International Journal of Advanced Technology in Engineering and Science
Vol. No.3, Issue 11, November 2015
www.ijates.com
ijates
ISSN 2348 - 7550

DATA MINING TECHNIQUES IN HEALTH CARE APPLICATION: A REVIEW OF SURVEY

Manaswini Pradhan

P.G. Department of Information and Communication Technology, Fakir Mohan University, Orissa, (India)

ABSTRACT

The issue of health care assumes prime importance for the society and is a significant indicator of social development. Health is clearly not the mere absence of disease but confers on a person or group's freedom from illness and the ability to realize one's potential. Health is therefore best understood as the indispensable basis for defining a person's sense of well-being. The delivery of health care services thus assumes greater proportion, and in this context the role played by information and communication technology has certainly a greater contribution for its effective delivery mechanism. The application of data mining is specifically relevant and it has been successfully applied in medical needs for its reliable precision accuracy and expeditious beneficial results. The various available application techniques have been discussed and analyzed for the purpose of the paper.

Keywords: Health care, Data mining, Electronic Medical Records (EMR), Medical decision support, Diagnosis and Treatment

I. INTRODUCTION

In India healthcare is delivered through both the public sector and private sector. The public healthcare system consists of healthcare facilities run by central and state government which provide services free of cost or at a subsidized rates to low income group in rural and urban areas. With the Indian economy enjoying a steady growth, the industry is heading towards growth phase by ***(M. Radha, 2011)***. According to **(Jeffrey M. Lackner *et al.* 2013)** the introduction of product patents in India is expected to boost the industry by encouraging multinational companies to launch specialized life-saving drugs. Attracted by the advantages such as lower costs of production and skilled workforce that India offers, these companies are looking to set up research and development as well as production centers there.

Health and health care need to be distinguished from each other for no better reason than that the former is often incorrectly seen as a direct function of the latter. Heath is clearly not the mere absence of disease by ***(Ilias Iakovidis, 1998)***. Good Health confers on a person or group's freedom from illness - and the ability to realize one's potential. Health is therefore best understood as the indispensable basis for defining a person's sense of well-being. The health of populations is a distinct key issue in public policy discourse in every mature society often determining the deployment of huge society. According to **(Shivalingappa B Javali *et al.* 2012)** include its cultural understanding of ill health and well-being, extent of socio-economic disparities, reach of health services and quality and costs of care. And current bio-medical understanding about health and illness.

1.1 Definitions of Health Care

A major challenge facing healthcare organizations (hospitals, medical centers) is the provision of quality services at affordable costs. Quality service implies diagnosing patients correctly and administering treatments that are effective. Poor clinical decisions can lead to disastrous consequences which are therefore unacceptable **(Alok Bhargava, *et al.* 2005)**. Hospitals must also minimize the cost of clinical tests. According to **(Hoosen Coovadia *et al.* 2009)** can achieve these results by employing appropriate computer-based information and/or decision support systems. Health care data is massive. It includes patient centric data, resource management data and transformed data by **(Cheon-Pyo Lee *et al.* 2007)**. Health care organizations must have ability to analyze data. Treatment records of millions of patients can be stored and computerized and data mining techniques may help in answering several important and critical questions related to health care.

The healthcare industry is expected to increase in size from its current €12.72 billion to €29.6 billion by 2012. India will spend €33.8 billion on healthcare in the next five years as the country, on an economic upsurge, is witnessing changes in its demographic profile accompanied with lifestyle diseases and increasing medical expenses by **(Patricia Pittman *et al.* 2007)**. Revenues from the healthcare sector account for 5.2% of the GDP and it employs over 4 million people. By 2012, revenues can reach 6.5 to 7.2% of GDP and direct and indirect employment can double, it said. According to **(Palaniappan S *et al.* 2008)** in private healthcare will continue to be the largest component in 2012 and is likely to double to €26.41 billion. It could rise by an additional €6.5 billion if health insurance cover is extended to the rich and middle class. Coupled with the expected increase in the pharmaceutical sector, the total healthcare market in the country could increase to €39.22 – 54 billion (6.2-8.5% of GDP) in the next five years.

1.2 Scope of the Healthcare Industry

Some of the macro factors for this industry's growth are the following:

Emerging private sector is more focused on tertiary-level as wells as preventive and diagnostic healthcare and is sensing a huge untapped opportunity in delivery of quality healthcare to the Indian masses by **(Leigh Turner *et al.* 2007)**. According to **(Harleen Kaur *et al.* 2009)** public sector is ramping up prevention and elimination of infectious diseases as accessibility of basic healthcare facilities to the rural masses; Global Private Equity and Venture Capitals are playing a vital and varied role in Indian healthcare delivery to increasing the global footprint of local pharmaceutical companies to aiding the rapidly growing contract research outsourcing industry **(Patricia Pittman *et al.* 2007)**.

- **Market situation:**

During the past decade there has been a great change in the availability of healthcare facilities in the country. The number of public hospitals grew from 4600 in 2000 to more than 7600 in 2006 that is an increase of more than 67% in just 6 years by **(Siri Krishan Wasan *et al.* 2006)**. The private hospitals have been growing at a similar pace during these years. Private healthcare boasts of superior quality and facilities. According to **(K.Srinivas *et al.* 2010)** an accounts for more than 65% of primary care and more than 40% of hospitals, resulting in personnel shortage in the public sector.

- **Medical infrastructure:**

It forms the largest portion of the healthcare pie. As per the current statistics (2006) bed per thousand population ratio for India stands at 1.03 as against an average 4.3 of comparable countries like China, Korea and Thailand **(2002 data).** Hence in spite of the phenomenal growth in the healthcare infrastructure, we are likely to reach a bed to thousand-population ratio of 1.85 and in a best-case scenario, a ratio of 2 by **(Latha Parthiban *et al.* 2008)**. Beds in excess of 1 million need to be added to reach a ratio of 1.85 per thousand, out of which about 896,500 beds will be added by the private sector with a total investment of €51 Billion over the next six years. According to **(Varun Kumar *et al.* 2008)** the gains are commensurate in this capital intensive industry, since the revenues generated by private hospitals in the year 2012 will be to the tune of €26.5 billion growing at a CAGR of 15%. Despite this investment, the bed to thousand population ratio would be far from comparison with other similar developing countries.

- **Telemedicine:**

It allows even the interiors to access quality healthcare and at the same time, according to the model proposed by us, significantly improves the productivity of medical personnel. In a country of over 1.1 Billion people, the Healthcare system will have to innovate to double the utilization of its existing resources just to reach a stage at which comparable developing countries by **(Aqueel Ahmed *et al.* 2012)**. According to **(Monica Chiarini Tremblay *et al.* 2009)** telemedicine is one such innovative technology, if used effectively can double utilization of scarce human resources **(Leigh Turner *et al.* 2007)**. Standalone telemedicine models may not be feasible, but if telemedicine models are integrated in a Healthcare model, such models can become viable. One important reason is that Telemedicine shall increase the patient base by **(Vikram Jeet Singh *et al.* 2013)**, which in turn will increase occupancy rates of hospitals in the integrated telemedicine model.

II. DATA MINING

Data mining can be considered a relatively recently developed methodology and technology, coming into prominence only by **(K. R. Lakshmi *et al.* 2013)**. It aims to identify valid, novel, potentially useful, and understandable correlations and patterns in data by combing through copious data sets to sniff out patterns that are too subtle or complex for humans to detect. According to **(Jinn-Yi Yeh *et al.* 2011)** have given data mining techniques can be broadly classified based on what they can do, namely description and visualization; association and clustering; and classification and estimation, which is predictive modeling.

2.1 Objectives

The objectives of this paper are the following:

1. To enumerate current uses and highlight the importance of data mining in medicine and public health,
2. To find data mining techniques used in other fields that may also be applied in the health sector.
3. To identify issues and challenges in data mining as applied to the medical practice.
4. To outline some recommendations for discovering knowledge in electronic databases through data mining.

2.2 Data Mining Applications in Healthcare

Data mining has been used intensively and extensively by many organizations. In healthcare, data mining is becoming increasingly popular, if not increasingly essential. Data mining applications can greatly benefit all parties involved in the healthcare industry by **(Petr Hájek *et al,* 2010)**. For example, data mining can help healthcare insurers detect fraud and abuse, healthcare organizations make customer relationship management decisions, physicians identify effective treatments and best practices, and patients receive better and more affordable healthcare services. Data mining can be defined as the process of finding previously unknown patterns and trends in databases and using that information to build predictive models by **(E.W.T. Ngai *et al.* 2011)**. Alternatively, it can be defined as the process of data selection and exploration and building models using vast data stores to uncover previously unknown patterns.

2.3 Why Data Mining Can Aid Healthcare

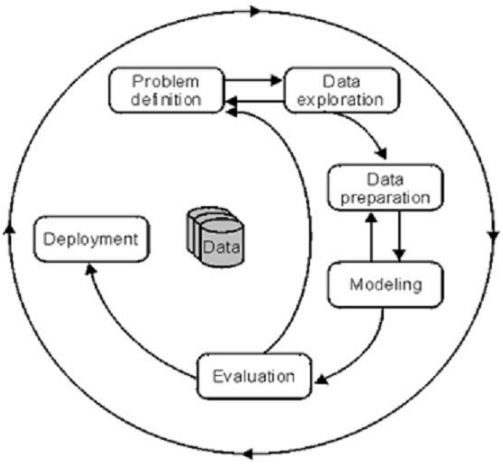
In healthcare, data mining is becoming increasingly popular, if not increasingly essential. Several factors have motivated the use of data mining applications in healthcare. According to **(Chao-Hui Lee *et al.* 2010)** existence of medical insurance fraud and abuse, for example, has led many healthcare insurers to attempt to reduce their losses by using data mining tools to help them find and track offenders **(Aqueel Ahmed *et al.* 2012)**. According to **(Monica Chiarini Tremblay *et al.* 2009)** Fraud detection using data mining applications is prevalent in the commercial world, for example, in the detection of fraudulent credit card transactions by **(R.S. Santos *et al.* 20010)**. Recently, there have been reports of successful data mining applications in healthcare fraud and abuse detection.

2.4 Data Mining can help

- Healthcare insurers detect fraud and abuse,
- Healthcare organizations make customer relationship management decisions,
- Physicians identify effective treatments and best practices, and
- Patients receive better and more affordable healthcare services.

III. REVIEW OF LITERATURE FOR APPLICATION WITH DATA MINING IN HEALTHCARE INDUSTRY

In this section of paper an attempt is made to comprehensively and systematically review the literature on Healthcare Industry and classify the applications with data mining. According to **(Monica Chiarini Tremblay *et al.* 2009)** it can be easily comprehended and applied. In this, around 100 papers from around 3 leading journals spread over last 10 years have been considered.



Source: Petr Hájek *et al.* [2010]

3.1 Importance Role of Healthcare Application with Data Mining

Despite the differences and clashes in approaches, the health sector has more need for data mining today. There are several arguments that could be advanced to support the use of data mining in the health sector, covering not just concerns of public health but also the private health sector by **(Ting-Ting Lee *et al.* 2011)**. Data overload. According to **(Wen-Yuan Jen *et al.* 2009)** there is a wealth of knowledge to be gained from computerized health records. Yet the overwhelming bulk of data stored in these databases makes it extremely difficult, if not impossible, for humans to sift through it and discover knowledge by **(Syed Sibte Raza Abidi *et al.* 2001)**. In fact, some experts believe that medical breakthroughs have slowed down, attributing this to the prohibitive scale and complexity of present-day medical information. Computers and data mining are best-suited for this purpose. According to **(Hsu-Hao Tsai *et al.* 2012)** evidence-based medicine and prevention of hospital errors. When medical institutions apply data mining on their existing data, they can discover new, useful and potentially life-saving knowledge that otherwise would have remained inert in their databases. For instance, an on-going study on hospitals and safety found that about 87% of hospital deaths in the United States could have been prevented, had hospital staff been more careful in avoiding errors. By **(Hyunjung Shin *et al.* 2012)** mining hospital records, such safety issues could be flagged and addressed by hospital management and government regulators.

3.2 Medical

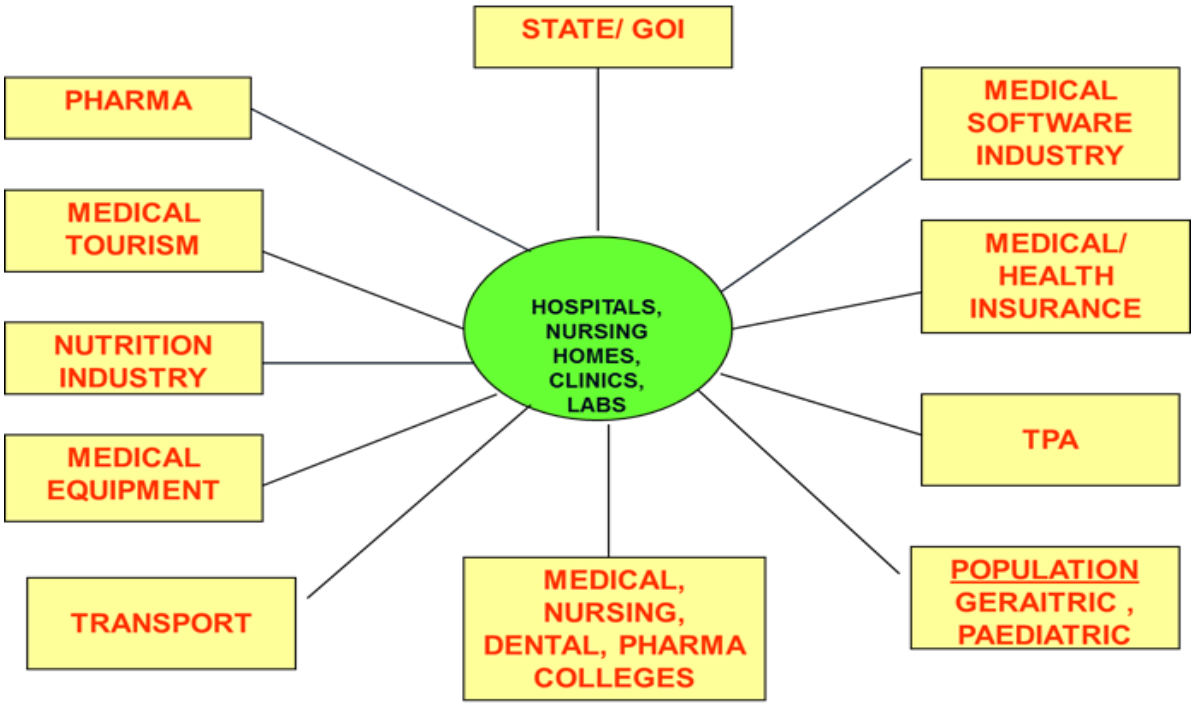
1. Picture Archival & Communication System (PACS)
2. DICOM (Digital Imaging & Communications in Medicine)
3. Telemedicine
4. Electronic Medical Records
5. Clinical Decision Support Systems
6. HL7 protocol
7. Epidemic prevention software

3.3 Non-Medical

1. Integrated hospital information system
2. Web-enabled Appointment Scheduling
3. Web-enabled applications for relatives to obtain conditions of the critical patients
4. Multi-media applications for patient education

IV. GRADUAL EVOLUTION OF THE HEALTHCARE INFRASTRUCTURE VERSUS ENDOGENOUS FACILITY PLACEMENT

Policy-making in public health. Combined GIS and data mining using among others, Weka with J48, to analyze similarities between community health centers in Slovenia by **(Ali Serhan Koyuncugil *et al.* 2012)**. Using data mining, they were able to discover patterns among health centers that led to policy recommendations to their Institute of Public Health.



The Integrated Emergency, Healthcare, and Medical Information System will be developed in the web-based multimedia environment, mobility and real-time technology by **(Ali Serhan Koyuncugil *et al.* 2012)**. The system provides an integrated medical database, which can provide stakeholders with related medical information. The registered users can log into the system to access or provide medical information based on their accessing privilege **(Sumana Sharma *et al.* 2009)**. The system will have the capabilities for finding the patient location based and suggest the nearest emergency center, arrange all necessary related patient information to be ready for the physician when the patient arrives, assigning a doctor to the patient based on the availability of the doctors and list all necessary requirements such as special devices or surgery room **(Dursun Delen *et al.* 2009)**. The system is an open cross-platform web-based real-time client-server environment with multiple language capabilities. The system provides mechanisms for exchange of image files, shared discussion

lists, textual information exchange, access to images and data exported from local data bases, voice and video transmission **(Gloria Phillips-Wren *et al.* 2008)**.

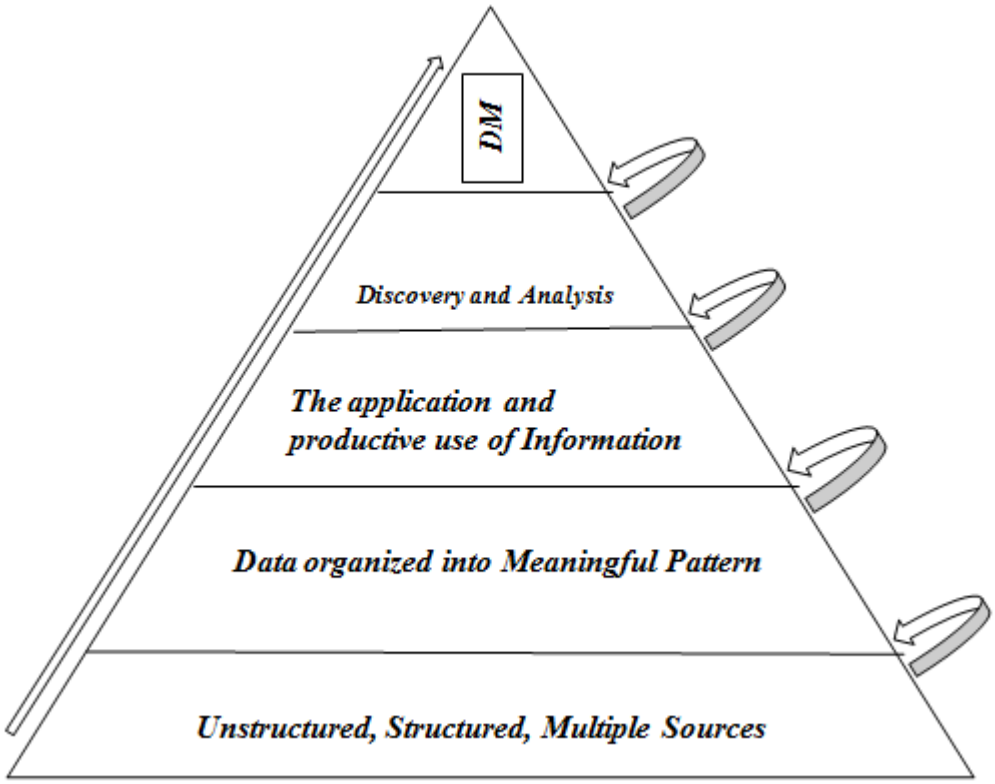


Table 1: Topology of Data Mining Technology of Health Care Application:

References	*Data Mining Technology*
(Aqueel Ahmed *et al.* 2012) **(R.S. Santos *et al.* 20010)**	System supports a function defined by standardized MOH forms can be entered for inclusion in the EMR.
(Hsu-Hao Tsai et al. 2012) **(Ali Serhan Koyuncugil *et al.* 2012)**	Health Care Eligibility Benefit Inquiry and Response
(Dursun Delen *et al.* 2009) **(Gloria Phillips-Wren *et al.* 2008).**	Kiosk Application
(Ting-Ting Lee *et al.* 2011) **(Wen-Yuan Jen *et al.* 2009)**	Electronic Billing and Claims Management
(Syed Sibte Raza Abidi *et al.* 2001) **(Hyunjung Shin *et al.* 2012)**	Laboratory Information Management Systems

4.1 Functional Requirements for EMR Systems

Electronic medical records can be developed to address different goals and health settings, and consequently emerge with different functions and capabilities **(Wan-Shiou Yang *et al.* 2006)**. However, it is desirable to maintain a core set of functions in each EMR system in order to support similar workflows and encourage best practices in clinical care. According to **(Zhengxing Huang *et al.* 2012)** the functional requirements defined in this chapter can be categorized into6 key functional areas that are critical to the definition of an EMR:

(i) Basic demographic and clinical health information;
(ii) Clinical decision support;
(iii) Order entry and prescribing;
(iv) Health information and reporting;
(v) Security and confidentiality, and;
(vi) Exchange of electronic information.

4.2 EMR Systems Must Address the Following 6 key Functional Areas

i. Basic Demographic and Clinical Health Information
ii. Clinical Decision Support
iii. Order Entry and Prescribing
iv. Health Information and Reporting
v. Security and Confidentiality
vi. Exchange of Electronic Information

Technology elements	***Year***	***Description of current technology***	***Function***
EMR User interface text template	*2010*	*User can enter pre-written text strings which can be edited*	*Text based interface prior to introduction of GUI software*
EMR User interface -point-and-click	*2008*	*Several interfaces in use no dominant player yet*	*Faster and more accurate billing process*

4.3 Medical Decision Support

Analysis of digitized images of skin lesions to diagnose melanoma. Computer-assisted texture analysis of ultrasound images aids monitoring of tumor response to chemotherapy. Predicting the presence of brain neoplasm with magnetic resonance spectroscopy **(Hyunjung Shin *et al.* 2012)**. Analysis of digital images of tissue sections to identify and quantify senile plagues for diagnosing and evaluating the severity of Alzheimer's disease.

4.4 Diagnosis and Treatment

Data mining could be particularly useful in medicine when there is no dispositive evidence favoring a particular treatment option. According to **(Krzysztof J. Cios *et al.* 2002)** Based on patients' profile, history, physical examination, diagnosis and utilizing previous treatment patterns, new treatment plans can be effectively suggested.

4.5 Healthcare Resource Management

Using logistic regression models to compare hospital profiles based on risk-adjusted death with 30 days of non-cardiac surgery **(David F. Motta Cabrera *et al.* 2013)**. According to **(A´lvaro Rebuge *et al.* 2012)** Neural network system to predict the disposition in children presenting to the emergency room with bronchiolitis. Predicting the risk of in-hospital mortality in cancer patients with nonterminal disease.

4.6 Prediction of Inpatient Length of Stay

Effectively manage the resource allocation by identifying high risk areas and predicting the need and usage of various resources. A key problem in the healthcare area is the measurement of flow of patients through hospitals and other health care facilities **(Thanh Kim Dao *et al.* 2010)**. If the inpatient length of stay (LOS) can be predicted efficiently, the planning and management of hospital resources can be greatly enhanced.

4.7 CRM is to Establish Close Customer Relationships

The focus shifts away from the breadth of customer base (product oriented view, mass marketing) to the depth of each customer's needs (customer-oriented view, one-to-one marketing) **(Dursun Delen *et al.* 2009).** CRM is built on an integrated view of the customer across the whole organization. Customers have a fractured view of an enterprise; the enterprise has a splintered view of the customer. **(Kohli *et al. 2010*)** demonstrate a web-based Physician Profiling System (PPS) to strengthen relationships with physicians and improve hospital profitability and quality. Development of total customer relationship in healthcare includes several tenets. According to **(S.Vijiyarani *et al.* 2013)** a helping profession, the ultimate judge of performance is the person helped. Most people, including sick people, are reasonable most of the time. Different people have different, legitimate needs. Pain and fear produce anxiety in both the victim and the helper **(Dursun Delen *et al.* 2009).** Meeting needs without waste is a strategic and moral imperative. Some demographic characteristics and institutional characteristics consistently have a significant effect on a patient's satisfaction scores. Chronic illnesses require self-management and a collaborative patient-physician relationship.

The principles of applying of data mining for customer relationship management in the other industries are also applicable to the healthcare industry **(Dursun Delen *et al.* 2009).** The identification of usage and purchase patterns and the eventual satisfaction can be used to improve overall customer satisfaction **(N. Aditya Sundar *et al.*, 2012)**. The customers could be patients, pharmacists, physicians or clinics. In many cases prediction of purchasing and usage behavior can help to provide proactive initiatives to reduce the overall cost and increase customer satisfaction.

4.8 Bolton and Hand Briefly Discuss Healthcare Insurance Fraud

Prescription fraud: claims for patients who do not exist

Upcoding: claims for a medical procedure which is more expensive or not performed at all.

The ability to detect anomalous behavior based on purchase, usage and other transactional behavior information has made data mining a key tool in variety of organizations to detect fraudulent claims, in appropriate prescriptions and other abnormal behavioral patterns **(Motilal C. Tayade *et al.* 2003)**. Another key area where data mining based fraud detection is useful is detection and prediction of faults in medical devices.

V. DATA QUALITY AND COMPLETENESS

Data quality and completeness are critical to the success of any information system. Achieving high standards is a particular challenge in sites with limited computer literacy and experience **(Aqueel Ahmed *et al.* 2012)**. It is important to design systems that are easy to use and have good instructions and training. The system should collect the minimum data necessary for the task, and data items should be structured and coded where possible to simplify data checking and optimize reuse. According to **(Ting-Ting Lee *et al.* 2011) t**his does not mean that free text must be excluded; doing so prevents the system from capturing any data that do not fit the normal pattern. Such data will either be lost or recorded in hard-to-locate paper records. Structured data such as laboratory test results might benefit from double entry **(Dursun Delen *et al.* 2009).** In some projects physicians and other staff enter data directly. This has the advantage of avoiding transcription errors, and also allows order entry systems to be deployed to check for potential medical errors.

5.1 The Obstacles for Data Mining In Healthcare

One of the biggest problems in data mining in medicine is that the raw medical data is voluminous, and heterogeneous. These data can be gathered from various sources such as from conversations with patients, laboratory results, review and interpretation of doctors **(Ali Serhan Koyuncugil *et al.* 2012)**. All these components can have a major impact on diagnosis, prognosis and treatment of the patient, and should not be ignored. According to **(Sumana Sharma *et al.* 2009) t**he scope and complexity of medical data is one of the barriers to successful data mining. Missing, incorrect, inconsistent or non-standard data such as pieces of information saved in different formats from different data sources create a major obstacle to successful data mining **(Dursun Delen *et al.* 2009).** It is very difficult for people to process gigabytes of records, although working with images is relatively easy, because doctors are being able to recognize patterns, to accept the basic trends in the data, and formulate a rational decisions. Stored information becomes less useful if they are not available in easily apprehensible format **(Gloria Phillips-Wren *et al.* 2008)**. The role of visualization techniques is increasing in this, as the picture are easiest for people to understand, and can provide plenty of information in a snapshot of the results.

Authors	**Year**	**Consequences**	**Application Technology**
(E.W.T. Ngai et al.) **(Petr Hájek et al,)**	2010 2010	Dirt, moisture, and bird droppings can contaminate window air conditioners, which can then introduce infectious material into hospital rooms.	EMR User interface text template
(Ting-Ting Lee et al.) **(Syed Sibte Raza Abidi et al.)**	2011 2001	Several interfaces in use –no dominant player yet	EMR User interface -point-and-click
(Chao-Hui Lee et al.) **(Hyunjung Shin et al.)**	2010 2012	socioeconomic, public spending and healthcare system	Laboratory Information Management Systems
(Aqueel Ahmed et al.)	2012	Particles in the air are larger than the	Evaluation and Management

(Hsu-Hao Tsai et al.)	2012	openings between the filter fibers, resulting in gross removal of large particles.	Services
(R.S. Santos et al.) **(Hsu-Hao Tsai et al.)**	2010 2012	Particles collide with filter fibers and remain attached to the filter. Fibers may be coated with adhesive.	Document Management System
(Wen-Yuan Jen et al.) **(Syed Sibte Raza Abidi et al.)**	2009 2001	Small particles, moving in erratic motion, collide with filter fibers and remain attached.	Continuity of Care Record
(Ali Serhan Koyuncugil et al.) **(Sumana Sharma et al.)**	2012 2009	Rigorous air filtration requires air flow resistance. Air stream will elude filtration if openings are present because of filter damage or poor fit.	Personal Health Record
(Dursun Delen et al.) **(Gloria Phillips-Wren et al.).**	2009 2008	Open windows can alter fan-induced pressure balance and allow dirty-to clean air flow	Lab Report Automation Software.

5.2 Data Security and Confidentiality

Views of medical data security and confidentiality vary in different developing countries. In some sites, the use of electronic databases is treated with great suspicion **(Chao-Hui Lee *et al.* 2010)**; in other sites staff think nothing of emailing sensitive medical data. Patients can face serious risk if their communities discover their HIV status or other sensitive medical information.

5.3 Estimation of Current HIT Adoption and Related Factors

Our primary data Source was the Healthcare Information and Management Systems Society -Dorenfest survey which represents a broad canvassing of acute care hospitals, chronic care facilities, and ambulatory practices on their adoption and plans to adopt various HIT components **(Chao-Hui Lee *et al.* 2010)**. They included in the adoption category the provider organizations that had contracted for but not yet installed an EMR system **(Sumana Sharma *et al.* 2009)**. To examine the factors related to differences in adoption, emerged additional data about the providers and then performed pro bit regression analysis **(Petr Hájek *et al,* 2010).** Our lower-bound estimate of HIT adoption as summed an integrated system that had an EMR, clinical decision support, and a central data repository from the same vendor to ensure inter-operability They adjusted the estimates according to the known under-representation of smaller providers in this survey **(E.W.T. Ngai *et al.* 2011).**

5.4 Challenges in Data Mining for Healthcare

5.4.1 Data Sets from Various Data Sources

1: Patient referral data can vary extensively between cases because structure of patient referrals is up to general practitioner who refers the patient

2: Dipti Patil *et al.* use neural networks to predict preterm birth on a heterogeneous maternal population

3: Traditional clinical-based prognosis models were discovered to contain some restrictions to address the heterogeneity of breast cancer.

5.4.2 Data from Heterogeneous Sources Present Challenges

1. *Sampling bias:* Clinical studies use diverse collecting methods, inclusion criteria, and sampling methods.
2. *Referral bias: Data represent a preselected group with a high prevalence of disease.*
3. *Selection bias:* Clinical data sets include patients with different demographics.
4. *Method bias:* Predictors have varied specifications, granularities, and precisions.
5. *Clinical spectrum bias:* Patient records represent varied severity of a disease and co-occurrence of other medical problems.

5.4.3 Missing Values, Noise, and Outliers

1. Cleaning data from noise and outliers and handling missing values, and then finding the right subset of data, prepares them for successful data mining.
2. Transcription and manipulation of patient records often result in a high volume of noise and a high portion of missing values.
3. Missing attribute values can impact the assessment of whether a particular combination of attribute-value pairs is significant within a dataset.

5.5 Advantages of Data Mining Application in Healthcare

Information technologies in healthcare have enabled the creation of electronic patient records obtained from monitoring of the patient visits **(Dursun Delen *et al.* 2009)**. 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 **(Sumana Sharma *et al.* 2009).** Privacy of documentation and ethical use of information about patients is a major obstacle for data mining in medicine. According to **(Motilal C. Tayade *et al.* 2013)** 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 **(V. Krishnaiah *et al.* 2013)**. According to **(Gloria Phillips-Wren *et al.* 2008)** before data mining process can begin, healthcare organizations must formulate a clear policy concerning privacy and security of patient records. This policy must be fully implemented in order to ensure patient privacy. According to **(Ishtake S. H *et al.* 2012)** 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.

VI. 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. 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.

REFERENCES

[1] M. Radha "Problems and Prospects Of Health Care services In India", Asian Journal Of Res E Arch In Business Economics And Management, Vol. 1, No. 3, Pp. 88-97, 2011.

[2] Jeffrey M. Lackner, James Jaccard and Charles Baum "Multidomain Patient-Reported Outcomes of Irritable Bowel Syndrome: Exploring Person-Centered Perspectives to Better Understand Symptom Severity Scores", Value in Health, Vol. 16, pp. 97 –103, 2013.

[3] Ilias Iakovidis, "Towards personal health record: current situation, obstacles and trends in implementation of electronic healthcare record in Europe", International Journal of Medical Informatics, Vol. 52, pp. 105 – 115, 1998.

[4] Shivalingappa B Javali. "A Structural Equation Model of the Determinants of Health Care in the Surveyed Households in Rural Area of Dharwad District, Karnataka State, India", National Journal of Medical Research, Vol. 2, No. 3, pp. 245-249, 2012.

[5] Alok Bhargava, Sadia Chowdhury and K.K. Singh "Healthcare infrastructure, contraceptive use and infant mortality in Uttar Pradesh, India", Economics and Human Biology, Vol. 3, pp. 388–404, 2005.

[6] Hoosen Coovadia, Rachel Jewkes, Peter Barron, David Sanders, Diane McIntyre, "The health and health system of South Africa: historical roots of current public health challenges", The Lancet, Vol. 374, No. 9692, pp. 817-834, 2009.

[7] Cheon-Pyo Lee and Jung P Shim "An exploratory study of radio frequency identification (RFID) adoption in the healthcare industry", European Journal of Information Systems, Vol. 16, pp. 712–724, 2007.

[8] P Pittman, LH Aiken, J Buchan, "International Migration of Nurses: Introduction", Health Services Research, Vol. 42, No. 3 pp. 1275–1280, 2007

[9] Palaniappan S. and Awang R. "Intelligent heart disease prediction system using data mining techniques", In preceding of IEEE/ACS International Conference on Computer Systems and Applications, Doha, pp.108-115, 2008.

[10] Leigh Turner, "First World Health Care at Third World Prices': Globalization, Bioethics and Medical Tourism",

[11] Harleen Kaur and Siri K. Wasan "Empirical Study on Applications of Data Mining Techniques in Healthcare", Journal of Computer Science, Vol. 2, No. 2, pp. 194-200,

[12] Siri Krishan Wasan, Vasudha Bhatnagar and Harleen Kaur "The impact of data mining techniques on medical diagnostics", Data Science Journal, Vol. 5, pp. 119-126, 2006.

[13] K.Srinivas B.Kavihta Rani and A. Govrdhan "Applications of Data Mining Techniques in Healthcare and Prediction of Heart Attacks", International Journal on Computer Science and Engineering, Vol. 02, No. 02, pp. 250-255, 2010.

[14] Latha Parthiban and R.Subramanian "Intelligent Heart Disease Prediction System using CANFIS and Genetic Algorithm", International Journal of Biological and Medical Sciences, Vol. 3, No. 3, 2008.

[15] Varun Kumar, Dharminder Kumar and R. K. Singh "Outlier Mining in Medical Databases: An Application of Data Mining in Health Care Management to Detect Abnormal Values Presented In Medical Databases", International Journal of Computer Science and Network Security, VOL.8 No.8, pp. 272-277, August 2008.

[16] Aqueel Ahmed, Shaikh Abdul Hannan "Data Mining Techniques to Find Out Heart Diseases: An Overview", International Journal of Innovative Technology and Exploring Engineering, Vol. 1, No. 4, pp. 18-21, 2012.

[17] Monica Chiarini Tremblay, Donald J. Berndt, Stephen L. Luther, Philip R. Foulis and Dustin D. French "Identifying fall-related injuries: Text mining the electronic medical record", Information Technology and Management, Vol. 10, No. 4, pp. 253-265, 2009.

[18] Vikram Jeet Singh, Dharam Pal Singh and Kishori Lal Bansal, "Proposed Architecture: Cloud Based Medical Information Retrieval Network", International Journal of Computer Science & Engineering Technology, Vol. 4 No. 05, pp. 485-496, May 2013.

[19] K.R. Lakshmi, M. Veera Krishna, and S. Prem Kumar, "Performance Comparison of Data Mining Techniques for Predicting of Heart Disease Survivability", International Journal of Scientific and Research Publications, Vol. 3, No. 6, pp. 1-10, 2013.

[20] Jinn-Yi Yeh, Tai-Hsi Wu and Chuan-Wei Tsao "Using data mining techniques to predict hospitalization of hemodialysis patients", Decision Support Systems, Vol. 50, pp. 439–448, 2011.

[21] Petr Hájek, Martin Holena and Jan Rauch, "The GUHA method and its meaning for data mining", Journal of Computer and System Sciences, Vol. 76, pp. 34–48, 2010.

[22] E.W.T. Ngai, Yong Hu, Y.H. Wong, Yijun Chen and Xin Sun "The application of data mining techniques in financial fraud detection: A classification framework and an academic review of literature", Decision Support Systems, Vol. 50, pp. 559–569, 2011.

[23] Chao-Hui Lee, Jessie Chia-Yu Chen and Vincent S. Tseng "A novel data mining mechanism considering bio-signal and environmental data with applications on asthma monitoring", computer methods and programs in biomedicine, Vol. 101, pp. 44–61, 2011.

[24] R.S. Santos, S.M.F. Malheiros, S. Cavalheiro and J.M. Parente de Oliveira "A data mining system for providing analytical information on brain tumors to public health decision makers", Computer Methods and Programs in Biomedicine, Vol. 109, No. 3, pp. 269-282, 2013

[25] Ting-Ting Lee, Chieh-Yu Liu, Ya-Hui Kuo, Mary Etta Mills, Jian-Guo Fong and Cheyu Hungd "Application of data mining to the identification of critical factors in patient falls using a web-based reporting system", International journal of medical informatics, Vol. 80, pp. 141–150, 2011.

[26] Wen-Yuan Jen "Mobile healthcare services in school-based health center", International Journal of Medical Informatics, Vol. 78, pp. 425–434, 2009.

[27] Syed Sibte Raza Abidi, "Knowledge management in healthcare: towards 'knowledge-driven' decision-support services", International Journal of Medical Informatics, Vol. 63, pp. 5–18, 2001.

[28] Hsu-Hao Tsai "Global data mining: An empirical study of current trends, future forecasts and technology diffusions", Expert Systems with Applications, Vol. 39, pp. 8172–8181, 2012.

[30] Hyunjung Shin, Hayoung Park, Junwoo Lee and Won Chul Jhee "A scoring model to detect abusive billing patterns in health insurance claims", Expert Systems with Applications, Vol. 39, pp. 7441–7450, 2012.

[31] Ali Serhan Koyuncugil and Nermin Ozgulbas "Financial early warning system model and data mining application for risk detection", Expert Systems with Applications, Vol. 39, pp. 6238–6253, 2012.

[32] Sumana Sharma and Kweku Muata Osei-Bryson, "Framework for formal implementation of the business understanding phase of data mining projects", Expert Systems with Applications, Vol. 36, pp. 4114–4124, 2009.

[33] Dursun Delen, Christie Fuller, Charles McCann and Deepa Ray "Analysis of healthcare coverage: A data mining approach", Expert Systems with Applications, Vol. 36, pp. 995–1003, 2009.

[34] Gloria Phillips-Wren, Phoebe Sharkey and Sydney Morss Dy "Mining lung cancer patient data to assess healthcare resource utilization", Expert Systems with Applications, Vol. 35, pp. 1611–1619, 2008.

[35] Wan-Shiou Yang and San-Yih Hwang "A process-mining framework for the detection of healthcare fraud and abuse", Expert Systems with Applications, Vol. 31, pp. 56–68, 2006.

[36] Zhengxing Huang, Xudong Lu and Huilong Duan "On mining clinical pathway patterns from medical behaviors", Artificial Intelligence in Medicine, Vol. 56, pp. 35–50, 2012.

[37] Krzysztof J. Cios and G. William Moore, "Uniqueness of medical data mining", Artificial Intelligence in Medicine, Vol. 26, pp. 1–24, 2002.

[38] David F. Motta Cabrera and Hamidreza Zareipour "Data association mining for identifying lighting energy waste patterns in educational institutes", Energy and Buildings, Vol. 62, pp. 210 –216, 2013.

[39] Alvaro Rebuge and DiogoR.Ferreira "Business process analysis in healthcare environments: A methodology based on process mining", Information Systems, Vol. 37, pp. 99–116, 2012.

[40] Thanh Kim Dao, Firas Zabaneh, Judy Holmes, Laura Disrude Margaret Price and Layne Gentry, "A practical data mining method to link hospital microbiology and an infection control database", American Journal of Infection Control, Vol. 36, No. 3, pp. 18-20, 2008.

[41] JE Rey "Evaluation of Cesarean Section Endometritis Case-Finding Methods", American Journal of Infection Control, Vol. 34, No. 5, pp. 157-165, 2006

[41] K.Srinivas B.Kavihta Rani and A.Govrdhan "Applications of Data Mining Techniques in Healthcare and Prediction of Heart Attacks", International Journal on Computer Science and Engineering, Vol. 02, No. 02, pp. 250-255, 2010.

[42] Nidhi Bhatla and Kiran Jyoti "An Analysis of Heart Disease Prediction using Different Data Mining Techniques", International Journal of Engineering Research & Technology (IJERT), Vol. 1, No. 8, pp. 1-4, 2012.

[43] S.Vijiyarani and S.Sudha "Disease Prediction in Data Mining Technique – A Survey", International Journal of Computer Applications & Information Technology, Vol. II, No. I, pp. 17-21, January 2013.

[44] V. Krishnaiah, G. Narsimha and N. Subhash Chandra, "A Study on Clinical Prediction Using Data Mining Techniques", International Journal of Computer Science Engineering and Information Technology Research, Vol. 3, No. 1, pp. 239-248, 2013.

[45] Mary K. Obenshain "Application of Data Mining Techniques to Healthcare Data", Infection Control and Hospital Epidemiology, Vol. 25, No. 8, pp. 690-695, 2004.

[46] M. Hemalatha and S. Megala "Mining Techniques in Health Care: A Survey of Immunization", Journal of Theoretical and Applied Information Technology, Vol. 25, No.2, 2011.

[47] Dipti Patil, Bhagyashree Agrawal, Snehal Andhalkar, Richa Biyani, Mayuri Gund and V. M. Wadhai, "An Adaptive parameter free data mining approach for healthcare application", International Journal of Advanced Computer Science and Applications, Vol. 3, No. 1, pp. 55-59, 2012.

[48] N. Aditya Sundar, P. Pushpa Latha and M. Rama Chandra "Performance Analysis of Various Data Mining Classification Techniques on Healthcare Data", International Journal of Engineering Science & Advanced Technology, Vol. 2, No. 3, pp. 470 – 478, 2012.

[49] Motilal C. Tayade and Pratibha M. Karandikar "Role of Data Mining Techniques in Healthcare sector in India", Scholars Journal of Applied Medical Sciences, Vol. 1, No. 3, pp. 158-160, 2013.

[50] Ishtake S. H and Sanap S. A. "Intelligent Heart Disease Prediction System Using Data Mining Techniques", International Journal of Healthcare & Biomedical Research, Vol. 1, No. 3, pp. 94-101, 2013.