICAL DIAGNOSIS USING DATA MINING march 2011.

[3] Sunitha soni, Associate professor PREDICTIVE DATAMINING FOR MEDICAL DIAGNOSIS AN OVERVIEW OF HEART DISEASE March 2011.

[4] Asha Rajkumar, G.Sophia Reena, Diagnosis Of Heart Disease Using Data mining Algorithm, Global Journal of Computer Science and Technology 38 Vol. 10 Issue 10 Ver.1.0 September 2010.

[5] Hall, J., First, make no mistakes. The New York Times(2009).

[6] Candelieri, A., Dolce, G., Riganello, F., & Sannita, W. G. (2011). Data Mining in Neurology. In Knowledge-Oriented Applications in Data Mining (pp. 261-276). InTech.

[7] Blake, C.L., Mertz, C.J.: "UCI Machine Learning Databases",http://mlearn.ics.uci.edu/databases/heart disease/,2004.

[8] Osigwelem K.U : Electronic Data Processing And Communication Foundation, Officeprint Publishers 2008.

[9] N.Satyanandam,Dr.Ch.Satyanarayana,Md.Reyazuddin and Amjan Shaik "Data Mining, Machine Learning Approaches and Medical Diagnose Systems: A Survey" International Journal of Computer and Organization Trends" volume 2, Issue 3,2012.

