Health Care Analysis using Data Mining
Dhinesh Babu L.D, Mohan Sai Kumar P, Manish G
Department of Information Technology and Engineering, VIT University, Vellore, Tamil Nadu
*Corresponding author: E-Mail: mohan.saiikumar2013@vit.ac.in

ABSTRACT

Data Mining is one of the most motivating area of research that is become increasingly popular in health organization. Data Mining plays an important part to uncover new patterns in medicinal services association which thusly accommodating for all the parties associated with this field. This project intend to build a diagnostic model of the various diseases based on the symptoms by using Data Mining technique such as classification in health domain. In this project, we are going to use algorithms like Decision Tree, Naive bayes that can be utilized for health care diagnosis. Performances of the classifiers are compared to each other to find out highest accuracy. This also helps us to find out persons who are affected by the infection. The test information is completed utilizing some benchmark information.

KEY WORDS: Health Care, Data Mining, Analysis, Disease.

1. INTRODUCTION

Liver Disease has become the primary reason of death in people in developed Countries. The most effective way to reduce liver disease deaths is to detect it earlier. Early diagnosis needs an accurate and reliable diagnosis procedure that can be used by physicians to distinguish liver tumors. In this project data mining methods that have been applied in liver disease dataset to know people who are affected by disease and problems are surveyed. We will represent the proper analysis of data mining techniques applied on data sets. i.e., accuracy of those dataset.

With this project, we will implement data mining algorithms to classify the given data to determine which algorithm is best suitable to classify it and to find out whether the person is having the disease or not. Data Mining also inspires many practical solutions that maximize resource utilization and prolong the lifespan of the system.

2. METHODS

In data mining, naive Bayes classifiers are a family of simple probabilistic classifiers based on applying Bayes' hypothesis with solid (credulous) autonomy presumptions between the components. Innocent Bayes has been studied extensively since the 1950's. It was introduced under a different name into the text retrieval community in the early 1960's, and remains a popular (baseline) method for text categorization, the problem of judging documents as belonging to one category or the other (such as spam or real, games or governmental issues, and so on.) with word frequencies as the elements.

With suitable pre-processing, it is competitive in this domain with more advanced methods including bolster vector machines. It likewise discovers application in programmed medicinal analysis. Credulous Bayes classifiers are highly scalable, requiring a number of parameters direct in the quantity of factors (elements/indicators) in a learning issue.

Greatest probability preparing should be possible by assessing a shut frame expression, which takes straight time, instead of by costly iterative guess as utilized for some different sorts of classifiers. In the measurements and software engineering writing, Naive Bayes models are known under an assortment of names, including straightforward Bayes and autonomy Bayes. Every one of these names reference the utilization of Bayes' hypothesis in the classifier's choice administer, yet innocent Bayes is not (really) a Bayesian strategy.

3. RESULTS
4. CONCLUSION

Here we will end up our project by implementing stacking of naive bayes and decision tree algorithm in Anaconda navigator by using data mining model created by us. This model can be created for any algorithm that are related to data sciences. We finally made visualisations of data collected by health care analysis through data mining.

REFERENCES