

A Survey of Health Care Prediction Using Data Mining

Sujatha R¹, Sumathy R², Anitha Nithya R³Assistant Professor, Dept. of Computer Science Engineering, Sri Krishna College of Technology, Coimbatore, India¹Assistant Professor, Dept. of Computer Science Engineering, Sri Krishna College of Technology, Coimbatore, India²Assistant Professor, Dept. of Computer Science Engineering, Sri Krishna College of Technology, Coimbatore, India³

ABSTRACT: Data Mining is one of the most motivating areas of research that is become increasingly popular in health organization. Mining plays an important role for uncovering new trends in healthcare organization which in turn helpful for all the parties associated with this field. It is a new powerful technology which is of high interest in computer world. It is a sub field of computer science that uses already existing data in different databases to transform it into new researches and results. It makes use of Artificial Intelligence, machine learning and database management to extract new patterns from large data sets and the knowledge associated with these patterns. The actual task is to extract data by automatic or semi-automatic means. The different parameters included in data mining include clustering, forecasting, path analysis and predictive analysis.

KEYWORDS: Data Mining, E-health care, Prediction, Effective Treatment

I. INTRODUCTION

Data Mining is one of the most vital and motivating area of research with the objective of finding meaningful information from huge data sets. In present era, Data Mining is becoming popular in healthcare field because there is a need of efficient analytical methodology for detecting unknown and valuable information in health data. In health industry, Data Mining provides several benefits such as detection of the fraud in health insurance, availability of medical solution to the patients at lower cost, detection of causes of diseases and identification of medical treatment methods. It also helps the healthcare researchers for making efficient healthcare policies, constructing drug recommendation systems, developing health profiles of individuals etc. [1].

The data generated by the health organizations is very vast and complex due to which it is difficult to analyze the data in order to make important decision regarding patient health. This data contains details regarding hospitals, patients, medical claims, treatment cost etc. So, there is a need to generate a powerful tool for analyzing and extracting important information from this complex data. The analysis of health data improves the healthcare by enhancing the performance of patient management tasks. The outcome of Data Mining technologies are to provide benefits to healthcare organization for grouping the patients having similar type of diseases or health issues so that healthcare organization provides them effective treatments[6]. It can also useful for predicting the length of stay of patients in hospital, for medical diagnosis and making plan for effective information system management. Recent technologies are used in medical field to enhance the medical services in cost effective manner. Data Mining techniques are also used to analyze the various factors that are responsible for diseases for example type of food, different working environment, education level, living conditions, availability of pure water, health care services, cultural ,environmental and agricultural factors as shown in Fig 1.

International Journal of Innovative Research in Science, Engineering and Technology

(An ISO 3297: 2007 Certified Organization)

Vol. 5, Issue 8, August 2016



Fig:1 Data Mining in Healthcare

Data mining has many applications in the fields of telecommunication industry, financial data analysis biological data analysis and much more. With the growing research in the field of health informatics a lot of data is being produced. The analysis of such a large amount of data is very hard and requires excessive knowledge. E-healthcare applies data mining and telecommunication techniques for health diagnosis. There are some patients who require continuous check-up and might need doctor help immediately. E-health was primarily used for patient data analysis and disease diagnosis at various levels.

II. RELATED WORK

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

Association Rule: The central task of association rule mining is to find sets of binary variables that co-occur together frequently in a transaction database[8], while the goal of feature selection problem is to identify groups of that are strongly correlated with each other with a specific target variable. Association rule has the several algorithms like: Apriori, CDA, DDA, interestingness measure etc.

A smart shirt has been designed in [10]. The shirt can measure electrocardiogram (ECG) and acceleration signals for continuous and real time health monitoring of a patient. The shirt mainly consists of sensors and conductive fabrics to get the body signal. The measured body signals are transmitted to a base station and server PC via IEEE 802.15.4 network. The wearable devices consume low power and they are small enough to fit into a shirt. To reduce the noise associated with the ECG signal an adaptive filtering method has also been proposed in this work.

In the home of the future, some devices will contribute physiological information about the patient (e.g., heart rate, blood pressure), while other devices in and around the home will contribute information about the patient's environment (e.g., humidity, temperature, carbon monoxide level). These physiological and environmental data will be collated to assess the patient's state of health and to identify external factors that may influence that state. In some cases, groups of devices will have enough collective awareness to function autonomously based on sensor data (e.g., A carbon monoxide detector may note levels above a safety threshold, initiating a protocol to open the windows, sound an alarm, and activate vital signs sensors for individuals in the house.

International Journal of Innovative Research in Science, Engineering and Technology

(An ISO 3297: 2007 Certified Organization)

Vol. 5, Issue 8, August 2016

By this we can analyze about the different kinds of smart health care prediction systems.

III. E-HEALTH CARE SYSTEMS

In today's world e-health is used to form patient and physicians blogs and monitoring data such as sensors attached to patient are exercised to assist better diagnosis of the patient and for a continuous check-up of sensitive patients. From [5] WHO defined E-health as, the use of information and communication technologies (ICT) for health to, for example, treat patients, pursue research, educate students, track diseases and monitor public health.

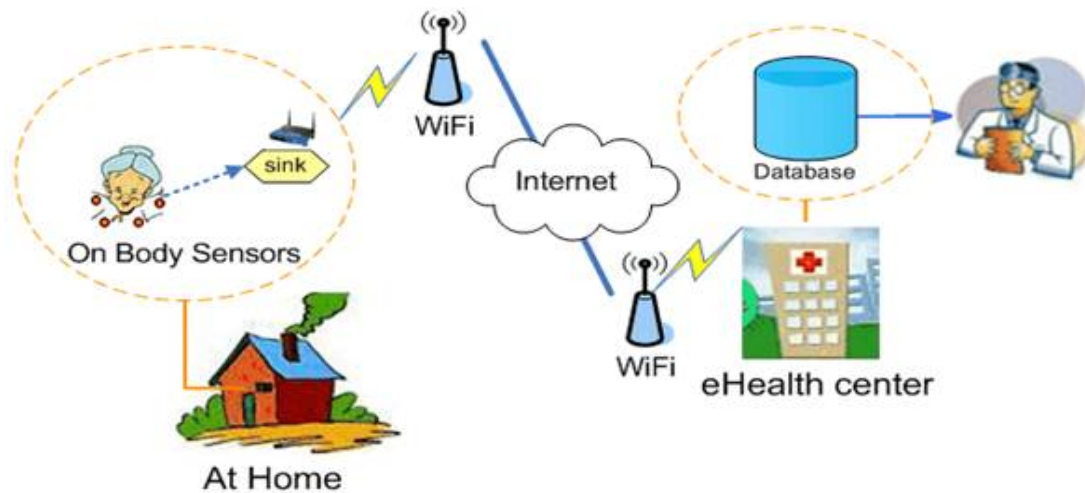


Fig2.E-Health Care Systems

Fig2 shows the advancement of technology more and more smart systems are being designed with better data mining technologies to give the most accurate results that could be associated with the disease. If the system after proper research is not able to provide the accurate results it notifies it to the patient the type of disease it feels that the user is associated with. If the symptoms do not exactly match any disease it displays the result the user symptoms might be related with. The system have information about the doctors phone number[9], address along with feedback and administrator control panel for system processes.

Smart health care system can be used to monitor PD patients. PD is a progressive neurological disorder. The patients different parameters such as voice, images, movements and daily activities are monitored and systems are generated using quantitative analysis and pattern recognition. E-health has shown significant improvement and advancement in the field of medicine. Health care is a data rich field. With increase in research more and more data is provided which would eventually increase the demand of data mining in this field.

IV. HEALTHCARE APPLICATIONS IN MINING

Tendency for data mining application in healthcare today is great, because healthcare sector is rich with information, and data mining is becoming a necessity. Healthcare organizations produce and collect large volumes of information on daily basis. Use of information technologies allows automatization [10] of processes for extraction of data that help to get interesting knowledge and regularities, which means the elimination of manual tasks and easier extraction of data directly from electronic records, transferring onto secure electronic system of medical records which

International Journal of Innovative Research in Science, Engineering and Technology

(An ISO 3297: 2007 Certified Organization)

Vol. 5, Issue 8, August 2016

will save lives and reduce the cost of the healthcare services, as well and early discovery of contagious diseases with the advanced collection of data. Data mining can enable healthcare organizations to predict trends in the patient conditions and their behaviors, which is accomplished by data analysis from different perspectives and discovering connections and relations from seemingly unrelated information. Raw data from healthcare organizations are voluminous and heterogeneous. They need to be collected and stored in the organized forms, and their integration enables forming of hospital information system. Healthcare data mining provides countless possibilities for hidden pattern investigation from these data sets. These patterns can be used by physicians to determine diagnoses, prognoses and treatments for patients in healthcare organizations.

There is vast potential for data mining applications in healthcare. Generally, these can be grouped as the evaluation of treatment effectiveness; management of healthcare; customer relationship management; and detection of fraud and abuse. More specialized medical data mining, such as predictive medicine and analysis of DNA micro-arrays, lies outside the scope of this paper. Other data mining applications related to treatments include associating the various side-effects of treatment, collating common symptoms to aid diagnosis, determining the most effective drug compounds for treating sub-populations that respond differently from the mainstream population to certain drugs, and determining proactive steps that can reduce the risk of affliction.

V. TREATMENT EFFECTIVENESS

Data mining applications can be developed to evaluate the effectiveness of medical treatments. By comparing and contrasting causes, symptoms, and courses of treatments, data mining can deliver an analysis of which courses of action prove effective. For example, the outcomes of patient groups treated with different drug regimens for the same disease or condition can be compared to determine which treatments work best and are most cost-effective. Along this line, United HealthCare has mined its treatment record data to explore ways to cut costs and deliver better medicine. It also has developed clinical profiles to give physicians information about their practice patterns and to compare these with those of other physicians and peer-reviewed industry standards. Similarly, data mining can help identify successful standardized treatments for specific diseases.

VI. HEALTHCARE MANAGEMENT

To aid healthcare management, data mining applications can be developed to better identify and track chronic disease states and high-risk patients, design appropriate interventions, and reduce the number of hospital admissions and claims. For example, to develop better diagnosis and treatment protocols[12], the Arkansas Data Network looks at readmission and resource utilization and compares its data with current scientific literature to determine the best treatment options, thus using evidence to support medical care[6]. Also, the Group Health Cooperative stratifies its patient populations by demographic characteristics and medical conditions to determine which groups use the most resources, enabling it to develop programs to help educate these populations and prevent or manage their conditions. Group Health Cooperative has been involved in several data mining efforts to give better healthcare at lower costs. Data mining can also be used to analyze massive volume of data and statistics to search for patterns that might indicate an attack by bio-terrorists.

VII. ADVANTAGES OF MINING APPLICATION IN HEALTH CARE

Information technologies in healthcare have enabled the creation of electronic patient records obtained from monitoring of the patient visits. This information includes patient demographics, records on the treatment progress, details of examination, prescribed drugs, previous medical history, lab results, etc. Information system simplifies and automates the workflow of health care institution. Privacy of documentation and ethical use of information about patients is a major obstacle for data mining in medicine. In order for data mining to be more exact, it is necessary to make a considerable amount of documentation. Health records are private information, yet the use of these private documents may help in treating deadly diseases. Before data mining process can begin, healthcare organizations must

International Journal of Innovative Research in Science, Engineering and Technology

(An ISO 3297: 2007 Certified Organization)

Vol. 5, Issue 8, August 2016

formulate a clear policy concerning privacy and security of patient records. This policy must be fully implemented in order to ensure patient privacy.

Health institutions are able to use data mining applications for a variety of areas, such as doctors who use patterns by measuring clinical indicators, quality indicators, customer satisfaction and economic indicators, performance of physicians from multiple perspectives to optimize use of resources, cost efficiency and decision making based on evidence, identifying high-risk patients and intervene proactively, optimize health care, etc. Integration of data mining in information systems, healthcare institutions reduce subjectivity in decision-making and provide a new useful medical knowledge.

Data mining provides the link between knowledge of continuous data[7], such as biomedical signals collected from patients in intensive care units, and it develops an intelligent monitoring system that sends reminders, warnings and alarms for the pre-selected critical conditions. Using association rules involves finding all the rules, or at least part of key subsets of rules that is characteristic of certain information as a consequences or as a antecedent. This type of problem is very interesting for health professionals who are searching for the relations between diseases and lifestyles or demographics or between survival rates and treatment.

The tasks of association are used to help strengthen the arguments regarding whether to engage or eliminate certain rules in the knowledge model. Tasks of the managers that manage quality of the healthcare services can be described as optimization of clinical processes in terms of medical and administrative quality as well as the cost/benefit relation. Key questions of the process of healthcare quality management are quality of data, standards, plans, and treatments[7]. Data mining can be used by quality managers to solve the following tasks: Discovering new hypothesis for indexes of quality for data, standards, plans and treatments; Checking if the given indexes of quality for data, standards, plans and treatments are still valid; Improving, strengthening and adjusting of quality indexes for data, standards, plans and treatments; and These tasks can be supported by data mining if the existing knowledge in domain is seriously considered in data mining process.

VIII. CONCLUSION

Data mining has great importance for area of medicine, and it represents comprehensive process that demands thorough understanding of needs of the healthcare organizations. Healthcare is one of the major sectors which can highly benefit from the implementation and use of information system. We have provided an overview of applications of data mining in infrastructure, administrative, financial and clinical Health care system. 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 has 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. These patterns can be used for clinical diagnosis. However, available raw medical data are widely distributed, different and voluminous by nature. These data must be collected and stored in data warehouses in organized forms, and they can be integrated in order to form hospital information system. Data mining technology provides customer oriented approach towards new and hidden patterns in data, from which the knowledge is being generated, the knowledge that can help in providing of medical and other services to the patients. Healthcare institutions that use data mining applications have the possibility to predict future requests, needs, desires, and conditions of the patients and to make adequate and optimal decisions about their treatments. With the future development of information communication technologies, data mining will achieve its full potential in the discovery of knowledge hidden in the medical data.

International Journal of Innovative Research in Science, Engineering and Technology

(An ISO 3297: 2007 Certified Organization)

Vol. 5, Issue 8, August 2016

REFERENCES

- [1] Han, J., Kamber, M.: "Data Mining Concepts and Techniques", Morgan Kaufmann Publishers, 2006
- [2] Mar y K. Obenshain, MAT "Application of Data Mining Techniques to Healthcare Data", Infection control and Hospital Epidemiology, 2004
- [3] Abdullah A. Aljumah, Mohammed Gulam Ahamad, Mohammad Khubeb Siddiqui, "Application of data mining: Diabetes health care in young and old patients", 2012.
- [4] Koh HC¹, Tan G., "Data mining applications in healthcare", 2005, Spring.
- [5] Bennett C, Doub T (2011) Data mining and electronic health records: selecting optimal clinical treatments in practice. CoRR abs/1112: 1668
- [6] Estella F, Delgado-Marquez BL, Rojas P, Valenzuela O, San Roman B, Rojas I (2012) Advanced system for autonomously classify brain MRI in neurodegenerative disease In: International Conference on Multimedia Computing and Systems (ICMCS 2012). IEEE, based in New York, USA, Tangiers, Morocco, pp 250–255
- [7] Mahboob khan, "Smart Health Prediction Using Data Mining", 2015, Spring.
- [8] Obenshain, M.K: "Application of Data Mining Techniques to Healthcare Data", Infection Control and Hospital Epidemiology, 25(8), 690–695, 2004
- [9] Ledbetter C, Morgan, M. "Toward Best Practice: Leveraging the Electronic Patient Record as a Clinical Data Warehouse". JOURNAL OF HEALTHCARE INFORMATION MANAGEMENT, VOL 15; PART 2, pages 119-132, 2001.
- [10] Strisland, F.; Sintef.; Svagard, I.; Seeberg, T.M.(2013) "ESUMS: A mobile system for continuous home monitoring of rehabilitation patient", In the proceedings of the 35th IEEE Annual International Conference on Engineering in Medicine and Biology Society, July 3-7, 2013, Osaka, pp. 4670-4673
- [10] Bourouis, A., Feham, M., and Bouchachia, A.(2011), " Ubiquitous Mobile Health Monitoring System for Elderly (UMHMSE)", International Journal of Computer Science and Information Technology, Vol.2, No. 3, June, pp. 74-82
- [11] Varbol, Peter. "Next-Generation Internet Appliances," *Computer Design*, March 1998, pp. 22-24
- [12] Lei Clifton, David A. Clifton, Marco A. F. Pimentel, Peter J. Watkinson, and Lionel Tarassenko (2014), " Predictive Monitoring of Mobile Patients by Combining Clinical Observations with Data From Wearable Sensors", IEEE Journal of Biomedical and Health Informatics, Vol. 18, No. 3, May, pp. 722-730
- [13] Xiaoliang Wang ; Qiong Gui ; Bingwei Liu ; Zhanpeng Jin et al (2014), "Enabling Smart Personalized Healthcare: A Hybrid Mobile-Cloud Approach for ECG Telemonitoring", IEEE Journal of Biomedical and Health Informatics, Vol. 18, No. 3, May, pp. 739 – 745
- [14] Orlando R. E. P., Caldeira, M. L. P. Lei S., and Rodrigues, J.P.C (2014), "An Efficient and Low Cost Windows Mobile BSN Monitoring System Based on TinyOS", Journal of Telecommunication Systems, Vol. 54, No. 1, pp. 1-9
- [15] Dunsmuir, D., Payne, B.; Cloete, G.; Petersen, C.(2014), "Development of m-Health Applications for Pre-eclampsia Triage", IEEE Journal of Biomedical and Health Informatics, Vol. PP, No. 99, January, pp. 2168-2194