

DEVELOPMENT OF A STRATEGIC HOSPITAL INFORMATION MANAGEMENT SYSTEM

^{*1}Odeniyi O. A.,

²Akinwumi A. R.,

³Makinde B. A. &

⁴Oyeniran O. R.

^{1,2,3} Department of Computer Science, Osun State College of Technology, Esa-Oke, Osun State, Nigeria

⁴Department of Electrical & Electronic Engineering, Osun State College of Technology, Esa-Oke, Osun State, Nigeria

*Corresponding author E-mail: olufemiodeniyi@gmail.com;
odeniyioa@oscotechesaoke.edu.ng;

Abstract

Wide spread use of information system in the delivery of managed healthcare and 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 development of a web-based Strategic Hospital Information Management System (SHIMS) based on data mining technology. The designed web-based SHIMS was developed using Hypertext Protocol (PHP) as the programming language, Microsoft SQL Server 2008 as the database server and Apache as web server, and implemented using three-tier architecture. The developed SHIMS 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: Data Mining, Information System, Health Information System, Hospital Information System, Decision Making, Strategic Management.

1. Introduction

For years, many industries have turned to information systems (IS) as a means of obtaining competitive edge; the health industry is no exception. In fact, improving health

information systems is one of the prime agendas in most health reform activities throughout the world (Abraham et al., 2011; Buntin et al., 2011; Health Information and Management Systems Society 2008; Novak and Judah 2011). The goal of health information systems is to deliver efficient patient care and quality services (Haux 2004). Health Information Systems, otherwise known as Health Informatics, is a discipline that combines information science, computer science and healthcare study. It covers all aspects of healthcare, such as nursing, dentistry, hospital and other healthcare related services. Hospital Information Systems (HIS) was chosen to limit the scope of the study; moreover, HIS as an instance of health information systems with the hospital as the healthcare environment (Haux 2006). As the name implies, this type of information systems was designed to manage hospital services like medical, administrative, financial and others.

The importance of strategic management information system in the delivery of managed healthcare can hardly be overemphasized (Hampshire and Rosborough, 1993). A variety of strategic management information system capabilities will be necessary to increase the productivity of medical personnel, analyze care outcomes, and continually refine care delivery processes to remain profitable while holding the line on costs and maintaining quality of care (Dutta and Heda, 2000).

HIS is faced with the challenges of identifying and disseminating relevant healthcare information (Sakthimurugan and Poonkuzhali, 2012), complex and diverse data and knowledge forms and tasks (Prins and Stegwee, 2000; Sheng, 2000), the lack of standardized terminology compared to basic sciences, the stringent performance and accuracy requirements and the prevalence of legacy systems (Sheng, 2000). A lot of information which is hidden in the legacy systems cannot be easily be extracted, because legacy systems are not build to answer questions that are unpredictable,

There is also a critical need for automated approaches for effective and efficient utilization of massive amount of data to support corporate and individuals 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.

Healthcare industry (hospitals) today generates large amounts of complex data about patients, hospitals resources, disease diagnosis, electronic patient records, medical devices among others (Aggarwal et al., 2012), this makes extracting hidden knowledge from medical database complex and more time consuming. Analyzing these data is critical for medical decision makers and managers, but few tools exist in the healthcare setting to analyze the data fully to determine the best practices and the most effective treatments (AlZoubi, 2013; Shaker et al., 2013).

The large amounts of data is a key resource to be processed and analyzed for knowledge extraction that enables support for cost-savings and decision making. Shaker et al (2013) noted that clinical decisions are often made based on doctors' intuition and experience rather than on the knowledge rich data hidden in the database. This practice

leads to unwanted biases, errors and excessive medical costs which affects the quality of service provided to patients.

The literature review reveals that Data Mining (DM) is progressively used in Information Systems (IS) as a technology to support decision making on the tactical level, as well as to enable decision activities within operational business processes. It is predominantly used to support analytical decision making, which is typically based on models acquired from a huge quantities of data and therefore makes possible to acquire patterns and knowledge (Rupnik et al., 2007; Rupnik and Jaklic, 2009).

In general, there are three categories of business decision making approaches: (i) decision making based on precisely defined business rules, (ii) analytical decision making based on the analysis of information and (iii) decision making based on intuition. In many cases all these three categories are used at a time. Business rule based decision making is typical for operational business processes, whereas other two are typical for managerial processes (Rupnik and Jaklic, 2009).

The use of DM to facilitate decision support enables new approaches to problem solving by discovering patterns and relationships hidden in data and therefore enabling an inductive approach to decision making. In addition, the decision to collaborate DM and decision making is also based on the fact that the traditional use of DM through its software tools does not bring it closer to business users due to the complexity of DM tools. The tools are very complex and needs expertise in DM technology. In the other words, understanding the DM algorithms and parameters for implementing in a large data set is a great task (Bawden, 2005; Lavrac et al., 2003).

In the integrating approach (collaborating DM and decision making), knowledge is automatically extracted from a large amount of row data by employing DM technologies, structured and transformed into appropriate forms, and finally retrieved by a hybrid reasoning engine in order to provide decision support for decision makings in special situations (Khademolqorani and Hamadani, 2013).

In recent years, data mining has been used in all kinds of areas of science and engineering, for instance in bioinformatics, genetics, medicine and electrical power engineering. Also people from business find more and more applications for data mining, most applications are found in finance and insurance, retail, telecommunication and security (Seng and Chen, 2010).

In the healthcare sector data mining application is however still barely used. This is odd as in these environments traditional HIS called legacy systems are used. A lot of information is hidden in the legacy systems which can easily be extracted. Most of the times this cannot be done directly from the legacy systems, because these are not build to answer questions that are unpredictable. Most healthcare institutions lack the appropriate information systems to produce reliable reports with respect to other information than purely financial and volume related statements (Prins and Stegwee, 2000).

Modern HIS gather and centrally store large amounts of data and therefore DM seems to be a good technology to support the data processing, information retrieval and

knowledge generation process in order to support decision making. DM brings a set of tools and techniques that can be applied to this processed data to discover hidden patterns. It provides healthcare professionals an additional source of knowledge for making decisions.

In the turmoil between costs, care-results and patient satisfaction the right balance is needed and can be found in collaborating DM and decision making within HIS. Research shows that successful decision systems enriched with analytical solutions are necessary for HIS (Zuckerman and Alan, 2006; Armoni, 2002; Rada, 2002). This could reduce medical errors, enhance patient safety, decrease unwanted practice variation, and improve patient outcome.

There are several algorithms and methods to mine medical/clinical data which are implemented in HIS. Using these methods have often resulted in surprising improvements in clinical decision-making. These algorithms or methods could be implemented in separate software environments (such as WEKA and IBM SPSS Modeler) and results imported to HIS, or designed as a part within HIS software engine.

Some recent technology achievements, such as SQL Server 2008 Analysis Services enables the possibility to develop application systems which utilize data mining methods and as such do not demand expertise in data mining technology for business users. The SQL Server 2008 Analysis Services which is an add-in for the widely used Microsoft SQL Server 2008 software has several benefits over traditional data mining applications (MacLennan et al., 2008). These include: ease of use, a simple but rich API and scalability (ability to handle large data sets), analysts do not have to build sample sets and the algorithms operate on the whole data set to provide the best results (Wang, 2010).

As evident from the discussion above, the classical data mining methodologies have to be extended for the cases where DM is deployed in HIS to support analytical decision making process. Hence, this paper presents the development of such a web-based HIS enriched with analytical solutions, called SHIMS. The rest of this paper is organized as following: Section 2 presents the background material on which the proposed methodology is based. Section 3 displays the methodology adopted to complete the proposed work. Section 4 shows the results of the proposed work, and Section 5 concludes the paper.

2. Background

In this section we briefly introduce areas that are important for the mission of the paper.

2.1 Data Mining

Data Mining (DM) 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 Data Base Management Systems. It uses variety of

tools like query and reporting tools, analytical processing tools, and Decision Support System (DSS) tools.

DM techniques have been used to uncover hidden patterns and relations, to summarize the data in novel ways that are both understandable and useful to the executives, to predict future 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, DM 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.

DM 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) reported that the goal of DM is to learn from data and there are two broad categories of DM strategies: Supervised Learning (Predictive) and Unsupervised Learning (Descriptive). Table 1 shows the summary comparing the characteristics and the techniques used for the two different learning methods, while Table 2 outlines DM strategies by modelling objective, categorized by supervised and unsupervised distinctions.

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 DM 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).

Prediction algorithms determine models or rules to predict continuous or discrete target values given input data; Classification algorithms determine models to predict discrete values given input data; Exploration uncovers dimensionality in input data; while Affinity analysis determines which events are likely to occur in conjunction with one another. Further details on these DM characteristics and the techniques used for the two different learning methods can be found in (Obenshain, 2004; Ting et al., 2009; Li et al., 2011; El-Sappagh, 2013; Guleria and Sood, 2014).

Table 1: Comparing the characteristics and the techniques of the unsupervised and supervised learning (Ting et al., 2009)

Characteristics		Techniques
Supervised Learning	With guidelines Class defined Common with vast literature and application	Classification Statistical Regression Artificial neural networks
Unsupervised Learning	No guidance Use to Define the class Seldom utilized (until recently)	Clustering Association Rule

Table 2: Modelling Objectives and Data Mining Techniques (Obenshain, 2004)

Modelling Objective	Supervised	Unsupervised
Prediction	Ordinary Least Squares Regression Logistic Regression Neural Networks Decision Trees Memory-based Reasoning Support Vector Machines Multi-Adaptive Regression Splines	Not feasible
Classification	Decision Trees Neural Networks Discriminant Analysis Bagging and Boosting Ensembles Naïve Bayes Classifiers	Clustering (e.g. K means) Kohonen Networks Self-Organizing Maps
Exploration	Decision Trees	Principal Components Clustering (e.g. K means) Link Analysis
Affinity		Associations Sequences Factor Analysis

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 DM 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 DM 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.

2.2 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 paper, 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 1 illustrates the diverse research areas in the IS domain.

Each IS research area has its own followers or research groups that sometimes intersect 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.

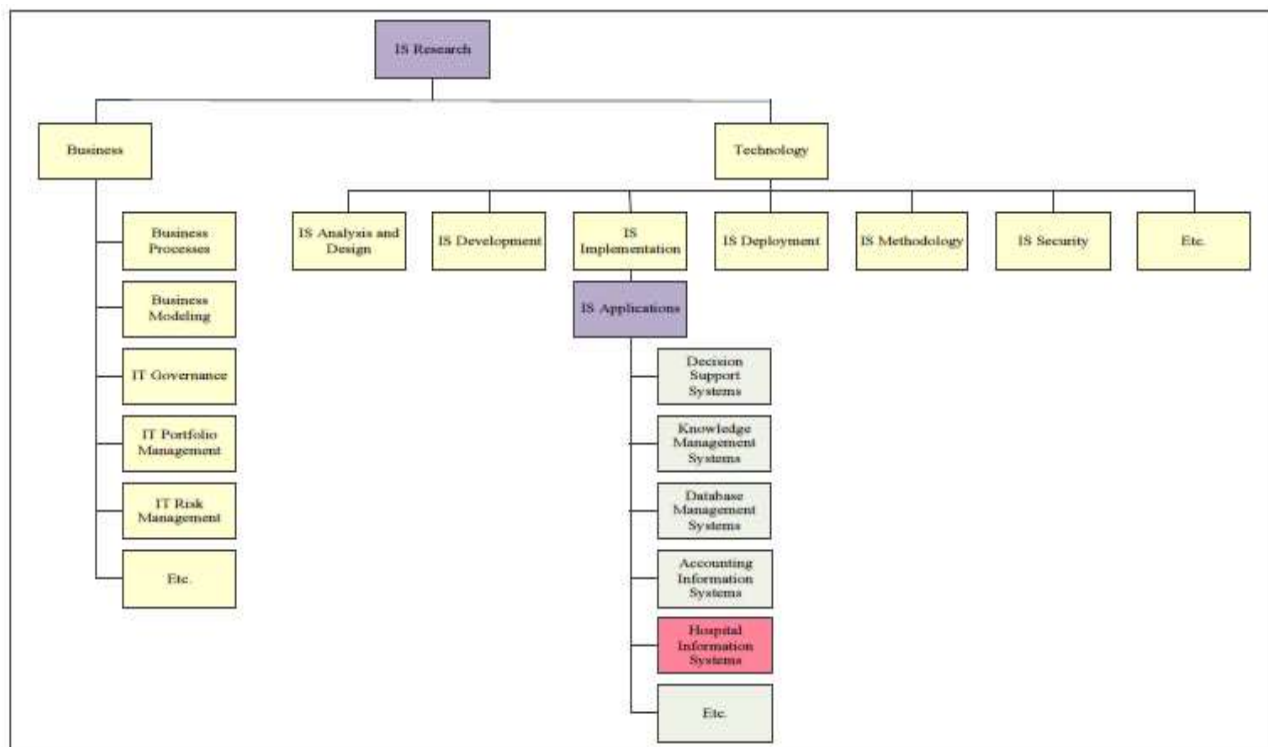


Figure 1: IS Research (Source: Adapted from Claver et al., 2000; Chiasson and Davidson, 2005; Walsham and Sahay, 2006; MIT Sloan, 2012; Abdullah, 2013)

2.3 Health Information System

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.

Typically, as depicted in Figure 2, 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). Figure 2 illustrates typical view of how the HIS subsystems connect. The number of subsystems to be installed is decided by the hospital management; implemented separately by an incremental approach or simultaneously by means of a big bang approach.

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). The descriptions for HIS subsystems and analogous health systems are described in Abdullah (2013).

Among the benefits of an HIS system are that it includes automate 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. Further, Goldschmidt (2005) supported the view that HIS lessens administrative expenses, improves the quality of patient care, automates sharing of health information and reduces the risk of negligence

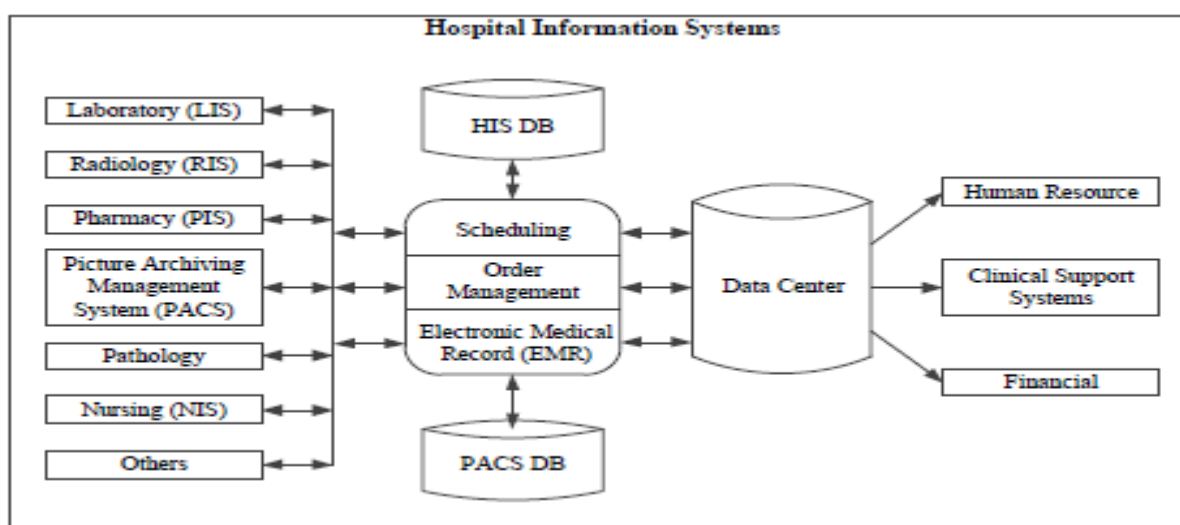


Figure 2: HIS and its Subsystems (Source: Adapted from Wainwright and Waring, 2000; Abdullah, 2013).

2.4 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 DM 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 DM 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 DM techniques for anomaly detection and classification of medical images. Delen et al. (2005) used two popular DM 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 DM 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.

Some works also discussed the use of DM in business processes in general and they also represent important basis for our research. Khademolqorani and Hamadani (2013) presents the recent study in this area. Chung and Gray (1999) argued that there is a lot of research done in the areas of DM model creation, but there is a lack of research done in the area of the use of DM models in operational business processes. Kohavi and Provost (2001) argued that it is important to enable the use of DM in business processes through automated solutions. They discussed the importance of the ease of integration of DM in business processes. In their paper, they discussed the integration of DM in business processes as the consequence of the need to incorporate background knowledge in business processes. They stated that deploying automated solutions to previously manual processes can be a rife with pitfalls. Authors also argued that one must deal with social issues when deploying automated solutions to previously manual processes.

Feelders et al. (2000) discussed the use of DM in business processes with the emphasis on the integration of DM models and solutions into existing application systems within IS. Authors argued that it is essential that the results of DM are used to support operational business processes like direct mailing for the selection of potential customers. Gray (2005) discussed DM as an option of knowledge sharing within the enterprise. Through the integration of DM in operational business processes, knowledge sharing is not only present in business processes on a tactical level, but also on operational level.

Rupnik et al. (2007) developed a DM process model and appropriate DM based decision support system to support decision processes in the telecommunication company. They explored the use of DM application systems approach. In this approach, the DM was not used in ad-hoc projects, but through DM based decision support systems. The aim of the work was to define how business users can use DM models to facilitate decision making where DM experts create DM models and business users use them.

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 DM techniques in operational business processes, such as in Hospital Information System (HIS). Therefore, this work developed a HIS that incorporated DM model called SHIMS in order to assist the strategic management mechanisms in hospital operational business processes.

3. Methodology

In this work, the System Development Life Cycle (SDLC) methodology or framework was adapted. From the engineering, information processing and information system development literature, the formal design of any information processing system is supposed to follow a set of steps labeled the System Development Life Cycle (SDLC). However, as Sprague and Carlson (1982) and other authors (Davis, 1983; Aktas, 1987) have noted, the traditional SDLC is somewhat difficult to use as originally specified. But there is agreement that the SDLC has four general or essential phases: (i) system analysis, (ii) system design, (iii) system construction or development and (iv) system implementation.

3.1 System Analysis

Analysis involved a detailed study of the current system and it includes sub dividing of complex process involving the entire system, identification of data store and manual process. During analysis, data are collected on available files, decision points and transactions handled by the present system. Interviews, case study, document and archives, on-site observations, questionnaire and review of literature are the tools used to establish the objectives, requirements and services expected from the system.

3.2 System Design

System design brings a proposed system much closer by describing the nature of input and output file as well as sharing the processing in which they are connected. These design elements consisted of functional hierarchy, business process diagrams, pseudo code and entity-relationship diagram. In this stage the design elements are defined with the help of interviews, case study, document and archives, on-site observations, questionnaire and review of literature conducted. These design elements provided detail description about the software and each element is related to a specific requirement

3.3 System Development

The development stage was initialized by the previous design stage. At this stage the code for the design elements of system was written using Hypertext Protocol (PHP) as the programming language and Microsoft SQL Server 2008 as the database server, and it provided the functional system components. PHP was chosen because it is an open source server-side scripting language designed for web development (but can also be used as a general-purpose programming language), and can interact with Microsoft SQL databases. Also, Microsoft SQL Server 2008 was chosen because it is a relational database management system (RDBMS) that supports a wide variety of transaction processing, business intelligence and analytics applications in corporate IT environments. It enables the creation, maintenance and management of database. Microsoft SQL Server 2008 is a Structured Query Language (SQL) based, client/server relational database.

3.4 System Implementation

To implement a web application, client-server architecture is required. The most popular client-server architectures are the two-tier and the three-tier architecture. The choice of architecture affects the development time and the future flexibility and maintenance of the application. While selecting the architecture most suitable for an application, many factors including the complexity of the application, the number of users and their geographical dispersion are considered.

This system was designed based on a traditional three-tier architecture used by many web applications. Three-tier architecture includes a presentation layer, business rules/logic layer, and the data layer. The Data layer is responsible for data storage. Primarily this tier (layer) consists of one or more relational databases and/or file systems. The Business Rules/Logic layer is the middleman between the presentation layer and the data layer. This middle tier was introduced to overcome the deployment limitation (whenever the application logic changed the application had to be redistributed at each and every client) in the two-tier architecture. The middle tier provides process management where business logic and rules are executed and can accommodate hundreds of users.

The Presentation Layer, also called the Client tier, is responsible for the presentation of data, receiving user events, and controlling the user interface. The user interaction with the system is entirely through this layer. The three-tier architecture is generally used when an effective distributed client/server design is needed that provides: increased performance, flexibility, maintainability, reusability and scalability. This model hides the complexity of distributed processing from the user. These features have made the three-tier architecture a popular choice over the two-tier architecture for Internet applications.

In addition, a web server is required to implement any web-based application. A web server is a piece of software that manages web pages and makes them available to the 'client' browser – via a local network or over the Internet. The web server can be accessed remotely or locally. There are many web servers available such as Apache,

Microsoft Internet Information Services (IIS), Netscape Web Server and many more. However, in this work Apache was chosen as web server (WAMP Server 1.7.0 running Apache Web Server version 2.2.4) because of its availability as open source software and compatibility with PHP, Microsoft SQL Server 2008.

4. Result and Discussion

The web-based Strategic Management Hospital Information System was developed using Hypertext Protocol as the programming language and Microsoft SQL Server 2008 as the database server and implemented using three-tier architecture as described in section 3 above. The system was tested using three testing techniques (White box, Usability and Black box) to check the validity, correctness, completeness and to evaluate the quality of the system.

The application provides features such as the passcode database, staff/employee database, patient database, patients' treatment database, extract patient information database, staff/employee information report, patients details report, patients' treatment report and extraction information report based on pattern. To gain access to the system every user should be authenticated. When the user types the website's URL in the browser, the login page is displayed as shown in Figure 3.

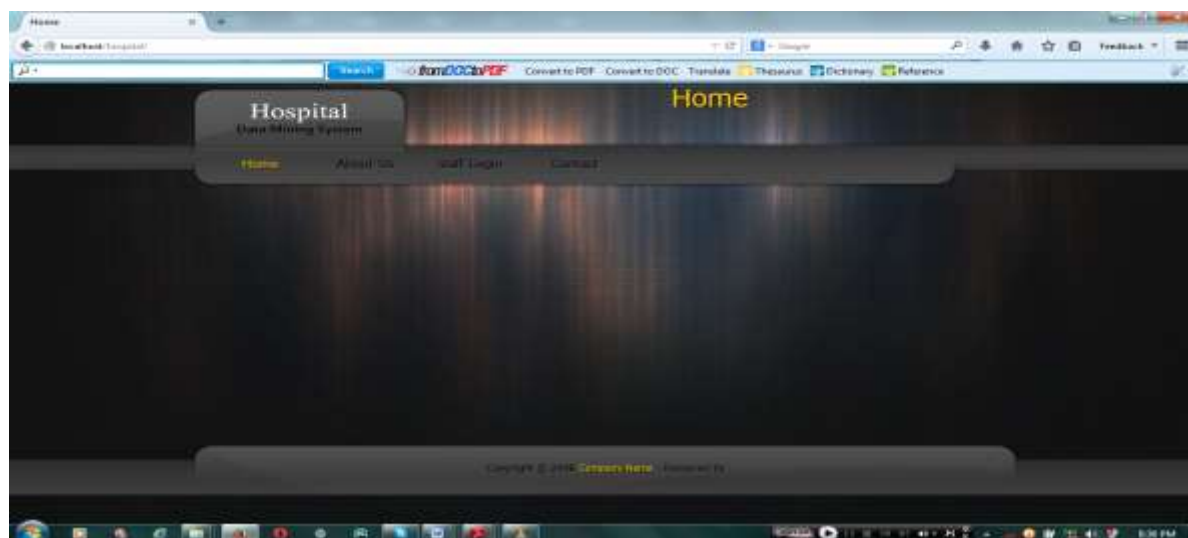


Figure 3: Home Page

Depending on the role selected in the login page the user is directed to the respective subsystem, the user (staff) subsystem, or the Administrator subsystem. Some of the other graphical user interface of the developed system is depicted in Figures 4 to 10.

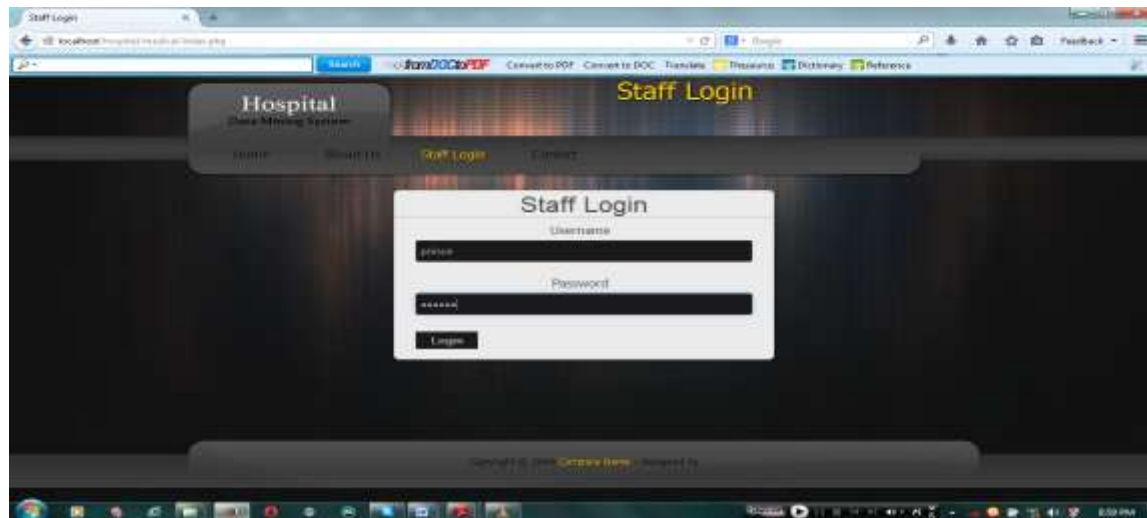


Figure 4: Login Page



Figure 5: Registered Patient Details

The screenshot shows a web browser window with the URL 'localhost:192.168.1.100/hospital_data_min...'. The page has a dark header with 'Hospital Data Mining System' on the left and 'User prince' on the right. Below the header is a 'New Patients Information' form with the following fields: Surname (text input), Other names (text input), Gender (dropdown menu showing 'MALE'), DOB (date picker showing '01/01/1980'), Phone Number (text input), and Address (text input). The form is set against a background image of a city skyline at night.

Figure 6: Patient's Biodata Registering

The screenshot shows the 'SEARCH RESULT' page. At the top, it says 'Total Number(s) of Patients: 953'. Below this is a search bar with a 'Search' button. Underneath is an 'Advanced Search' section with a dropdown menu and a 'Search' button. The main part of the page is a table with the following columns: ID, SURNAME, OTHER NAMES, SEX, DOCTOR, DIAGNOSIS, PRESCRIPTION, STATUS, DATE ADMITTED, DATE DISCHARGED, No(s) of Days, and ACTION. The table contains two rows of data.

ID	SURNAME	OTHER NAMES	SEX	DOCTOR	DIAGNOSIS	PRESCRIPTION	STATUS	DATE ADMITTED	DATE DISCHARGED	No(s) of Days	ACTION
PAT-141030210	LUANA	ANNA	Female	Dr. KARA (STP-141030473)	Diabetes	Drugs	Discharged	3/1/2000	3/1/2000		VIEW RECORD
PAT-141030223	ANNA	ANNA	Male	Dr. KARA (STP-141030473)	Diabetes	Drugs	Discharged	3/1/2000	3/1/2000		VIEW RECORD

Figure 7: Advance Search

The screenshot shows the 'Patient's Record' page for Patient ID (PAT-141030305). The page displays the following information: NAME: OLUFEMI ABILUJAMI, STATUS: Discharged, DOB: 01/01/1980, Sex: male, Date of Birth: 01/01/1980, Phone Number: 08170341078, Address: 45th Street, FMS, ABILUJAMI@GMAIL.COM. Below this is a 'MEDICAL RECORD' section with the following details: DOCTOR INCHARGE: DR. KARA (STP-141030473), DATE OF ADMISSION: 3/1/2000, DATE OF DISCHARGE: 3/1/2000, DIAGNOSIS: Diabetes, DOCTOR'S REPORT: Patient needs follow-up care, PRESCRIPTION: Drugs.

Figure 8: View Registered Patients' Information



Figure 9: Add New Staff



Figure 10: Create New Department

The analytical decision-making ability of the developed web-based Strategic Management Hospital Information System was tested based on three selected patterns (parameters): Doctors, Ailments and Year. Doctor is a parameter that was used to mining the doctor performance based on the attendance to the patients and the doctor ability on the patients in relate to ailments and years. Doctor parameter can be used to determine the rate of each status (Death, Discharged, Transferred and On-admission) of which each Doctor have on their patient each year.

Ailments was a parameter that was used to determine the most occurrences ailments in each year and to determine the rate of each status (Death, Discharged, Transferred and On-admission) and the Doctor performances at each ailment in each year. Year was a parameter that was used determine the rate of status (Death, Discharged, Transferred and On-admission) of each ailment in each year provided or choose between

the years used in the application. Year was also used to determine and show the performance of each Doctor according to the year selected.

Figures 11, 12 and 13 depict the performances of the four Doctor used in this application. The performances of these Doctors are showing for each year, from 2017 to 2019 and the parameter used are: (i) Ailment Diagnosis (Malaria, Cholesterol, Arthritis, High Blood Pressure, Fever Bladder Infection, Ashma, Cancer, Chicken Pot, Diabetes, Heart Disease, HIV / AIDS, STD, Injury / Accident) and (ii) Treatment Status (Discharged, Transferred, Death, On Admission)

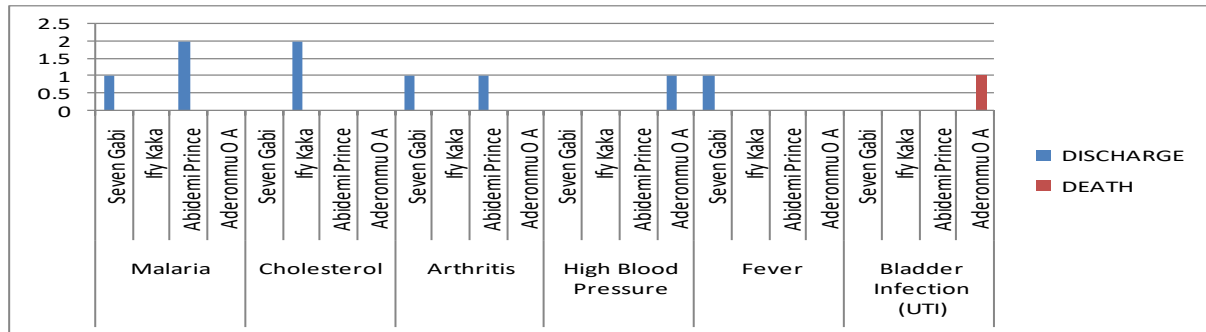


Figure 11: 2017 Doctor, Diagnosis and Status

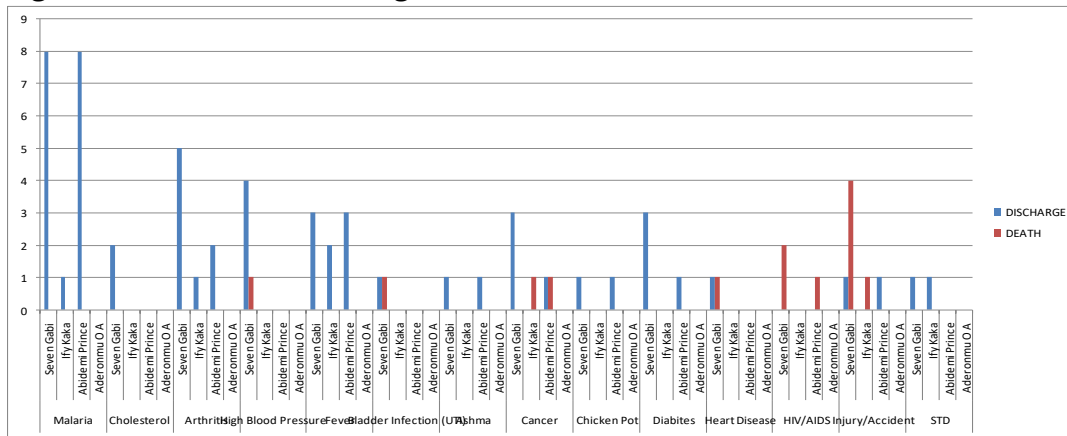


Figure 12: 2018 Doctor, Diagnosis and Status

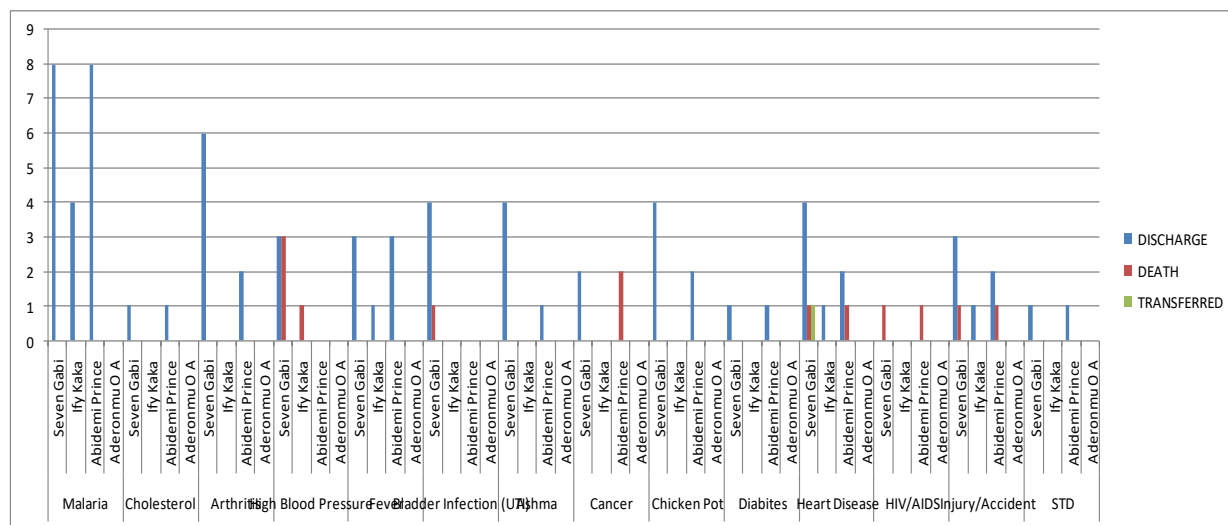


Figure 13: 2019 Doctor, Diagnosis and Status

In 2017, Figure 11 illustrates that Doctor Abidemi Prince have the high rate of Discharge Malaria Patients, Doctor Ify Kaka have high rate of Discharge Cholesterol Patients and Doctor Aderonmu O. A have Death Patient of Bladder Infection, while in 2018, Figure 12 depicts that Doctor Abidemi Prince and Doctor Steve Gabi have the high rate of Discharge Malaria and Fever Patients, Doctor Steve Gabi have high rate of Discharge Cholesterol, Arthritis, High Blood Pressure, Cancer, Diabetes Patients and have high rate of Death patients in HIV/AIDS and Injury/Accident.

In 2019, Figure 13 shows that Doctor Abidemi Prince and Doctor Steve Gabi have the high rate of Discharge Malaria and Fever Patients, Doctor Steve Gabi have high rate of Discharge Asthma, Arthritis, High Blood Pressure, Bladder Infection, Chicken Pox and Heart Disease Patients and have high rate of Death patients in High Blood Pressure and Doctor Abidemi Prince have high rate of Death Patients in Cancer. This shows that the developed web-based HIS (SHIMS) is enriched with analytical solutions to support in strategic planning and decision-making and thereby assisted the strategic management mechanisms of the HIS.

5. Conclusion and Recommendations

This paper presents the development of web-based Strategic Hospital Information Management System (SHIMS). The use of DM technology to support the data processing, information retrieval and knowledge generation process in order to support decision making in HIS offer obvious advantages over traditional HIS called legacy systems. When analytical technologies are embedded in Modern HIS that gather and centrally store large amounts of data, this will assist the strategic management mechanisms and improvements in clinical decision-making. The developed SHIMS 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 support the analytical decision-

making process.

Further exploration of data mining for research related to health information systems seems in order, especially where the data volume exceeds capabilities of traditional statistical techniques. Models produced via data mining and predictive analysis can form the backbone of future health information systems.

6. References

- Abdullah, Z. A. (2013). Hospital Information Systems Implementation Framework: Critical Success Factors for Malaysian Public Hospitals. Unpublished Ph.D Thesis, Curtin University
- 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.
- Aggelidis, V.P. and Chatzoglou, P. D. (2008). Methods for Evaluating Hospital Information Systems: A Literature Review. *EuroMed Journal of Business*, 3 (1): 99-118.
- Aktas, A. Z. (1987). Structured analysis and design of information systems. Englewood Cliffs, NJ:Prentice Hall.
- AlZoubi, W. A. (2013). Mining Medical Databases Using Graph based Association Rules. *International Journal of Machine Learning and Computing*, 3(3): 294-296.
- Anderson, L.K. and Stafford, C. J. (2002). The "Big Bang" Implementation: Not for the Faint of Heart. *Computers in Nursing*, 20 (1): 14-20.
- Armoni, A. (2002). Effective Healthcare information systems, IRM Press.
- Bawden, D. (2005). Data Mining and Decision Support: Integration and Collaboration. *Journal of Documentation*, 61(3): 443-445.
- Biohealthmatics.com. (2006). Hospital Information Systems. Accessed 24 Dec 2020, <http://www.biohealthmatics.com>.
- Bulgiba, A.M. (2004). Information Technology in Health Care – What the Future Holds. *Asia Pacific Journal of Public Health*, 16 (1): 64-71.
- Buntin, M.B., Burke, M. F., Hoaglin, M. C. and Blumenthal, D. (2011). The Benefits of Health Information 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 Systems Research. *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.
- Davis, W. S. (1983). Systems analysis and design: A structured approach. Reading, MA: Addison Wesley.
- Dutta, A, and Heda, S (2000). Information systems architecture to support managed care business process. *Decision Support Systems*, 30: 217-225.
- El-Sappagh, S. H., El-Masri, S., Riad, A. M., and Elmogy, M. (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.

- Feelders, A., Daniels, H. and Holsheimer, M. (2000). Methodological and Practical Aspects of Data Mining. *Information and Management*, 271-281.
- 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.
- Goldschmidt, P.G. (2005). HIT and MIS: Implications of Health Information Technology and Medical Information Systems. *Communications of the ACM*, 48 (10): 68-74.
- 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.
- Hampshire, D. A. and Rosborough, B. J. (1993). The evolution of decision support in a managed care organization. *Topics in Healthcare Financing*, 20(2):26-37.
- 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>.
- Jones, M.R. (2003). Computers Can Land People on Mars, Why Can't They Get Them to Work in a Hospital. *Methods Inf Med*, 42: 410-415.
- 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
- Koh, H. C. and Tan, G. (2005): Data Mining Applications in Healthcare. *Journal of Healthcare Information Management*, 9(2): 64-72.
- Krivda, C. D. (1995). Data-Mining Dynamine. Byte.
- Kuhn, K.A. and Giuse. D. A. (2001). From Hospital Information Systems to Health Information Systems. *Methods of Information in Medicine*, 40 (4): 275-287.
- 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, A.S. (2004). Thinking About Social Theory and Philosophy for Information Systems. *Social Theory and Philosophy for Information Systems*: 1-26.
- 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.
- MacIennan, J., Tang, Z. and Crivat, B. (2008). Data mining with Microsoft SQL server 2008, Indianapolis, Wiley.
- Matkovsky, I. P. and Nauta, K. R. (1998). Overview of data mining techniques. Presented at the Federal Database Colloquium and Exposition; September 9-11; San Diego, CA.
- MIT Sloan. (2012). Center for Information Systems Research - MIT Sloan School of Management. Accessed September 05, 2020, <http://cistr.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.
- Prins, S and Stegwee, R. A. (2000). 'Zorgproducten en geïntegreerde informatiesystemen', (in Dutch) *Handboek sturen met zorgproducten*, F3100-3, december.
- 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 Adem Karahoca (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.
- Silver, M.S., Markus, M. L. and Beath, C. M. (1995). The Information Technology Interaction Model: A Foundation for the MBA Core Course. *MIS Quarterly*, 19 (3): 361-390.
- Sprague, R. H. and Carlson, E. D. (1982). Building effective decision support systems. Englewood Cliffs, NJ: Prentice Hall.
- 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

- Wainwright, D. and Waring, T. 2000. The Information Management and Technology Strategy of the UK National Health Service – Determining Progress in the NHS Acute Hospital Sector. *International Journal of Public Sector Management*, 13 (3):241-259.
- Walsham, G. and Sahay, S. (2006). Research on Information Systems in Developing Countries: Current Landscape and Future Prospects. *Information Technology for Development*, 12 (1): 7-24.
- Wang, D. Y. (2010). Advanced Studying on Microsoft SQL Server 2008 Data Mining. *Applied Mechanics and Materials*, 20 - 23, 748 - 752.
- Zuckerman and Alan, M. (2006). 'Healthcare Strategic Planning', Prentice Hall of India