Integrating Data Mining Techniques into Telemedicine Systems

Mihaela GHEORGHE, Ruxandra PETRE
The Bucharest University of Economic Studies
mihaela.gheorghe@ie.ase.ro, ruxandra_stefania.petre@yahoo.com

The medical system is facing a wide range of challenges nowadays due to changes that are taking place in the global healthcare systems. These challenges are represented mostly by economic constraints (spiraling costs, financial issues), but also, by the increased emphasis on accountability and transparency, changes that were made in the education field, the fact that the biomedical research keeps growing in what concerns the complexities of the specific studies etc. Also the new partnerships that were made in medical care systems and the great advances in IT industry suggest that a predominant paradigm shift is occurring. This needs a focus on interaction, collaboration and increased sharing of information and knowledge, all of these may is in turn be leading healthcare organizations to embrace the techniques of data mining in order to create and sustain optimal healthcare outcomes. Data mining is a domain of great importance nowadays as it provides advanced data analysis techniques for extracting the knowledge from the huge volumes of data collected and stored by every system of a daily basis. In the healthcare organizations data mining can provide valuable information for patient’s diagnosis and treatment planning, customer relationship management, organization resources management or fraud detection. In this article we focus on describing the importance of data mining techniques and systems for healthcare organizations with a focus on developing and implementing telemedicine solution in order to improve the healthcare services provided to the patients. We provide architecture for integrating data mining techniques into telemedicine systems and also offer an overview on understanding and improving the implemented solution by using Business Process Management methods.

Keywords: Telemedicine, Data Mining, Healthcare, Data Warehouse, Business Process Management, Architecture

1 Introduction

In our current society, the healthcare domain is at a critical crossroads and faces a wide number of challenges. There are for example, the threatening of the continuous increase of costs, the expand of the government-funded healthcare coverage, the continuous number of the uninsured, a very slow process of quality care improvement, the demographic trends, especially the number of senior population that keeps increasing and therefore putting strain on the system, a non-optimized health insurance framework etc.

The research that is being done in the innovation direction, especially in what concerns the provisioning and administration of offering remote healthcare services has conducted to the appearance and development of a new domain which represents a key enabler that can improve the quality of medical services by enhancing the operational effectiveness and achieving improved outcomes. This domain was named telemedicine and represents a solution to the large number of challenges that the medical system confronts with.

Telemedicine represents a combination of expertise and technology that provides services and information over distance. [1] IT technology delivers this information in the form of voice, data or video image and its main objective consists in the utility and efficacy of the technology itself, as it is applied to hospitals problems and their infrastructure. The use of telemedicine for providing care services has been rapidly increasing and with the latest innovation in the information and computer technologies, it has become an important part of the medical development process with a high potential of improving the economic prosperity of the country.

Information and technology resources are having a very important contribution in the

DOI: 10.12948/issn14531305/18.1.2014.11
process of avoiding and reducing clinical errors but also in the optimization process of the medical health provider’s processes and workflows. [2] These resources are mostly being used for activities concerning the process of:
- collecting medical data from patients;
- providing the required medical information to the physicians, patients or other members of the medical community;
- real-time monitoring of patients;
- diagnosing patients using expert systems.

Systems and applications in telemedicine broadly involve communication between a client (generally represented by the patient or by a person working in the medical field) and a specialist in the medical field, where the two parties are not in the same place. [3] There are many types of services provided by telemedicine, including:
- telemonitoring (remote monitoring of vital signs);
- teleradiology (production and transmission of digital image compression respectively decoding);
- telecardiology (access to medical services for patients presenting with heart disease but also in educating them in order to track the correct steps required to therapy;
- telepathology (the practice of pathology at a distance and uses telecommunications technology to facilitate the transfer of image-rich pathology data between distant locations for the purposes of diagnosis, education, and research);
- teleophthalmology (delivers eye care through digital medical equipment and telecommunications technology).

Telemedicine covers a wide spectrum of benefits through the use of video conferencing technology in areas such as consultations, diagnostics, therapeutic, transfer of patient-related records, case management, training, and meetings. Researchers envision telemedicine to be an important building block in the strategic plan of many health care organizations. In a rural setting, telemedicine could help health providers in supplying quality, fast, and economical medical services to rural patients and hence save doctors and patients valuable time wasted in commuting long distances. Specialists could utilize this extra time in seeing more patients at the main hospital.

In the clinical workflow process, the medical informatics has an important role in what concerns controlling the activities that are needed to be followed in order to obtain resources, to maintain their insurance, storage, retrieval and also to use them in order to provide benefits for patients and doctors. Some of these benefits include the following [4]:
- optimization of management appointments;
- easy and unlimited access to patient history in real-time;
- use of expert systems that can lead to making a medical diagnosis;
- reducing the time taken for the transmission of medical investigations to qualified medical personnel for their analysis and interpretation.

This kind of integration between medicine and technologies is needed due to the fact that the medical services market faces different issues related to lack of funds, overcrowding and low quality among managed services and treatments. [3] Using a specific telemedicine solution can streamline activities in a manner that allows a better overview which leads to a better control over the needed resources. By controlling costs and resources an integrated flow automation solution is ensured. This can provide accurate information exchange and improve medical institution profitability with a positive impact on the economy.

Thus implementing a mechanism for reducing physical flows for document management within a medical institution and promotion of electronic documents, leads to cost control by reducing them (it eliminates the cost of papers, film, use of space and other physical devices). [5] Integrating technologies aimed at streamlining medical processes, leading to reduce patients waiting time in terms of consultation and/or its medical diagnosis is reflected ultimately in a decentralization of health services. Reducing document flows

DOI: 10.12948/issn14531305/18.1.2014.11
has a significant impact on physician competence, they can get real-time information, regarding patient’s history or they can pass medical observations really quickly to a specialized center for their management. Therefore, the main objectives and opportunities of the healthcare field are [6]:

- transforming business and clinical processes to optimize workflows;
- controlling medical costs;
- ensuring quality and providing time access to care;
- encouraging the use of telemedicine applications that enables greater productivity and collaboration;
- ensuring security maintenance and data protection;
- helping patients live healthier lives with fewer hospitals and clinical visits;
- triggering new business models and opportunities to enhance quality of care like diagnosis or patients management solution for vital signs irregularities etc.

The high accessibility and imminent growth of the volumes of data collected and stored in every application led to the necessity of using a diversity of data analysis tools in order to explore and extract the significant information from the data. Many domains of human life, like business, education or science, need new and advanced data analysis techniques in order to obtain relevant knowledge from the available data. Therefore, in the past two decades a new data analysis subfield emerged and developed, that focuses on applying a diversity of techniques for knowledge discovery for both descriptive and predictive purposes. This particular subfield of data analysis is data mining, which applies advanced data analysis techniques to explore the data and identify useful and unexpected patterns and rules that provide relevant knowledge for predicting future outcomes.

Data mining can be defined as a process of exploring and analysis for large amounts of data with a specific target on discovering significantly important patterns and rules. Data mining helps finding knowledge from raw, unprocessed data. Using data mining techniques allows extracting knowledge from the data mart, data warehouse and, in particular cases, even from operational databases. [7]

Data mining became an important discovery process for various domains including business, science or education, where it can be used for the following activities:

- business – data mining can be used in retail, banking or insurances, for purposes like customer segmentation, market basket analysis or fraud detection;
- science – data mining can be applied in medicine, where it can help with diagnosis prognosis or treatment planning, bioinformatics, genetics, electrical power or astronomy;
- education – data mining can be used for predicting student performance, understanding student behavior or student segmentation.

Data mining has many successful applications, such as business intelligence, Web search, bioinformatics, health informatics, finance, digital libraries, and digital governments. [8]

Data mining methods are incorporating techniques from statistics, artificial intelligence, machine learning and database systems. According to the Gartner Group (The world’s leading information technology research and advisory company), data mining is the process of discovering meaningful new correlations, patterns and trends by sifting through large amounts of data stored in repositories, using pattern recognition technologies as well as statistical and mathematical techniques. [9]

Therefore, the main purpose of data mining consists in applying advanced data analysis techniques on large volumes of data stored in databases, data marts or data warehouses, in order to discover significant patterns, trends, rules, relations and correlations contained within the data, information which would be impossible or very difficult to observe using other data analysis techniques.

The process of applying data mining techniques to the data, presented in Figure 1, takes the following steps in order to extract relevant knowledge:

DOI: 10.12948/issn14531305/18.1.2014.11
Data mining applies a variety of advanced data analysis techniques for extracting useful knowledge from the data. The main data mining techniques used for identifying hidden patterns and trends are: [10]

- classification – consists of a function that maps (classifies) a data item into one of several predefined classes;
- regression – involves a function that maps a data item to a real-valued prediction variable;
- clustering – is a common descriptive task where one seeks to identify a finite set of categories or clusters to describe the data;
- association rule learning (Dependency modeling) – consists of finding a model that describes significant dependencies between variables;
- anomaly detection (Change and deviation detection) – focuses on discovering the most significant changes in the data from previously measured or normative values.

The diversity of data and the multitude of data mining techniques provide various applications for data mining, which have improved many fields of human life. Data mining is a dynamic and fast-expanding domain which offers great benefits, the main advantages of using data mining techniques being the possibility to predict future trends and the help with decision making, based on the patterns and trends discovered.

Also, there are many challenging issues in data mining research. Areas include mining methodology, user interaction, efficiency and scalability, and dealing with diverse data types. Data mining research has strongly impacted society and will continue to do so in the future. [11]

The ongoing remarkable growth in the field of data mining and knowledge discovery has been fuelled by a fortunate confluence of a variety of factors:

- the explosive growth in data collection;
- the storing of the data in data warehouses, so that the entire enterprise has access to a reliable current database;
- the availability of increased access to data from Web navigation and intranets;
- the competitive pressure to increase market share in a globalized economy;
- the development of off-the-shelf commercial data mining software suites;
- the tremendous growth in computing power and storage capacity.

One of the most important domains that can apply data mining is the health care. The high availability and impressive amounts of data
collected and stored in the medical databases require appropriate tools to be developed for analyzing and discovering knowledge in the data. With the culminant growth of the volumes of data contained by the medical applications came the difficulty of extracting significant information to base decisions upon. This led to the requirement of integrating data mining techniques in the medical systems. Data mining in medicine is most often used for building classification models, these being used for diagnosis, prognosis or treatment planning. [11] Therefore, data mining can be applied in the medical domain for purposes like improving the decision making such as diagnostic and prognostic problems in various fields, such as oncology, liver pathology, neuropsychology or gynecology.

2 Integrating Data Mining Techniques into Telemedicine Systems

Today, in the healthcare industry huge volumes of data are collected and stored, data referring to patients, medical devices, hospital resources, diagnostics and treatments etc. Data mining provides a set of techniques that can be applied to process and analyze the data in order to discover patterns that would enable healthcare personnel to make decisions based on the obtained knowledge. All healthcare institutions need an expert analysis of their medical data, the project that is time consuming and expensive. There is a great potential for data mining application in healthcare. Healthcare institutions are very oriented on use of patient’s information. Ability to use a data in databases in order to extract useful information for quality health care is a key of success of healthcare institutions. Healthcare information systems contain large volumes of information that include information on patients, data from laboratories that are constantly growing. With the use of data mining methods, useful patterns of information can be found in this data that will later be used for further research and report evaluation. [12]

The key domains for which data mining can provide support in healthcare are:
- diagnosis and treatment, by enabling doctors to identify diagnosis and provide effective treatment plans;
- customer relationship, by making better decisions regarding customer services and offering patients improved services;
- detecting fraud, used by insurers in order to identify fraud.

Therefore, integrating data mining techniques into telemedicine systems would help improve the efficiency and effectiveness of the healthcare organizations activity. For this purpose we propose, in Figure 2 below, the following architecture for a telemedicine system that integrates data mining tools:

![Fig. 2. Architecture for integrating data mining techniques into a telemedicine system](image-url)
According to our architecture, the medical data is extracted from various data sources, both internal and external to the healthcare organization, and loaded through an ETL (Extract, Transform and Load) process to the medical data warehouse. The data sources contain information referring to patients, hospital resources, customer relationship data, medical devices, diagnostics and treatments information etc. The medical data warehouse will contain raw and summary data, and also metadata, data that provides information about the raw and summary data belonging to the data warehouse.

The proposed solution provides a data mining engine that may be used to perform advanced analysis on the medical data collected. The graphical user interface provides the possibility to apply data mining methods for the healthcare professionals, by accessing the solution from various devices, like the desktop, laptop, tablet or mobile.

After generating the results the healthcare personnel can use the obtained knowledge to make medical decisions, like providing diagnosis and treatment planning to the patients. Through telemedicine the communication between the doctors and patients is improved, since patients are provided with quick access to physicians and the patient’s medical records are available to the physicians though the organization’s software.

Data mining techniques can help doctors to make medical decisions by identifying diagnosis. Also, by analyzing the profile, history, diagnosis and previous treatments of the patients, new treatment plans can be suggested using data mining methods.

Moreover, telemedicine brings new benefits for integrating data mining in the healthcare information systems, as it increases the volumes of data collected in all the key domains for which data mining is successfully applied.

Telemedical programs, tend to produce far greater amount of diagnostic data than in-person diagnostics. Data-mining analysis performed on such data may not only reveal new information on the diagnosed diseases but also further optimize settings of the applied diagnostic tests. [13]

Therefore, data mining has great importance for the healthcare domain, as the knowledge obtained through it provides means of better understanding the needs of the healthcare organizations, as well as making decisions that would contribute to improving the health of the patients.

The statistics represented in Figure 3 are showing that healthcare domain is growing and this could be due to the following:

- flexibility: in this way, healthcare resources can be accessed from any environments (rural or urban), anytime;
- accessibility: telemedicine application are accessible from any environment.
...and provides access to many health resources;

- portability: it refers to the adaptability and coexistence of all the common resources of the IT devices with the application installed;
- reduce of time: based on the fact that technology expends time and compresses the space, there is no need for doctors and patients to be on the same place in the same time, they could just communicate in real-time through the IT device applications.

The most common telemedicine applications are represented in the following chart illustrated in Figure 4.

Cardio applications are designed to monitor blood pressure by means of specific sensors for each cardiac cycle in order to identify the minimum and respectively maximum, necessary for the diagnosis of cardio respiratory disorders. Another category of applications is the one that provides chronic disease management, with the purpose to measure vital parameters of the patient and their transmission specialist doctors, a specific example being those that measure blood glucose levels to monitor people with diabetes [14].

Regarding the other categories of features implemented through telemedicine solution the following are the most commonly:

- information - application for public awareness campaigns on health;
- alerting service - applications that aim to inform patients at regular intervals on the actions needed during their treatment;
- remote consulting - allow communication between patients and specialist doctors when they are not in the same location;
- remote monitoring - program that allows keeping the medical history of patients and follow their progress;
- medical diagnostics - tools for physicians and patients that aim to facilitate the process of establishing a preliminary diagnosis;
- emergency services - provides the ability to call on medical services and ambulance or other specialized organizations;
- helpline - a service that can be called for documentation on drugs, equipment and other medical information.

In what concerns the medical institution which is represented by the physicians, the main advantage consists of the speed of information processing. Therefore, all the informational solutions provided by the data mining integration can ensure that the process of capturing and processing medical information will be very rapid with the help of computerization of the entire procedure. Moreover, this integration ensure that all the medical stages starting with patient scheduling, proper consultation and diagnosis, access to laboratory analysis, scheduling further consultation or hospitalization, and also medication prescription, are covered by the obtained workflows. The two main advantages
of this solution, e.g. the process of controlling flow and automation solution that provides accurate information exchanges and a growing trend of medical institution’s profitability. Implementing a mechanism for reducing physical flows for document management within an institution and promotion of electronic document shall ensure reducing internal costs (it eliminates the cost of paper, film, use of space and other physical devices).

On the other hand, there are other significant advantages of this integration which are represented by the efficient working time that will be obtained. This consists of the following [15]:

- optimization of medical appointments and reduction of the downtime during the day due to an efficient allocation of medical staff program;
- reducing the time taken for the transmission of medical results or investigations due to the fact that they are integrated by the means of data mining technologies that can capture and automatically transmits via the Internet, in various formats (video, audio, image etc) such information directly to qualified medical personnel for their analysis and interpretation;
- reducing the time and human errors in the registration process of patients arriving at the emergency room. Through telemedicine solutions can achieve the patient's electronic record that is to go to the emergency room for rapid intervention.[5]

Therefore, the implementation of such solutions leads to reduced patient waiting time in terms of consultation and / or its diagnosis by a doctor, which is reflected in the end by a decentralization of health services. Remove flow documents have a significant impact on doctor’s skills due to the fact that they can get information about patients in real time or forward them to a specialized medical observation center for their management and review. Therefore, this automation process leads to some economic advantages such as [6]:

- improving the quality of medical services;
- improved productivity and efficiency in the health sector;
- ensuring customer care data security;
- update information quickly on the health of patients;
- improving the working conditions of staff / doctors;
- improving care at home or in the field of medical;
- reduce costs;
- reducing the problems facing the health sector and are derived from the lack of funds, overcrowding or poor quality services, treatments administered by a physician;
- reduce complexity and automation of activities and related medical processes.

Another benefit is familiar to the doctors who are performing field work, facilitating access to patient information by eliminating the physical transfer of files that would be their medical records, and the possibility of patient’s registration in the system. Specific application can be used in both urban and rural environments, with specialized information that could be accessed directly on different devices, also, with no need for consultation of specialized documents, connect to a network or the internet to provide important data for the medical diagnosis. [16]

There are also advantages for the doctors working in the ambulance service. They can access useful application to obtain medical records when they are on move, in order to determine the treatment necessary to be provided for first help. In this sense it can be seen of the patient is allergic to different medication or suffer from a particular medical condition that would have caused the symptoms for which he is treated. Following the diagnosis of the patient in question, the doctor may send him to the hospital by ambulance along with the information that represents the preliminary diagnosis, and may announce, if appropriate, a specialist that will prepare in order to take over the patient or to
make a video for an urgent medical intervention.
We can conclude that integrating data mining techniques into telemedicine systems is beneficial for both physicians and patients, as the telemedicine solutions improves the communication between them and by integrating data mining tools the decision making process for the healthcare professionals is better supported.

3 Business Process Management for a Telemedicine System That Integrates Data Mining Techniques
Business process management is an emerging topic that provides the desired development and competitiveness for improving the business effectiveness and efficiency in achieving the objectives of the organization. According to the Gartner Group, business process management (BPM) is the discipline of managing processes (rather than tasks) as the means for improving business performance outcomes and operational agility. Processes span organizational boundaries, linking together people, information flows, systems and other assets to create and deliver value to customers and constituents. [17]
Business process management focuses on understanding, managing and improving the processes in an organization. The six main activities in business process management are: vision, design, modeling, execution, monitoring, and optimization.
Business process management (BPM) is a field of knowledge at the intersection between management and information technology, encompassing methods, techniques and tools to design, enact, control, and analyze operational business processes involving humans, organizations, applications, documents and other sources of information. [18]
Both telemedicine and data mining are successful emerging fields, so understanding and improving the processes involved in the implementation of a solution integrating them is of high importance.
Applying business process management to a system that integrates data mining techniques with telemedicine provides the following benefits:
- providing a single consistent view of the processes involved that can be understood by both the healthcare professionals and developers implementing the system;
- ceasing to have information that is not communicated completely and properly between the healthcare professionals and the developers;
- effectively and efficiently integrating the tools and techniques within the system;
- significantly shorter periods for delivering and running the processes that are part of the system.
Therefore, using business process management for implementing a system that integrates data mining techniques into telemedicine is beneficial for improving understanding of the processes involved for both healthcare professionals, that use the system, and developers, that implement it.

4 Conclusions
In recent years, providing care services through telemedicine has become an important part of the medical development process, due to the latest innovation in the information and computer technologies. Meanwhile, data mining, a dynamic and fast-expanding domain, has improved many fields of human life by offering the possibility of predicting future trends and helping with decision making, based on the patterns and trends discovered.
The diversity of data and the multitude of data mining techniques provide various applications for data mining, including in the healthcare organization. Integrating data mining techniques into telemedicine systems would help improve the efficiency and effectiveness of the healthcare organizations activity, contributing to the development and refinement of the healthcare services offered as part of the medical development process.
The integration of data mining techniques into telemedicine systems offers benefits for both healthcare professionals and for patients. Through a telemedicine system the communication between them is improved.
and the data mining tools integrated into the system support the decision making process of the physicians.

As business process management focuses on understanding, managing and improving the processes in an organization, using business process management for implementing a system that integrates data mining techniques into telemedicine refines the processes involved by the system.

References


DOI: 10.12948/issn14531305/18.1.2014.11

Mihaela GHEORGHE has graduated the Faculty of Economic Cybernetics, Statistics and Informatics of the Bucharest University of Economic Studies in 2010 as a promotion leader. In 2012 she graduated the Informatics Economics Master program also as a promotion leader. She is currently conducting research in Economic Informatics at Bucharest University of Economic Studies, coordinated by Professor dr. Bogdan GHILIC-MICU for her PhD thesis “Architecture and technologies in telemedicine”. She was a pre-Assistant within the Department of Economic Informatics between 2012 and 2013. Currently she is working as a software engineer at IBM Corporation, since 2012. Her main scientific interests include: telemedicine system, mobile technologies.

Ruxandra PETRE graduated from the Faculty of Cybernetics, Statistics and Economic Informatics of the Bucharest University of Economic Studies in 2010. In 2012 she graduated the Business Support Databases Master program. Currently, she is a PhD candidate, coordinated by Professor Ion LUNGU in the field of Economic Informatics at the Bucharest University of Economic Studies. The title of her PhD thesis is “Information Solutions for Data Analysis”. Her scientific fields of interest include: Databases, Data Warehouses, Business Intelligence, Decision Support Systems and Data Mining.