AMERICAN JOURNAL OF AGRICULTURAL SCIENCE, ENGINEERING AND TECHNOLOGY (AJASET)

ISSN: 2158-8104 (ONLINE), 2164-0920 (PRINT)

VOL: 5 ISSUE: 2 (2021)

PUBLISHED BY: E-PALLI, FLORIDA, USA

The American Journal of Agricultural Science, Engineering and Technology (AJASET) is blind peer reviewed international journal publishing articles that emphasize research, development and application within the fields of agricultural science, engineering and technology. The AJASET covers all areas of Agricultural Science, Engineering and Technology, publishing original research articles. The AJASET reviews article within approximately two weeks of submission and publishes accepted articles online immediately upon receiving the final versions.

Published Media: ISSN: 2158-8104 (Online), 2164-0920 (Print).

Frequency: 2 issues per year (January, July)

Area of publication: Agricultural Science, Any Engineering and Technology related original and innovative works.

DATA MINING TECHNOLOGY AS A TOOL FOR SUPPORTING ANALYTICAL DECISION MAKING PROCESS IN HEALTH INFORMATION MANAGEMENT SYSTEM (HIMS)

Asere Gbenga Femi[1], Dung Emmanuel Botson[2]

DOI: https://doi.org/10.54536/ajaset.v5i2.98

ABSTRACT

Wide spread use of information system in the delivery of managed healthcare system and the challenges of identifying and disseminating relevant healthcare information, complex and diverse data and knowledge forms and tasks coupled with the prevalence of legacy systems require automated approaches for effective and efficient utilization of massive amount of data to support in strategic planning and decision-making and assist the strategic management mechanisms. Despite the fact that data mining is progressively used in information systems as a technology to support analytical decision making, it is however still barely used in hospital information system to support analytical decision making process. Hence, this paper presents the usefulness of data mining technology in Hospital Information Management System (HIMS). Data mining technology offered capabilities to increase the productivity of medical personnel, analyze care outcomes, lower healthcare costs, improve healthcare quality by using fast and better clinical decision making and generally assist the strategic management mechanisms.

Keywords: Hospital, HIMS, Data Mining, DMHIMS

[1]Department of Computer Science, Federal School of Statistics, Manchok, Kaduna State, Nigeria. Email: aseregbenga@gmail.com, [2]Department of Computer Science, Federal School of Statistics, Manchok, Kaduna State, Nigeria. Email: botemma08@gmail.com

INTRODUCTION

Data play a significant role in carrying out day-to-day activities of an organization, in fact every business or organization needs data which form the basis of generating useful information. Organizations are set up to achieve certain objectives and these objectives will need to be met within various financial constraints and available limited resources. Data can be described as basic facts (numbers, figures, symbols or aggregates of one or more of these) which are raw or undesired in their present state or form and need to be transformed into desired form. The desired form into which data is finally transformed after undergoing series of processing is called information or simply put; information is generated when data have been subjected to processing or manipulation. This is the act of processing or manipulating data to get information. It can also be defined as the process of collecting all items of source original data together and converting these data into information or reports. The best data within the universe isn't much use if it is not processed. Processing refers to methods that take the data and switch it into usable information. Paper and pencil can work, but within the 21st century, data analysis usually relies on computers. To process data by computer, it's to be collected, checked for accuracy and entered into the pc first. This data can be processed in four ways: (I) Batch Processing, (II) Real-Time Processing, (III) Data Mining, (IV) Statistical Processing.

Before any data could be transformed into any meaningful information it must have been analyze by a specialist called data analyst. Data analysis is the complete process of constructing a data model from scratch and Data Models is a representation of the properties of the data within an existing or proposed system. There is also a critical need for automated approaches for effective and efficient utilization of massive amount of data to support corporate and individual's in strategic planning and decision-making. Hence, many strategic management information system models and relevant systems are still being developed in order to assist the strategic management mechanisms of organizations. Health Data Systems, otherwise referred to as Health information science, may be a discipline that mixes information processing, engineering and tending study. It covers all aspects of tending, like nursing, dentistry, hospital and alternative tending connected services. Hospital Data Systems (HDS) was chosen to limit the scope of the study; what is more, Health Information System has an associate degree instance of health data systems with the hospital because the tending setting

(Haux 2006). Hence, several strategic management data system models and relevant systems are still being developed so as to help the strategic management mechanisms of organizations. Healthcare business (hospitals) these days generates massive amounts of complicated knowledge concerning patients, hospitals resources, sickness identification, electronic patient records, medical devices among others (Aggarwal et al., 2012), this makes extracting hidden information from medical information complicated and longer intense. Analyzing this knowledge is vital for medical call manufacturers and managers; however, few tools exist within the tending setting to research the information totally to work out simplest practices (AlZoubi, 2013; Shaker et al., 2013). The large amounts of data or information are key resource to be processed and analyzed for knowledge extraction that allows support for cost-savings and deciding. Shaker et al (2013) noted that clinical choices are typically created supported doctors' intuition and experience skill, instead of on the knowledge wealthy knowledge hidden within the information. This observes ends up in unwanted biases, errors and excessive medical prices that affect the standard of service provided to patients. In recent years, data processing has been utilized in all areas of science and engineering, as an example in bioinformatics, genetics, drugs and power engineering. Conjointly individuals from business notice additional and additional applications for data processing, most applications are found in finance and insurance, retail, telecommunication and security (Seng and bird genus, 2010). In the tending sector data mining application is but still barely used. Plenty of data is hidden within the inheritance systems which might simply be extracted. Data Mining technology brings a group of tools and techniques which will be applied to the current processed knowledge to get hidden patterns and provides tending professionals an extra supply of data for creating choices. Research shows that successful decision systems enriched with analytical solutions are necessary for HIS (Zuckerman and Alan, 2006; Armoni, 2002; Rada, 2002). Hence, Data mining technology offered capabilities to increase the productivity of medical personnel, analyze care outcomes, lower healthcare costs, improve healthcare quality by using fast and better clinical decision making and generally assist the strategic management mechanisms.

Related Works

Numerous researches have been carried on Clinical Decision Support System (CDSS). Surveys of such work can be found in Shahsavarani et al (2015). Some works also reported the use of data mining techniques to find new patterns and knowledge from biomedical data. Such includes the following:

Megalooikonomou et al. (2000) introduced statistical methods that aid the discovery of interesting associations and patterns between brain images and other clinical data. Brossette et al. (2000) designed a Data Mining Surveillance System (DMSS) that uses novel data mining techniques to discover unsuspected, useful patterns of nosocomial infections and antimicrobial resistance from the analysis of hospital laboratory data. Antonie et al. (2001) investigated the use of different data mining techniques for anomaly detection and classification of medical images. Delen et al. (2005) used two popular data mining algorithms (artificial neural networks and decision trees) along with a most commonly used statistical method (logistic regression) to develop the prediction models on breast cancer using a large dataset. Su et al. (2006) used four different data mining approaches to select the relevant features from the data to predict diabetes. Phillips-Wren et al. (2008) assessed the utilization of healthcare resources by lung cancer patients' related to their demographic characteristics, socioeconomic markers, ethnic backgrounds, medical histories, and access to healthcare resources in order to guide medical decision making and public policy. A review of the above related works with literature analysis revealed that little or no work has been done in the area of the use of data mining in operational business processes, such as in Hospital Information System (HIS). Therefore, this work developed a HIS that incorporated data mining model called SHIMS in order to assist the strategic management mechanisms in hospital operational business processes.

RESULT AND DISCUSSION

Data Mining

Data mining fondly called patterns analysis on large sets of data, is the process of finding the patterns, associations or relationships among data using different analytical techniques involving the creation of a model and the concluded result will become useful information or knowledge (Ting et al., 2009). Lawal et al (2015) and Feelders et al (2000) noted that data mining is a process that allows users to extract patterns of knowledge from large data sets through the use of algorithms and techniques drawn from the field of Statistics, Machine Learning and Database Management Systems. It uses variety of tools like query and reporting tools, analytical processing tools, and Decision Support System (DSS) tools. Data mining technique have been used to uncover hidden patterns and relations, to summarize the data in novel ways that is both understandable and useful to the executives, to forecast for the trends and behaviors in business (Goksen et al., 2011). Koh and Tan (2005) reported that data mining applications can greatly benefit all parties involved in the healthcare industry. 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 techniques can be classified based on the database, the knowledge to be discovered and the techniques to be utilized (El-Sappagh, 2013). Matkovsky and Nauta (1998), Ting et al. (2009) and El-Sappagh (2013) narrated that the aim of data mining is to learn from data and there are two broad categories of data mining strategies: Supervised Learning (Predictive) and Unsupervised Learning (Descriptive). In supervised learning, the goal is to predict the value of an outcome based on a number of input measures. It involves using some variables in data sets in order to predict unknown values of other relevant variables. The models and attributes are known and are applied to the data to predict and discover information (Obenshain, 2004; Ting, et al., 2009; Gorunescu, 2011; El-Sappagh, 2013). Classification, statistical regression and association rules building (dependency modelling) are very common supervised learning techniques used in medical and clinical research.In unsupervised learning, there is no outcome measure, and the goal is to describe associations and patterns among a set of input measures. It involves finding human understandable patterns and trends in the data. The attributes and models are not known, but the patterns and clusters of data uncovered by data mining can lead to new discoveries (Obenshain, 2004; Rupnik et al., 2007; Ting, et al., 2009; Gorunescu, 2011; El-Sappagh, 2013). Clustering and summarisation the common unsupervised techniques used in medical and clinical research (Gorunescu, 2011). During the past few years, researchers have tried to combine both unsupervised and supervised methods for the analysis (Hastie et al., 2001). Some examples of advanced supervised learning models are Classification and Regression Trees (CART) and Support Vector Machines (SVM) (Lee et al., 2005). Advanced examples of the unsupervised learning models are Hierarchical Clustering, C-means Clustering, Self-organizing Maps (SOM) and Multidimensional Scaling Techniques. Liao et al (2012) reported that data mining techniques developed recently includes generalization, characterization, classification, clustering, association, evolution, pattern matching, data visualization and meta-rule guided mining. The most used software tools that provide these techniques (algorithms) are SPSS/SPSS Clementine, Salford Systems CART/MARS/TreeNet/RF, Yale (open source), SAS/SAS Enterprise Miner, Angoss Knowledge Studio/Knowledge Seeker, KXEN, Weka (open source), R (open source), Microsoft SQL Server, MATLAB and Oracle Data Miner (Shaker et al., 2013). Krivda (1995) and Ting et al. (2009) noted that the typical data mining process involves transferring data originally collected in production systems (such as electronic medical records) into data warehouse, cleaning or scrubbing the data to remove errors and check for format consistency, and then searching the data using statistical model, artificial intelligence (such as neural networks), and other machine learning methods.

Data Mining Applications in Healthcare

There is vast potential for data mining applications in healthcare. Generally, these can be grouped as the evaluation of treatment effectiveness; management of healthcare; and Customer Relationship Management (CRM). More specialized medical data mining, such as analysis of DNA micro-arrays, lies outside the scope of this work.

1. Evaluation of treatment effectiveness: Data mining applications can be developed to evaluate the effectiveness of medical treatments (evidence based medicine), by comparing and contrasting the causes, the symptoms and the course of treatments, data mining can deliver an analysis of which courses of action prove effective such as predict optimum medication dosage.

2. 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. It can search for patterns that might indicate an attack by bio-terrorists. Moreover, this system can be used for hospital infection control, or as an automated early-warning system in the event of epidemics. Accurate prognosis and risk assessment as survival analysis for AIDS patients, predict pre-term birth risk, determine cardiac surgical risk, predict ambulation following spinal cord injury, and breast cancer prognosis.

3. CRM: Customer interactions may occur through call centers, physicians' offices, billing departments, inpatient settings, and ambulatory care settings. It determines the preferences, usage patterns, and current and future needs of individuals to improve their level of satisfaction.

Data Mining Process Models in Healthcare

Knowledge discovery is a process, and not a one-time response of the KDD system to a user's action. As any other process, it has its environment, its phases, and runs under certain assumptions and constraints. Figure 1. shows a typical decision making environment. In the healthcare environment, the source data in clinical databases and/or EHRs can be queried directly using SQL. A Data Warehouse (DW) can be created to integrate data from many sources and enhance data quality. Analysts can apply On-Line Analytical Processing (OLAP) tools on the DW. Data warehouse is not enough for data analysis. Data mining is required to discover hidden patterns in either EHR or DW. KDD is an iterative or cyclic process that involves a number of stages. Although the specific techniques may vary from project to project, the basic process is the same for all KDD problems.

Data Mining Challenges in Healthcare

Application of KDD in healthcare faces many challenges as (Yang, 2006; Ruben and Canlas, 2009):

1. Need for algorithms with very high accuracy because it is an issue of life or death. The problems of missing, noisy and inconsistent data complicate this objective. Moreover, unlike other (arguably simpler) domains, the medical discipline itself is diverse, complex and, to an outsider, relatively opaque.
2. Active data mining: It needs two types of triggers. First type used to fire data mining technique to analyze the data automatically after some time or making some updates. The second type is to enforce the discovered knowledge by embedding the discovered knowledge within clinical information systems.

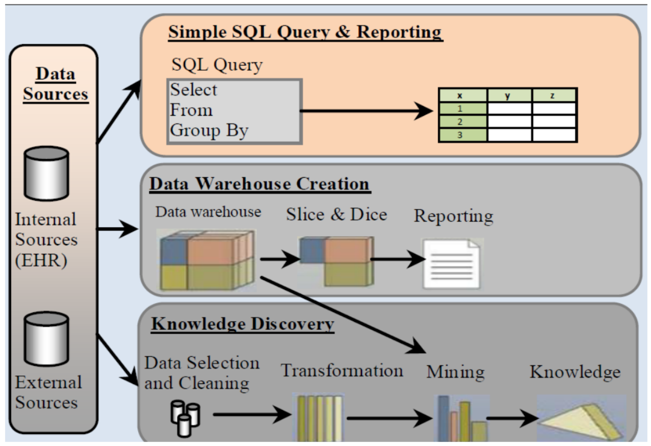


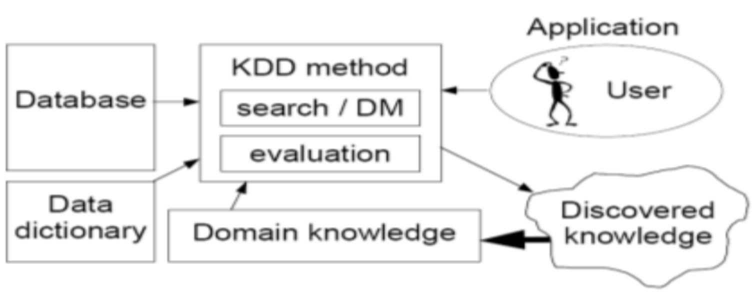
Figure 1: Decision Making Environment

3. For the same problem, there is need to apply many data mining techniques, compare their results and select the most interesting one.

4. Automation or semi-automation KDD system: this challenge is critical since EHR data and medical knowledge is changed continuously. KDD process needs to start automatically or with a minor help from analyst.

5. Previous (background) knowledge as concept hierarchy, domain expert knowledge, previously mined knowledge and rule template must be taken in accounting data preparation, modeling and model evaluation phases (Alonso, Martínez, Pérez, and Valente, 2012). Background knowledge can be expressed in different formats: examples may be found in the aspect of decision rules, the Bayesian models, the fuzzy sets and concept hierarchies. ***Figure 2.*** shows that background knowledge must be taken in to account in the KDD process. The result of data mining system must be appended in the existing knowledge base (knowledge fusion). The new knowledge may update, remove, or add constraints to the existing knowledge.

6. Longitudinal, temporal and spatial support: There is a need for advancement in data mining techniques to deal with EHR environment (Hripcsak, Knirsch, Zhou, Wilcox and Melton, 2011). EHR has clinical data collected and summarized from many healthcare systems, temporal data forming patient history, social network of patients' (Patient ID, Mother ID, Father ID, and Family ID), and spatial data as patient addresses. EHR may contain data with different formats as audio, video, image and text data. These formats also need advanced data mining techniques. The use of static techniques thus oversimplifies (or may hide) possible relationships and thus support for longitudinal, temporal and spatial semantics within the mining process is highly desirable. Episodic



data is often the key to good data mining. Techniques as anomaly detection (Li, Wu, Jajodia, and Wang, 2002), difference detection (Anthony, Hongjun, Jiawei and Ling 2003), longitudinal X analysis (Anthony, Hongjun, Jiawei and Ling, 2003) and temporal/spatial X analysis (Freksa, 1992) are used.

Figure 2: Background Knowledge Role in KDD Process

7. Mining complex knowledge from complex data: Using multi-relational data mining, knowledge mining in the form of graphs and mining non-relational data as text and image are future issues. Using text mining is

challenging in analysis of physician free text describing patient diagnoses and free text prescription. It can be used to summarize patient conditions.

8. Utilizing data mining in CDSS creates a system of real-time data-driven clinical decision support, or "adaptive decision support." It can "learn" over time, and can adapt to the variation seen in the actual real-world population. The approach is two pronged – developing new knowledge about effective clinical practices as well as modifying existing knowledge and evidence-based models to fit real-world settings (Bennett, and Doub, 2011).

9. Most KDD systems today work with relational databases. KDD and data mining need to be extended to object-oriented and multimedia databases.

10. Various security and privacy aspects: how to ensure the users' privacy while their data is being mined? One solution is *Anonymousness and identification transformation.* Besides data preprocessing, in order to separate the relative between patients' and their data which may have some private information, anonymousness and identification transformation are also needed.

11. Distributed data mining, mining heterogeneous and multi-agent data: many schemas can be utilized as collecting data from all sites and then mine it. In mining data distribution, one problem is how to mine across multiple heterogeneous data sources such as multi-database and multi-relational mining.

12. Large databases: clinical databases are very large and massive with hundreds of tables and fields, millions of records. Using more efficient algorithms, sampling, approximation, feature selection, and parallel processing can mitigate the problem.

13. High dimensionality: clinical environment has a very large number of fields. It can be solved by using prior knowledge to know irrelevant variables.

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. This information includes patient demographics, records on the treatment progress, details of examination, prescribed drugs, previous medical history, lab results, and so on. 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 (Canlas, 2009). Before data mining process can begin, healthcare organizations must formulate a clear policy concerning patient records confidentiality and security. 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, and so on (Koyuncugil, and Ozgulbas, 2010).

Integration of data mining in information systems, healthcare institutions reduces subjectivity in decision-making and provides new useful medical knowledge. Predictive models provide the best knowledge support and experience to healthcare workers. Data mining is using a technique of predictive modeling to determine which diseases and conditions are the leading trends. This requires a review of medical documentation of a healthcare institution and prescription drugs to determine which problems are the most common amongst patients. The problem of prediction in medicine can be divided into two phases: learning phase and the phase of decision making. In the learning phase, a large data set is transformed into a reduced (simplified) data set. Number of features and objects in this new set is much smaller than the original set in several different ways. The rules generated in this phase is used later to make accurate decisions. Newly formed data set is used to make predictions when the new instances with unknown outcomes occur with the predictive algorithm. This algorithm compares the characteristics of a new object with the characteristics of objects represented in the selected data set. If the match is found, the new object gets the outcome which is equal to the corresponding object in the set.

The goal of predictive data mining in medicine is to develop a predictive model that is clear, makes reliable predictions and helps doctors to improve their prognosis, diagnosis and treatment planning procedures. Important questions arise here (Bellazzi, and Zupan, 2008): if the data and the corresponding predictive characteristics are sufficient to build a predictive model of acceptable performance; what is the relation between attribute and outcome; can it be found an interesting combination, or the relation between attributes; whether the immediate factors can be drawn from the original attributes that can increase the performance of predictive models. A very important application of data mining is for biomedical signal processing expressed by internal regulations and responses to the stimulus conditions, whenever there is a lack of detailed knowledge about the interactions between different subsystems, and when the standard analysis techniques are ineffective, as it is often the case with non-linear associations. Data mining provides the link between knowledge of continuous data, 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 (Candelieri, Dolce, Riganello, and Sannita, 2011). Using association rules involves finding all the rules, or at least part of key subsets of rules that is characteristic of certain information as consequences or as 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 is used to help strengthen the arguments regarding whether to engage or eliminate certain rules in the knowledge model (Houston, Chen, Hubbard, Schatz, Ng, Sewell, and et al 1999).

Tasks of the managers' that manages 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 is quality of data, standards, plans, and treatments. Data mining can be used by quality managers' to solve the following tasks (Stühlinger, Hogl, Stoyan, and Müller, 2000):

1. Discovering new hypothesis for indexes of quality for data, standards, plans and treatments.
2. Checking if the given indexes of quality for data, standards, plans and treatments are still valid.
3. Improving, strengthening and adjusting of quality indexes for data, standards, plans and treatments.
4. These tasks can be supported by data mining if the existing knowledge in domain is seriously considered in data mining process.

Building the Structure of Data Mining in Hospital Information Management System (DMHIMS)

No system works on its own without some basic components. For example, the human digestive system has physical human parts like the mouth, the esophagus, the intestines and the anal canal as its structure. Equally, the human flesh and other organs in the body are hung on a structure called *skeleton*. In the same vein, for every DMHIMS there is some functionalities that constitute the system (Callon, 1996). The DMHIMS is therefore anchored on computers and telecommunication components including the human elements, rules and policies. These components are basically hardware and software. According to Leary and Leary (2005), there are five components that form the backbone of DMHIMS. They listed and explained them as follows:

- **a. Hardware:** Hardware consists of all tangible elements of a computer system. Typical examples are the input devices, the components that store and process data and perform required calculations and the output devices that present the results to information users are hardware's. Input devices allow users to enter data and commands for processing, storage, and output. Common input devices include the keyboard, mouse, scanner, modem, microphone, touch screen, etc. Storage and processing components consist of the hard drive, diskette drive, Zip drive, and CD-ROM drive. The newer CD-RW or DVD-RW drives can write disks as well as read them. This small portable device can store up to 128 MB of data when plugged into a computer USB port (Boone & Curtz, 2006).
- **b. Software:** Software is all programmes, routine and computer languages that control a computer and tell it how to operate. Over 80 per cent of personal computers use a version of Microsoft's popular Windows operating system. Personal computers made by Apple use the Mac operating system. Handheld computers use either the Palm operating system or a special version of Windows called Pocket PC. Other operating system include UNIX, which runs many microcomputers, and Linux, which is available for free in the public domain (Boone & Kurtz, 2009).
- **c. Connectivity:** Connectivity allows computers to connect and share data and information. They are mainly communication facilities such as telephones lines, modulator (modem) and demodulators. It can also be cables and wireless means. All put together, connectivity facilities are called Information Technology (IT). Examples include telephone lines, coaxial wire, fibre optic, etc.
- **d. Procedures:** Procedures are the rules or regulation that guides people when using DMHIMS such as the software's and hardware's manuals, policies guiding users and laws they should be aware of. While procedures and manual are written by software and hardware manufacturers, laws and policies do emanate from government and companies that have DMHIMS.
- **e. People:** People are the end users like you and me, who interface between the structure and the work environment. For organizations, they are mostly the employees who use the DMHIMS to achieve the organization's information management objectives.

Information Systems

Information Technology (IT) has become the cutting edge of global competition. Companies and organizations are keen to invest in information technology due to its potentials as a strategic enabling tool to support growth and enhance quality. IT is the area that manages technology which typically comprises of computer science, information systems, computer hardware, software, programming languages, network and many more. Conversely, Information Systems (IS) is a discipline that unites the business and computer science domain (Abdullah, 2013). Silver et al. (1995) stated that "information systems are implemented within an organization for the purpose of improving the effectiveness and efficiency of that organization". Hence, the blend between people, organization and technology is the major concern in IS. Despite the differences between IT and IS, in most literature, these two terms are used interchangeably (Lee, 2004). In this study, IS and IT are treated alike. There is an abundance of IS domain literature merging between the business and technology realms. The business realm covers researchers' in relation to business processes, business modeling, IT governance, IT management and others. The technology domain on the other hand, encompasses areas from IS development to IS deployment. IS research studies also varies according to applications such as decision support systems, knowledge management systems, database management systems, accounting information systems, manufacturing information systems, health information systems, transaction processing systems and many more. Other areas of IS studies include research on methodology, analysis and design, and security (Abdullah, 2013). Figure 2.3 illustrates the diverse research areas in the IS domain. Each IS research area has its own followers' or research groups that sometimes intersects with other domains. This makes it difficult to streamline the areas of research in IS. The current study was devised to focus only on IS development, particularly for the health domain. In the next section, Health Information System concept is considered.

Health Information System (HIS)

The term Health Information Systems is the label given to information systems in the entire healthcare domain and the term Hospital Information Systems (HIS) focuses specifically on the information systems within hospitals, the scope and definition of Hospital Information Systems (HIS) is seen as a subset of Health Information Systems. In the current study the term Hospital Information Systems (HIS) is preferred, though some researchers use the terms interchangeably (Kuhn and Giuse, 2001). There are alternative definitions of HIS. For example, Aggelidis and Chatzoglou (2008) defined HIS as a computer-based system designed to facilitate the management of the administrative and medical information within a hospital. Biohealthmatics.com (2006) defined HIS as a computer system that is designed to manage all the hospital's medical and administrative information in order to enable health professional perform their jobs effectively and efficiently. While Haux (2004) defined HIS as the socio-technical subsystem of a hospital, which comprises all information processing as well as the associated human or technical actors' in their respective information processing roles. Basically, HIS have subsystems such as Clinical Information System (CIS), Financial Information System (FIS), Laboratory Information System (LIS), Pharmacy Information System (PIS), Radiology Information System (RIS), Picture Archiving Management System (PACS) and the Nursing Information System (NIS) (Biohealthmatics.com 2006). The 'big bang' installs the entire system at once whereas an incremental approach allows users to implement a few subsystems at a time. An incremental approach is greatly recommended because it allows the users to become accustomed to the new system gradually. Many authors caution about the danger of the 'big bang' approach since it does not permit users to adapt progressively to change (Anderson and Stafford, 2002; Jones, 2003; Ludwick and Doucette, 2009). Among the benefits of an HIS system are automation tasks, fast retrieval of records, simplifies projections tasks, improves productivity, speedily assesses bed vacancies, hastens lab test results, shortens waiting time, schedules appointments, avoids misplaced records, reduces errors and provides a repository of valuable information (Bulgiba, 2004). Bulgiba (2004) added that the benefits of HIS aid in achieving enhanced patient safety aspects because life-saving decisions may be performed faster and inaccurate drug prescriptions may be minimized.

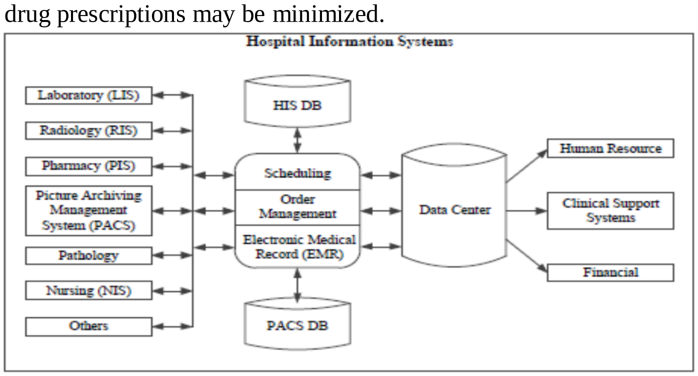


Figure 3. HIS and its Subsystems (Source: Adapted from Wainwright and Waring, 2000; Abdullah, 2013).

CONCLUSIONS

This research work has shown that the classical data mining methodologies can be deployed in HIMS to support analytical decision making process. Hence, the use of data mining technology in HIMS offer obvious advantages over traditional HIMS called legacy systems. When analytical technologies are embedded in Modern HIMS that gather and centrally store large amounts of data, this will assist the strategic management mechanisms and improvements in clinical decision-making. Adopting DMHIMS will offers capabilities to increase the productivity of medical personnel, analyze care outcomes, lower healthcare costs, improve healthcare quality by using fast and better clinical decision making and generally support the analytical decision-making process.

REFERENCES

Abraham, C., Nishihara, E. and Akiyama. M. (2011). Transforming Healthcare with Information Technology in Japan: A Review of Policy, People, and Progress. *International Journal of Medical Informatics*,80 (3): 157-170.

Aggarwal, N., Kumar, A., Khatter, H. and Aggarwal, V. (2012). Analysis the effect of data mining techniques on database. *Advances in Engineering Software*, 47(1): 164-169.

Buntin, M.B., Burke, M. F., Hoaglin, M. C. and Blumenthal, D. (2011). The Benefits of HealthInformation Technology: A Review of the Recent Literature Shows Predominantly Positive Results. *Health Affairs*,30 (3): 464-471.

Chiasson, M.W. and Davidson, E. (2005). Taking Industry Seriously in Information SystemsResearch. *MIS Quarterly*, 29 (4): 591-605.

Claver, E., González, R. and Llopis, J. (2000). An Analysis of Research in Information Systems(1981–1997). Information & Management, 37 (4): 181-195.

Dutta, A, and Heda, S (2000). Information systems architecture to support managed care business process. *Decision Support Systems*, 30: 217-225.

Goksen, Y., Eminagaoglu, M. and Dogan, O. (2011). Data Mining in Medical Records for the Enhancement of Strategic Decisions: A Case Study. *Scientific Bulletin – Economic Sciences*, 10(1): 135-144.

Gorunescu, F. (2011). Data Mining: Concepts, Models, and Techniques. India: Springer.

Guleria, P. and Sood, M. (2014). Data Mining in Education: A Review on the Knowledge Discovery Perspective. *International Journal of Data Mining & Knowledge Management Process*, 4(5): 47- 60.

Hastie, T., Tibshirani, R. and Friedman, J. (2009). The elements of statistical learning: Data mining, inference, and prediction. Second Edition. Springer, New York, USA.

Haux, R. (2004). Strategic Information Management in Hospitals: An Introduction to Hospital Information Systems. New York, NY, USA: Springer Verlag.

Haux, R. (2006). Health Information Systems - Past, Present, Future. *International Journal of Medical Informatics*,75 (3-4): 268-281.

Health Information and Management Systems Society (2008). Enabling Healthcare Reform using Information Technology. http://www.himss.org.

Khademolqorani, S. and Hamadani, A. Z. (2013). An Adjusted Decision Support System through Data Mining and Multiple Criteria Decision Making. *Procedia - Social and Behavioral Sciences*, 3: 388-395.

Lavrac, N., Mladenic, D., Bohanec, M. and Moyle, S. (2003). Data Mining and Decision Support: Integration and Collaboration. Springer.

Lawal, N. T. A, Odeniyi, O. A. and Kayode, A. A. (2015): Application of Data Mining and Knowledge Management for Business improvement: An Exploratory Study. International *Journal of Applied Information Systems*, 8(3):13-19.

Lee, J. W., Lee, J. B., Park, M. and Song, S. H. (2005). An extensive comparison of recent classification tools applied to microarray data. *Computational Statistics & Data Analysis*, 48(4): 869–885.

Liao, S-H., Chu, P-H. and Hsiao, P-Y. (2012). Data mining techniques and applications – A decade review from 2000 to 2011. *Expert Systems with Applications*, 39:11303- 11311

Li, J., Wei, L., Li, G. and Xu, W. (2011) An evolution strategy-based multiple kernels multi criteria programming approach: The case of credit decision making, *Decision Support System*, 51: 292-298.

Ludwick, D.A. and Doucette, J. (2009). Adopting Electronic Medical Records in Primary Care:Lessons Learned from Health Information Systems Implementation Experience in Seven Countries. *International Journal of Medical Informatics*, 78(1): 22-31.

Maclennan, J., Tang, Z. and Crivat, B. (2008). Data mining with Microsoft SQL server 2008,Indianapolis, Wiley.

MIT Sloan. (2012). Center for Information Systems Research - MIT Sloan School of Management. Accessed September 05, 2020, http://cisr.mit.edu.

Novak, J. and Judah, A. (2011). Towards a Health Productivity Reform Agenda for Australia.http://www.achr.com.au.

Obenshain, M. K. (2004). Application of Data Mining Techniques to Healthcare Data. In: Statistics for Hospital Epidemiology, Birnbaum, D. (ed.), 25(8): 690-695.

Rada, R. (2002). Information systems for Healthcare enterprises. Hypermedia Solutions Limited.

Rupnik. R., Kukar. M. and Krisper, M., (2007) Integrating data mining and decision support through data mining based decision support system. *Journal of Computer Information Systems*, 47(3):89-104.

Rupnik, R. and Jaklic, J. (2009). The Deployment of Data Mining into Operational Business Processes, In Data Mining and Knowledge Discovery in Real Life Applications, Julio Ponce and AdemKarahoca (Ed.), InTech, pp. 373-388.

Sakthimurugan, T. and Poonkuzhali, S. (2012). An Effective Retrieval of Medical Records using Data Mining Techniques. *International Journal of Pharmaceutical Science and Health Care*, 2 (2): 72–77.

Seng, J.-L. and Chen, T. C. (2010). An analytic approach to select data mining for business decision. *Expert Systems with Applications*, 37, 8042-8057.

Shaker, H., El-Sappagh, S., El-Masri, A. M. and Riad, M. E. (2013). Data Mining and Knowledge Discovery: Applications, Techniques, Challenges and Process Models in Healthcare. *International Journal of Engineering Research and Applications*, 3(3): 900-906.

Sheng, O. R. Liu (2000). Decision support for healthcare in a new information age. *Decision Support Systems,* 30:101-103.

Ting, S. L., Shum, C. C., Kwok, S. K., Tsang, A. H. C. and Lee, W. B. (2009). Data Mining in Biomedicine: Current Applications and Further Directions for Research. *Journal of Software Engineering & Applications*, 2, 150-159