UNIVERSAL HEALTHCARE SYSTEM

[1]Shivom kaul, [2]Abhishek Anand, [3]Dishant Desai,[4]Sohail Ahmed,[5]Shirish Sabnis

[1,2,3,4] Dept of I.T, Rajiv Gandhi Institute of Technology, Mumbai ,India
[5]Assistant Professor, Dept of I.T, , Rajiv Gandhi Institute of Technology, Mumbai ,India

Abstract : The healthcare industry collects a large amount of data that is not properly mined and not optimally used. Discovery of these hidden patterns and relationships often goes unexploited. Information and communication technologies (ICTs) have great potential to improve health in both developed and developing countries by enhancing access to health information and making health services more efficient; they can also contribute to improving the quality of services and reducing their cost. Patient information systems, for example, have the ability to track individual health problems and treatment over time, giving insight into optimal diagnosis and treatment of the individual as well as improving the delivery of services. This is particularly useful in chronic diseases, such as diabetes and cardiovascular diseases, as well as health services for mothers and children, where a record of health and treatment is required within a given time period. Our research focuses on this aspect of Medical diagnosis by learning pattern through the collected data of diabetes and heart diseases and to develop intelligent medical decision support systems to help the doctors. Analyzing data in patient information systems can lead to a new insight and understanding of health and disease, both chronic and acute. In this paper, we propose the use of decision trees ID3 algorithm classify these diseases and compare the effectiveness, correction rate among them.

Index Terms - **Active learning, decision support system, data mining, medical engineering, ID3 algorithm, CART algorithm, C4.5 algorithm.**

I. INTRODUCTION

The major challenge in the healthcare industry is Poor clinical reports that are not up to the mark and can lead to disastrous results which is unacceptable. Keeping these data personalized rather than anonymous is facilitated by using electronic systems which can more easily store, access, analyze and share data easily. As in the usual way, the collection of such data is on paper and the registration book. Advanced data mining modelling techniques can help in maintain and generating online records and help us to achieve the objective of including IT facilities into medical domain. The universal health care system, which will be developed to assist clinicians in the diagnostic process, often based on static data that can be obsolete. The concept of universal health care system is very wide, because there are many different approaches and a wide range of areas in which decisions are made. Even the most technologically advanced hospitals in India do not have such software, which predicts the disease through data extraction methods. There is a huge amount of unused data, which can be converted into useful information. Medical Diagnosis, as is well known, the subjective; it depends on the doctor that makes the diagnosis. Secondly , and most importantly, the volume of data that must be analyzed in order to make a good forecast and good analysis.

II. Implementation Details

A. ID3 Algorithm

Itemized Dichotomize 3 algorithm or better known as ID3 algorithm was first introduced by J.R Quinlan in the late 1970's. It is a greedy algorithm that selects the next attributes based on the information gain associated with the attributes. The information gain is measured by entropy, ID3 algorithm prefers that the generated tree is shorter and the attributes with lower entropies are put near the top of the tree.ID3 algorithm is an example of Symbolic Learning and Rule Induction. It is also a supervised learner which means it looks at examples like a training data set to make its decisions. It was developed by J. Ross Quinlan back in 1979. It is a decision tree that is based on mathematical calculations.

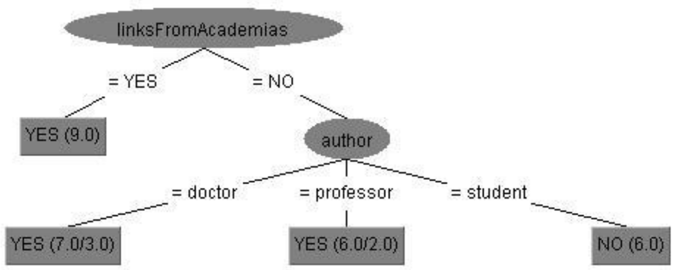


Figure 1: Decision Tree

The system uses the crystal-DM (the standard process data mining) for building data mining models. It consists of six basic stages: understanding, understanding, preparing data, modeling, assessment and deployment. Business understanding of the phase focuses on the understanding of the purpose and requirements from a business perspective, transforming this knowledge in the problem definition and formulation of a preliminary plan to achieve these goals.

The stage of the understanding of the uses the raw data, and income, in order to understand the data, determine its quality, get preliminary conclusions and interesting subsets in the form of hypotheses for hidden information. The data preparation phase builds the final set of data, which will be submitted to the modeling tools. This includes the table entries and selecting the attribute, as well as data cleaning and transformation. At the stage of modeling, selects and uses a variety of methods, and calibrates their parameters for optimal values. Assessment of the phase determines the model, to ensure that it reaches the business objectives. During the deployment phase identifies the tasks that must be performed for the application of the model.Intelligent Data Analysis (DMX), SQL-style Query Language to retrieve data, is used to create and access to model". Tabular and graphical renderings are included in the analysis and interpretation of results.

B. Data sets

Important rules will be retrieved that are useful for understanding the data pattern and behavior of experimental data. The following pattern is retrieved through the use of cart decision tree algorithm.
The rules for the hearts of the data set as follows:

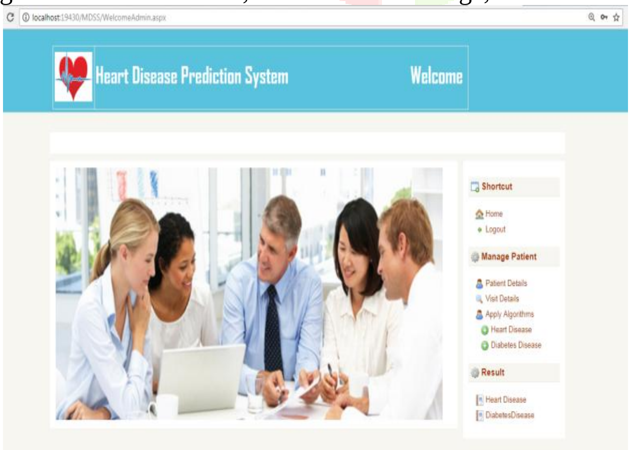
1. Heartdisease(lack of):
Thal=fixed_defect,Number_ships=0, Cholestoral =126-213.
2. heart disease_(availability):-
Thal=normal,Number_ships=0, Old_Peak=0-1.5, Max_heart rate=137-169, Cholestoral=126-213.
3. heart disease_(lack of):
Thal=normal,Number_ships=0, Old_Peak=0-1.5, Max_heart rate=137169,Cholestoral=214-301, the rest=0, the pressure=121-147.

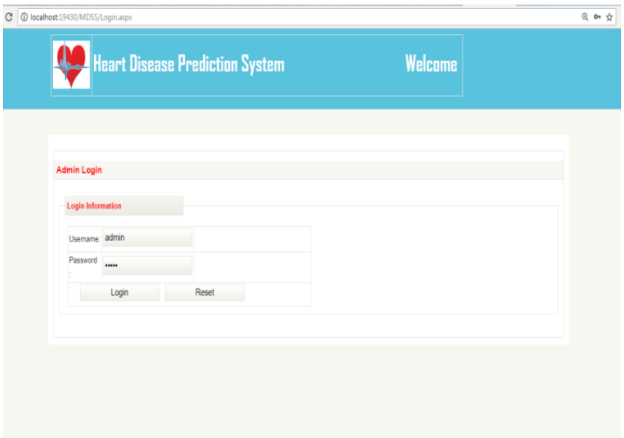
The rules for the hepatitis data sets will be retrieved and some of them are the following
1. Ascites = Yes and histology = No: Live (46.0/1.0)
2. Anorexia = Yes and Proteome > 47 and fatigue = No: Live (8.0)
3. Anorexia = Yes and malaise = Yes and ASCITES = Yes: Live (10.0/2.0)
4. Anorexia = Yes: Die (10.0) : Live (6.0)

Some rules of classification for diabetes data sets as follows:
1. The age of the <= 28 and Triceps skin thickness > 0 and Triceps skin thickness of <= 34 and age > 22, timespreg No <= 3 and plasma gc(2) <= 127: No (61.0/7.0)
2. Plasma gc(2) <= 99 and 2 hours of insulin in the blood serum <= 88 and 2 hours of insulin in the blood serum <= 18 and Triceps skin thickness of <= 21: No (26.0/1.0)
3. The age of the <= 24 and Triceps skin thickness > 0 and the body of the MI <= 33.3: No (37.0) The diastolic blood pressure <= 40 and plasma gc(2) > 130: Yes (10.0)
4. Plasma gc(2) <= 107 Diabetes pf and <= 0,229 and diastolic blood pressure <= 80: No (23.0)
5. no.time spreg <= 6 and plasma gc(2) <= 112 and diastolic blood pressure <= 88 AND Age <= 35: No (44.0/8.0)
6. The age of the <= 30, and the diastolic blood pressure > 72 and the body of the Mi <= 42.8: No (41.0/7.0).

C. Software Implementation

Clinical decisions are often accepted on the basis of "intuition and experience of doctors", and not data, the rich knowledge, hidden in the database. This practice leads to undesirable displacements, errors and excessive medical costs, which affect the quality of service provided to patients. Wu and others suggested that the integration of clinical decisions with the computer account, patients can reduce medical errors, improve patient safety, reduce unwanted changes in practice and improve the results of treatment of patients. This proposal is promising, because the tools of simulation and analysis of data, such as data mining can generate environment, rich in knowledge, which can help to significantly improve the quality of clinical decision.





III. CONCLUSION

Intelligent data analysis techniques and methods used in medical data has led to innovation, standards and decision support systems, significant progress in improving the health of patients and quality of medical services. The results of research and accompanying literature provide a convincing example of the wide acceptance of patient information systems in developed and developing countries. Implementation of this crucial technology is not just reliant on available resources; national health system priorities and institutional will also play key roles in the successful implementation of patient information systems, which contribute to improved patient health, more thorough understanding of disease.

.

REFERENCES

[1]. Qu Jianning, Huang Xiaoming, and Mei Zhechuan, “Study on the application of electronic signature and timestamp technology in medical electronic instrument,” Chinese Journal of health information management, vol. 30, pp. 43-46, 2014.

[2]. Sun Jianyu, and Wang Guoping, “Standardized medical documents are the basic requirements to adapt to the rules of law,” Chinese hospital management, pp. 8, 2003.

[3]. Wang Yu, Interpretation of the Functional Specification and Grading Evaluation Standard of Electronic Medical Record System, Beijing: People's army press, 2012, pp.260-261.

[4]. Jiao Yongben, "Discussion on the legitimacy of electronic medical records and other applications using electronic signature technology, “Journal of China health information management, vol. 6, pp. 23-26, 2009.