Scholars Journal of Applied Medical Sciences (SJAMS) **ISSN 2320-6691 (O)**
Sch. J. App. Med. Sci., 2013; 1(3):158-160 **ISSN 2347-954X (P)**

(An International Publisher for Academic and Scientific Resources)
www.saspublishers.com DOI: 10.36347/sjams.2013.v01i03.005

Review Article

Role of Data Mining Techniques in Healthcare sector in India

[1] Dr.Motilal C. Tayade, [2] Mrs. Pratibha M. Karandikar
[1]Department of Physiology, RMC, Pravara Institute of Medical Sciences (PIMS), Loni, Maharashtra, India.
[2]Mechnical Department, PREC' S College of Engineering,University of Pune , Loni, Maharashtra, India.

Corresponding author
Dr.Motilal C. Tayade
Email: drmctayade@gmail.com

Abstract: It is an essential to interpret the correct diagnosis of patient with the help of clinical examination and investigations. Computer information based decision support systems can play an important role in accurate diagnosis and cost effective treatment. Data mining is the computational process of discovering patterns in large data sets involving methods at the intersection of artificial intelligence, machine learning, statistics, and database systems. Our healthcare sector daily collects a huge data including clinical examination, vital parameters, investigation reports, treatment follow-up and drug decisions etc. But very unfortunately it is not analyze and mine in an appropriate way. Basically data mining is concerned with data processing, identifying patterns and trends in information. Increasing computer data analysis awareness, better online education availability and developing an integrated learning approach among medical professionals will definitely helpful for accurate diagnosis and effective treatment management plan in India.
Keywords: Decision tree, neural network, cluster formation

INTRODUCTION

Today diagnosis of a disease is a vital job in medicine. It is an essential to interpret the correct diagnosis of patient with the help clinical examination and investigations. Computer information based decision support system can play an important role in accurate diagnosis and cost effective treatment. Most hospitals have a huge amount of patient data, which is rarely used to support clinical diagnosis [1]. It is question why we cannot use this data in clinical diagnosis and patient management? Is it really a hectic job? Is it possible to formulate own area based prediction system concerned with specific disease by using data mining techniques? Data mining is the computational process of discovering patterns in large data sets involving methods at the intersection of artificial intelligence, machine learning, statistics, and database systems [2].

Current scenario

Our healthcare sector daily collects a huge data concerned with patients including clinical examination, vital parameters, investigation reports, treatment follow-ups, and drug decisions etc. But very unfortunately it is not analyze and mine in an appropriate way. It is stored either in record room in the form of bunches of paper sheet or occupying hard disc space. The doctors as well as researchers are careless concerned with these valuable data. In government as well as in some organizations the data is handle mainly by statisticians at professional level. The development of automated systems and their accuracy will definitely guide us in future. It will definitely helpful in various diseases management including effectiveness of surgical procedures, medical tests, medication, and the discovery of relationships among clinical and diagnosis data as well employ Data Mining methodologies [3].

Data mining technique

Basically data mining technique is concerned with data processing, identifying patterns and trends in information. In other words, data mining simply means collection and processing data in systemic manner by using computer based programs and subsequent formation of disease prediction or patient management system aid.

Data mining principles have been known around for many years, but, with the advent of information technology, nowadays it is even more prevalent. Data mining is not all about the database software that you are using. You can perform data mining with comparatively modest database systems and simple tools, including creating and writing your own, or using off the shelf software packages. Complex data mining benefits from the past experience and algorithms defined with existing software and packages, with certain tools gaining a greater affinity or reputation with different techniques [4].

This technique is routinely use in large number of industries like engineering, medicine, crime analysis, expert prediction, Web mining, and mobile computing, besides others utilize Data mining [5].

KEY PHASES IN DATA MINING PROCESS [4]

Information association

This is the most familiar and straightforward feature of this system. Here we made association between two

or more items or often of the same type to formulate specific pattern. Example, it is very well known etiological association between smoking and lung cancer. We have to collect data concerned with smoking habit details including numbers of smoke per day, duration of smoking, type of smoking either bidis, cigarettes, specific brands, lifestyle and age of patient etc.

Information classification

This is second most popular phase. Here we can classify the collected information according to our objectives like etiological factors, investigation purpose, drug treatment plans and results. Example, the etiological information collected from lung cancer patients can be classified on the basis of duration of smoking habit, type of exposure, number of exposure, age of patient etc.

Information clustering

If we put collected data in the form of cluster by using computer graphics, it is easy to locate data trends. These identified trends may helpful in preparation of prediction system. Example, information collected from lung cancer patients about the onset and duration of exposure, clustering pattern guide us in future regarding prevention strategies.

Pattern Sequencing

The data pattern sequencing is next step in module preparation. The pattern sequencing can be prepared with the help of readymade software packages available in market.

Preparation of decision tree

This is final step of prediction system.

Implementation

This is base feature directly concerned with objectives. You may have option either long term or short term data processing. Each data mining system can be set with different objectives. Data mining process are broadly formulated either as supervised or 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) [3]. These are broadly dived either classification or prediction based pattern [6]. Decision Trees and Neural Networks use classification algorithms while Regression, Association Rules and Clustering use prediction algorithms [7].

MINING METHODS

Nowadays there are large numbers of mining software available in the market. The most commonly used methods are decision tree algorithms and neural networks technique.

Decision Tree algorithms

It includes CART (Classification and Regression Tree), ID3 (Iterative Dichotomized 3) and C4.5. These algorithms differ in selection of splits, when to stop a node from splitting, and assignment of class to a non-split node [8].

Neural Networks

It 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. Neural Network algorithms use Linear and Sigmoid transfer functions. Neural Networks are suitable for training large amounts of data with few inputs. It is used when other techniques are unsatisfactory [1].

BENEFITS OF DEVELOPMENTS OF DISEASE PREDICTION SYSTEM USING DATA MINING TECHNIQUES

1. **Prevention and diagnosis:** Data mining technique made prediction system plays a vital role in strategy preparation for prevention of communicable as well as non communicable diseases in located area. Lifestyle related diseases like hypertension, diabetes mellitus, cardiovascular diseases; stroke etc can be easily and accurately classified and possible to locate their etiological area cluster patterns.

   These techniques are also useful in disease diagnosis. Ms. Ishtake et al. developed a prototype Intelligent Heart Disease Prediction System (IHDPS) using three data mining modelling techniques, namely, Decision Trees, Naïve Bayes and Neural Network. IHDPS can discover and extract hidden knowledge (patterns and relationships) associated with heart disease from a historical heart disease database. It can answer complex queries for diagnosing heart disease and thus assist healthcare practitioners to make intelligent clinical decisions which traditional decision support systems cannot. By providing effective treatments, it also helps to reduce treatment costs. To enhance visualization and ease of interpretation, it displays the results both in tabular and graphical forms [1].

2. **Workout of treatment plan:** The data mining techniques play an important role in treatment plan workout, surgical procedures, rehabitation, chronic diseases management plan etc. Long term follow up plan may be easily guided and keen supervision is possible. Example, a patient of hypertension can be long term manage and back through record of number of patients will guide in implementing future strategies.

3. **Reduction of cost of patient management**: These systems may definitely helpful in reduction of cost of patient management by avoiding unnecessary investigations and patients follow up.

4. These prediction systems will add **accuracy and time management**.

5. **Discovery of hidden etiological factors:** This is most excitable objective planed by using these systems. This will be helpful for confirmation of geographical variations. Most of our health strategies are planned on the basis of data interpretations from developed countries. We can formulate our own systems and can avoid geographical errors.

6. Computer-based patient support systems benefit patients by providing **informational support** that increases their participation in health care [9].

SCOPE OF DATA MINING TECHNIQUES

Increasing computer based data analysis awareness, online educational availability and developing integrated learning approach among medical professionals will definitely helpful for accurate diagnosis and effective treatment management plan in India. Innovative medical technologies are essential for patient care. This is also true for prevention of various diseases related to hygiene, communicable diseases, addiction related diseases like lung cancer, oral cancer, liver cirrhosis etc. In future the scope of technology applications like data mining techniques based systems in the healthcare system in India will really make dramatic changes at every level [1]0. Today the internet is a gateway to world knowledge as well as a massive platform for national media and documentation [9].This will help a lot in future in implementation of data mining techniques.

REFERENCES

1. Ishtake SH and Sanap SA; Intelligent Heart Disease Prediction System Using Data Mining Techniques. International Journal of Healthcare & Biomedical Research, 2013; 1(3): 94-101.
2. Soumen Chakrabarti, Martin Ester, Usama Fayyad, Johannes Gehrke, Jiawei Han, Shinichi Morishita, Gregory Piatetsky-Shapiro, Wei Wang; Data Mining Curriculum: A Proposal (Version 1.0). tp://www.sigkdd.org/sites/default/files/ CUR May 06.pdf. 2006-04-30.
3. Obenshain MK; Application of Data Mining Techniques to Healthcare Data. Infection Control and Hospital Epidemiology, 2004; 25(8): 690–695.
4. IBM, Data mining techniques, http://www.ibm.com/developerworks/opensource/library/ba-data-mining-techniques/index.html?ca=drs- , downloaded on 04 April 2013.
5. Microsoft Developer Network (MSDN). http://msdn2.microsoft.com/en us/virtuallabs/aa740409.aspx, 2007.
6. Han J and Kamber M; Data Mining Concepts and Techniques. Morgan Kaufmann Publishers, 2006.
7. Charly K; Data Mining for the Enterprise. 31st Annual Hawaii Int. Conf. on System Sciences, IEEE Computer, 1998; 7: 295-304.
8. Ho TJ; Data Mining and Data Warehousing. Prentice Hall, 2005.
9. Malukani K, Keswani B, Chaturvedi S and Nandedkar S; Impact of Informational and Social Support Services on Patient's *Perceived Satisfaction* from Virtual Pathology Communities. Indian Journal of Basic & Applied Medical Research, 2013; 2(6): 477-482.
10. Tayade MC and Kulkarni NB; The Interface of Technology and Medical Education in India: Current Trends and Scope. Indian Journal of Basic & Applied Medical Research, 2011; 1(1): 8-12.