

Electronic Health Records: Applications, Techniques and Challenges

Abdel Nasser H. Zaied
Dean, Faculty of Computers &
Informatics, Zagazig University,
Egypt

Mohammed Elmogy
Faculty of Computers &
Information, Mansoura
University, Egypt

Seham Abd Elkader
Teaching Assistant, Faculty of
Computers & Informatics,
Zagazig University, Egypt

ABSTRACT

With the fast evolution of information technology, traditional healthcare is moving towards a more electronic stage. As a result, the e-Health term appears, and Electronic Health Records (EHRs) become a critical application of e-Health management systems. These computerized records have been widely used by clinicians, healthcare providers, patients, and health insurance companies for the purpose of creating, managing, and accessing the health information of patients everywhere. Moreover, these information resources can be shared by different healthcare parties for monitoring the patients' health, delivering effective treatments, and decreasing costs. In this paper, we present a survey of various EHR applications in e-Health systems. These applications include using EHRs for diagnosing, monitoring diseases, and selecting the most efficient paths of treatments. In addition, we discuss the usage of EHRs as a source for building a knowledge base for Clinical Decision Support Systems (CDSS). Finally, the challenges of EHR implementations in the healthcare environment and current research topics will be highlighted.

Keywords: Electronic Health Records (EHR), E-Health Systems, Clinical Decision Support Systems (CDSS), Data Mining (DM) Techniques.

1. INTRODUCTION

It is well-known that traditional health care systems face many challenges. Some challenges are that these systems include massive work done manually on papers. The increased size of the database is another problem, and therefore it's hard to update it. In addition, there is a need for massive space to store these paper files. Thus, the rapid progress of information technology changed the conventional healthcare, and electronic health systems began to be adopted on a larger scale. Individuals consider these electronic systems as a key for solving the problems of traditional systems [1].

E-Health is a recent concept that facilitates delivering information and services using the Internet technology and other technological tools [2]. The e-Health term is widely used by many academic institutions, individuals, funding organizations, and professionals. E-Health can improve the traditional services delivery and channels of communication for the aim of providing individuals with all services they need [3]. Additionally, e-Health systems encourage patients to be involved in making accurate health decisions, and also empower the communication between patients and physicians [4].

E-Health involves various systems or services that are at the edge of healthcare and technology. One of these typical applications is EHRs that enable the storage and sharing of the patient's data between healthcare professionals. EHRs are at the forefront of healthcare IT. Consequently, many physicians and healthcare providers turned towards using EHR systems [5].

EHR systems have been constructed or built in many countries, but other countries have not yet reached an agreement on EHR definition. However, EHR can be defined as a place that stores information about patients. These electronic records are prepared by the hospitals, clinics, and other healthcare institutions. They store the past health history, radiology reports, treatments, lab tests, treatments given, immunization, and other relevant data that describe the health status of patients. Thus, EHR is a combination of patients' information systems [6]. In addition, EHR involves the network that connects these databases, systems, and interfaces, order entry by physicians, clinical workstations, and electronic communications systems. EHR systems are being recommended by health institutions and professionals for improving care quality of these diseases and providing effective, safe and efficient healthcare services, and reduce expenses [7].

Due to the inconsistency, incompleteness, and the large volumes of EHRs data that are too complicated and voluminous to be processed and analyzed by traditional methods. Therefore, some machine learning algorithms and techniques are applied to these data. These techniques transform EHR raw data into new and high-quality data set that are called a case-base in Case Base Reasoning (CBR) systems. The CBR systems are the most appropriate CDSS that improve decision-making process [8].

This paper is organized into five sections as follows. Section 2 discusses the different previous studies that talk about the numerous application fields of the EHRs in the e-Health systems. Section 3 presents the challenges of implementing EHRs. In Section 4, current research topic and future trends are introduced. Finally, the conclusion and the future directions are presented in Section 5.

2. RELATED WORK

EHRs systems are developed by many healthcare institutions. The researchers asserted that using EHRs systems are very beneficial to help healthcare organizations making accurate, correct decisions and improve the quality of healthcare services. The applications of EHR systems are numerous in the healthcare environment. These application fields include

diagnosis of diseases, making effective treatments, monitoring of patient health condition, and prevention of diseases. Some studies showed that EHRs are a critical source for building CDSS based on high-quality knowledge. Other efforts indicated the benefits of applying EHR on cloud networks to be accessed and shared by all concerned parties. Finally, good efforts confirmed the importance of applying the privacy and access control mechanisms to secure sharing process of EHRs data.

2.1 EHR Adoption for Diagnosis and Prediction of Diseases

Sikder et al. [9] designed user-friendly EHR system with a new approach that is phenotypic term binding, which help physicians predict the disease and as well improve facilities of healthcare. This application allows patients to register, access EHR system to do some functionality, such as view personal data, alter data when needed, and so on. The health worker also plays a significant role in this system. On the other hand, the physician is the most important part that uses a new concept appears in this system that is phenotypic term binding that retrieves phenotype that physicians use to predict the disease. This system is a good system that utilized a new approach for diagnosing diseases. When it was evaluated on a number of institutions, good results were obtained. On the other hand, this system needs automatic disease prediction feature that will help physicians make faster decisions.

Lin et al. [10] proposed a time-to-event approach for diagnosing the chronic illnesses that use EHR as a data source. This proposed framework helps the physician to predict the chronic condition of patients, expect if there are severe complications will occur, and plan for possible interventions for reducing event risks that will occur. The proposed time-to-event modeling differs from other prediction modeling studies in that it enables an integrating meaningful set of features. Integration of data abstraction techniques exists to improve prediction accuracy and decreasing data dimensionality. The drawbacks of this system are it did not handle clinical texts, and the guideline-based feature selection includes manually massive work done which is not scalable as automatic feature selection schemes. A comparison of guideline-based method and another statistical feature selection methods is necessary.

2.1.1 Classical Data Mining (DM) Techniques for Diseases Diagnosis

Kumari and Godara [11] proposed applying some DM techniques of classification for diagnosing cardiovascular diseases. These proposed methods were RIPPER, Decision Trees (DT), Support Vector Machines (SVM), and Neural Networks (ANN). The criteria that were utilized to compare the performance of these techniques are specificity, sensitivity, error rate, accuracy, false positive rate, and true positive rate. This study worked on 14 attributes and 296 dataset instances. These prediction models were estimated using 10-fold cross validation. When these techniques were analyzed, results indicated that SVM could predict cardiovascular diseases with the lower error rate and had the best performance.

Dangare and Apte [12] presented DM classification techniques which are ANN, Naive Bayes, and DT for predicting heart disease. In addition, they proposed a prediction system of heart diseases using backpropagation algorithm together with Multilayer Perception Neural Network (MLPNN). They employed 573 dataset instances and

15 attributes. The results were superlative and were nearly 100% of accuracy.

Guru et al. [13] developed an application for detecting heart diseases, sugar, and blood pressure by exploiting ANN. There are 78 records and 13 attributes; those represent the dataset used in this study for applying the training and test experiments. This study suggested diagnosing heart diseases by the supervised network and using backpropagation algorithm for training it. The system had the ability to identify the unknown data entered by physicians from training data. Then it could produce a list of diseases that patients may have.

Krishniah et al. [14] proposed using K-Mean, K-Nearest Neighbor (K-NN), and Entropy-based Mean Clustering (EMC) algorithms to anticipate heart attacks. They utilized some 14 features and 3000 instances. The result of this study was that EMC had the highest accuracy.

Palaniappan and Awang [15] proposed using Naïve Bayes and ANN to develop an Intelligent Heart Disease Prediction System (IHDP) prototype. ANN achieved 85.53% of accuracy while Naïve Bayes has 86.53%.

Sundar et al. [16] tried to design a system used for the diagnosis of heart diseases through utilizing two DM methods like Naïve Bayes and Weighted Associative Classifier (WAC). This system was evaluated on 15 attributes and a large number of instances about 2268 and concluded that WAC technique had higher results than Naïve Bayes in heart disease prediction.

Shahbaz et al. [17] used Naïve Bayesian, SVM, and KNN algorithms to build a DSS that can identify different types of cancer on the Genes dataset. The precision results were KNN had 90.72%; SVM had 90.27% while Naïve Bayes had the highest results in accuracy which about 95%.

Prasad et al. [18] conducted a comparative study of Auto-Associative Memory Neural Networks, ID3 (Iterative Dichotomized3), C4.5, and Bayesian networks algorithms. The comparison aims at identifying the best technique for diagnosing Asthma diseases.

Rajesh and Sangeetha [19] implemented the C4.5 technique. In addition, Kumar et al. [20] proposed using an ID3 algorithm, C4.5 algorithm, and Classification and Regression Trees (CART) algorithm for classifying hepatitis, diabetes, and heart diseases. The research focused on medical diagnosis by knowing manner through the gathered data related to these diseases and build intelligent medical systems for supporting decisions. It was concluded that the performance of CART DT was better than C4.5 and ID3.

Khan et al. [21] applied two DT algorithms namely C4.5 and C5.0 to develop a prototype of heart and cancer disease prediction model. Results indicated that C5.0 algorithm generates rules in a very simplified and readable form. While C4.5 algorithm produces the rule set in a decision tree form.

Wang et al. [22] utilized logistic regression and DT methods of data mining to check 5-year survivability of patients with breast cancer. Some parameters were adopted to measure the methods' performance like accuracy, specificity, sensitivity, and g-mean. The results showed that there is no remarkable difference in accuracy between the two methods. Whereas g-mean values and area under the curve showed that logistic regression had better performance than the decision tree model in the prediction of breast cancer survivability.

2.1.2 DM Techniques for Diseases Prediction

Chao and Wong [23] suggested complementing traditional incremental decision tree learning algorithms by applying the I+ Learning theory that is about incremental, intelligent, and interactive learning. This theory is concerned with newly available attributes besides the new incoming instances. This study also proposed a model of practical DM diagnostic workbench that supports real medical diagnosis using technologies of DM and multi-agent [24].

McCormick et al. [25] implemented a statistical modeling mechanism, named the Hierarchical Association Rule Model (HARM). This statistical model predicts the possible symptoms of a patient that may occur in the future through the available past and current history of reported symptoms.

Chandra et al. [26] suggested detection of association rules on medical data using a new method for the aim of predicting heart disease using Matrix Based technique to decrease the number searching a database.

Ullah [27] applied Binary Association Rules to detect relationships that exist among items in medical databases. The researcher used binarization to convert numerical data to binary (0,1) ones, after that he implemented Apriori algorithm and Swarm Intelligence Algorithm for the purpose of practical implementation.

Hesar [28] focused on employing Probabilistic Bayesian Classifier (BPCs) for diagnosis and prognosis of breast cancer disease on Wisconsin Diagnostic Breast Cancer dataset. BPCs are DM tool that encodes relationships among a collection of variables under uncertainty. In structure or building phase, various heuristic algorithms are evaluated, where Tabu is selected for graphical modeling of proposed BPC and Genie simulator (using Maximize Likelihood Estimation) is used for numerical modeling of BPN.

Patil et al. [29] applied D-stream and K-means classification algorithms to implement a system that predicts the patient's current health condition based on the past and real time data of a patient. The results of their system showed that D-stream algorithm improved the performance and its results are more accurate than the other tested algorithm, and therefore it solves the problems of K-means algorithm.

Shukla et al. [30] applied computing tools such as Neuro-Fuzzy Systems and ANN for building a knowledge system that can diagnose breast cancer diseases.

Jimenez et al. [31] designed an expert system to diagnose problems of health care. The knowledge base of expert systems is about the collection of production rules and possible facts. These rules are written in classic logic that has two values. The inference engine of the expert system employs Gröbner Bases and normal form to get the diagnosis from the information kept in the knowledge base.

2.1.3 Hybrid DM Techniques for Diseases Prediction

Tsipouras et al. [32] applied a fuzzy model and C4.5 DT to develop a DSS based on the fuzzy rule-based architecture for diagnosing coronary artery disease. The results showed that the proposed DSS based on optimized fuzzy model improves the outcome of the crisp rule-based classifier by 15% in terms of accuracy.

Bhatla and Jyoti [33] presented two classifier algorithms namely, Naïve Bayes and DT Classifiers using Fuzzy Logic for predicting the existence of heart disease with few attributes. In this study, there are six characteristics were included in predicting the heart disease, then they reduced

these characteristics to four attributes and then automatically reduce the number of tests to be taken by a patient. It was concluded that this proposed system could predict with 100 % efficiency.

Ha and Joo [34] suggested hybrid algorithms for classifying chest pain. The two hybrid algorithms were Apriori and C5.0 DT algorithms. Apriori algorithm was used to detect lab tests. C5.0 DT produced the classification rule base for classifying chest pain. They aimed to allow physicians take decisions quicker and more precisely.

Kavitha et al. [35] prepared an ANN model for diagnosing heart illnesses. This proposed research applied hybrid approach using a Genetic algorithm to train the neural network.

While Soni et al. [36] conducted a study of Bayesian classification, DT, classification based on clustering, KNN Algorithm, and ANN for a heart diseases diagnosis. The purpose of applying Genetic Algorithm is to improve the prediction accuracy of both Bayesian classification and DT.

Shouman et al. [37] examined the integration of k-means clustering with naïve bayes for diagnosing patients with heart disease. They confirmed that when Naïve Bayes and K-means clustering were integrated together, this improves the accuracy of Naïve Bayes in diagnosing heart diseases.

Shouman et al. [38] discussed using KNN and integrate voting with KNN for a heart diseases diagnosis. The results proved that using KNN techniques could improve the performance, and therefore 97.4% of accuracy was obtained. While, voting could not improve the precision in heart diseases diagnosis.

Patil and Kumaraswamy [39] designed an efficient and powerful system for detecting heart attack by operating DM and MLPNN. First, significant patterns were discovered from a data warehouse of heart diseases. Then K-means clustering was used to cluster the preprocessed data warehouse and MAFLIA algorithm successfully extracted the frequent items. There is a weight age calculated for frequent patterns. This weight age helped in identifying when the items had a value higher than a predefined threshold. Therefore, these patterns could be selected for the diagnosis of heart diseases. Finally, the chosen patterns were adopted in training the MLPNN algorithm using Backpropagation training algorithm.

The following studies done on cardiovascular diseases [40] [41]; which can be better managed by analyzing health records data [42]. Sofia et al. [43] developed an EHR system for cardiovascular medicine. The purpose of this study is to list, analyze and evaluate the currently certified EHRs by certification body named certification commission for healthcare information technology (CCHIT). CCHIT put testing criteria for EHRs to test their functionality, interoperability, and security issues. Each EHR chosen was assessed using quality assessment process (QA). The results indicated that 22 CCHIT-certified EHRs from 77 EHRs were selected. The quality evaluation process may include some criteria that make a group of EHRs has an advantage than others; Secondly, adding more criteria can produce a different result in quality assessment classification process. A future research requires analyzing the specification to recognize the best EHR for cardiovascular. In addition, security and privacy requirements are also needed to be addressed.

2.2 Diseases Treatments

Aljumah et al. [44] proposed a regression analysis algorithm to detect the most efficient treatment for hypertension. Yeh et al. [45] developed a predictive model for improving the diagnosis and treatments accuracy of cerebrovascular disease. Three algorithms for classification were implemented for building the classification models; these are Backpropagation Neural Network, Bayesian classifier, and DT.

Duan et al. [46] suggested a clinical recommendation system. This proposed application applied a prefix tree structure known in the item set mining for building a ranked list of care plan items suggested that depend on entered items previously. There are two evaluation techniques were utilized, namely the greedy selection and random selection.

Sakthimurugan and Poonkuzhali [47] represented a framework with the use of KNN classifier for classifying the semantic relations between treatment and disease.

Badria et al. [48] proposed a rough set technique based on the granular computing framework. The proposed technique is used for selecting the most relevant features of the dataset and produces a minimal set of decision rules. These decision rules classify the impact of a new treatment given for patients with Hepatitis C Virus (HCV). The overall accuracy of classification of the proposed approach depends highly on the rules generated during the treatment of HCV. Rough set approach with granular computing framework achieves good classification accuracy, but the limitations were processing time of data reduction.

Eissa et al. [49] proposed a hybrid model that takes advantage of both rough set and genetic algorithms. This hybrid model applied the rough set for identifying the most relevant attributes and generates a minimal set of attributes called reducts and produces a group of rules that support taking decisions. Genetic algorithms were also used for optimizing generated rules by the use of rough sets that perform classification of the new medication impacts for HCV. The overall accuracy of that hybrid model is superlative, rough sets and Genetic algorithms together can produce a positive effect, enhancing model performance. In addition, the processing time of data reduction is a limitation. Thus, it is suggested to apply new hybrid approaches with larger datasets and compared them with the proposed approach and other techniques.

Lo et al. [50] proposed an algorithm to facilitate detecting adverse drug reaction (ADR) using EHRs. This detection was done by using detection methods. These detection methods extract and mine EHRs data. They were monitored by spontaneous reporting systems (SRS) developed by known organizations. The Poisson method was proposed to make a comparison of condition occurrence rates, such as disproportionality methods. The exploratory results indicated that the proposed method has better performance than other ADR detection methods. However, it was discovered that Poisson method detects only single cause-effects relationships. In the future, more extension needed for detecting required causes and effects associations in other longitudinal data types. Furthermore, dealing with multiple causes/effects and select observation methods automatically.

Cintho et al. [51] proposed an EHR system for kidney disease prevention through the use of archetypes. These archetypes allow representing information and rules included in clinical practice guidelines (CPGs). The CPGs are integrated into EHR for chronic kidney disease prevention. In this system,

data and rules contained in guidelines are identified in the beginning. Archetype is prepared, and rules in inference engine were determined and collected to build the DSS. Then, DSS and EHR are integrated together. Finally, the validity of DSS is checked. Archetypes, especially for disease prevention area, are a new concept appear in recent years and has an advantage than traditional systems. However, it is not easy to make the CPGs as a formal way for prevention. In addition, selecting the type of structure is a problem in this study. Therefore, more work is needed on these issues and also validation phase is required to complete the work.

2.3 EHR for Tracking and Monitoring Diseases

Navaneethan et al. [52] proposed an EHR for monitoring chronic kidney illness. They used data from EHRs to develop chronic kidney disease registry from social security and kidney data system. This registry permit is conducting EHRs-based clinical trials that seek whether providing patients with knowledge about their personal records improves chronic kidney care. The results were that better collection of data concerning patients will facilitate providing better care for patients, simplify clinical research. Data associated with dialysis and dissolution might not be immediately available. In addition, the value of the personal health records (PHRs) depends on patient's interest of their health condition [53]. Thus, dependence on health records should involve patient awareness and education about provider education to improve EHRs efficient use.

Khan [54] built a PHR System for the purpose of managing the health status of chronic patients. It allowed data to be stored securely. This application was innovative and flexible. It used a preventative approach that involves the patient, monitor multiple chronic conditions over a long-term period and provide timely feedback on patient condition. This system improved the diagnosis efficiency, deliver high-quality healthcare services. However, it was restricted by resources available and the requirements, existing infrastructure, limited time, and limited funding. Therefore, it was needed to modify thoroughly and develop the application functionalities, perform validity and usability test for efficiency and bugs of application algorithm; it also required to expand the system later.

Gupta et al. [55] compared Genetic Algorithms, Rough Sets, Fuzzy Sets, SVM, ANN, and DT on tracking patients who suffer from breast cancer diseases. Results proved that all methods achieved a higher percentage of accuracy. The DT algorithm had the highest percentage that was 100% on breast cancer monitoring.

Delen [56] suggested DT, Logistic Regression, SVM, and ANN methods for analyzing data of prostate cancer. Results indicated that when these methods were compared, SVM was the best performing algorithm with 92.85% of accuracy higher than other methods.

Osofisan et al. [57] applied ANN method for kidney failure prognostic. The researchers suggested using two learning algorithms. These were Backpropagation Delta-Delta Learning Algorithm and Backpropagation. The Pattern Mode algorithm extracts features that could expect a patient life, the survival of kidney dialysis and discover kidney failure existence.

Jacob and Ramani [58] proposed an efficient classifier for Prognostic of Breast Cancer in Wisconsin (WPBC) data set from the repository of UCI machine learning. Twenty

algorithms of classification were applied; they concluded that C4.5 classification algorithm was the best algorithm in performance with 100% accuracy.

2.4 EHRs Integration to Support Physician Collaboration

Noteboom et al. [59] proposed the meaningful integration of EHRs for facilitating collaboration among care providers. It improves the care provided to patients, minimizes the increased costs of healthcare, and achieves service continuity. The qualitative research method was applied to ask physicians about their perception and experience of EHRs. The results of Open Coding analysis process showed that the majority of physicians who use EHRs for collaboration were greatly negative. Thus, they proposed a model that integrates health records together with clinical processes by using EHRs with Web services allowed patients, health care providers' access necessary information and knowledge for improving care. In future research, it is necessary to evaluate patient-centered care outcomes by applying technological products that use EHR data.

Strauss et al. [60] conducted a study on effects of collaboration and notified physicians to apply EHRs for coordination, adaptation, collaboration, and communication. In addition, this research notified the adaptation of EHR at work, technology, and social areas [61] [62] [63].

Qureshi et al. [64] presented a model that describe the effects of E-collaboration. It assured that when technology used by persons for working with each other, there will be improvements in patient-centered care. New technology adaptation in interactive groups collaborated with each other occurs if the members know how a new technology impacts their professional work and relationships [65] [66].

2.5 EHR Application of CBR Systems

El-Sappagh et al. [67] proposed a framework for preparing a case-base for CBR systems. CBR systems are considered the most suitable systems for making more efficient decisions [68] [69]. EHRs are used as a starting point or a complete source for building that case-base. Three major phases were applied on EHR data; these are data preprocessing, data fuzzification and coding stage. This study concentrates only on data preprocessing phase that includes many steps like feature selection and normalization [70]. A number of machine learning algorithms are used to apply these preprocessing steps using the diagnosis dataset of diabetes diseases. When testing this application, it is noted that preparing EHR data improves the accuracy of CBR systems retrieval phase. The future study includes completing the remaining phases of case-base preparation using SNOMED CT, UMLS OR ICD ontologies, also increasing case-base size to enhance CBR system decision accuracy.

2.5.1 Preprocessing EHR Data

Many efforts have been done to preprocess EHR data as data quality needs to be measured as a knowledge source. Among these studies, Weiskopf and Weng [71] stated that five factors can be used for measuring EHR data quality. Xie et al. [72] handled the unmatched feature and missing values problem in case retrieval algorithms. Guessouma et al. [73] proposed five techniques for dealing with the problem of missing data in CBR system.

Data preprocessing steps include data transformation, handling missing values, data cleaning, and feature selection steps that can be applied in any order [74]. These data pre-

processing steps are applied on EHR for converting database structure of EHR to case base structure and transforming EHR generic data to specified case base.

Han et al. [75] have suggested rapid miner for processing of diabetes management dataset. Pla et al. [76] proposed an eXit*CBR system that include necessary preprocessing steps such as normalization, discretize feature selection techniques. Therefore, the list of preprocessing steps is according to the data's nature and purpose of CBR system [77]. There are so many techniques for handling missing values such as K-NN Algorithm [78]. In addition, coding process can be done using Rapid Miner Studio 6.0 by applying discretize operator. The discretize operator will perform discretization of selected attributes into specified classes through a simple mapping process.

The generated features of patients are huge in most cases, and they have different importance levels. Consequently, algorithms for feature selection must be applied to the whole range of features to determine and select the most important ones. Two attractive and complementary methods are applied for selecting the most important features, namely filter and wrapper methods [79]. There is also some machine learning algorithms can be used, such as Naïve Bayes, K-NN classifier, C4.5, and DT. The collected features from these algorithms will participate in CBR system case representation. The Rough set approach is also considered a powerful analysis tool for generating the most significant and relevant features from the whole set of features.

Weights vectors are assigned to attributes. Consequently, essential characteristics will have higher weights. There is a variety of algorithms calculate feature weightings, such as Genetic algorithms, rule induction, DT, and correlation techniques.

It is well known that the performance of algorithms especially retrieval algorithms in case base reasoning will be affected adversely if there are an outlier and extreme values in the dataset. Outliers affect the normalization process. Rapid Miner detection outlier operator checks for outliers. This operator identifies the n outliers in the dataset on the basis of distance to K -nearest neighbors. Thus, extreme values of feature can be replaced by domain experts.

2.6 Standardization of EHR

In the cloud computing environment, cloud-based EHR systems have several advantages over client-server or traditional EHR systems. However, data interoperability standards were considered a critical issue for exchanging healthcare data between all stakeholders in the healthcare system.

Bahga and Madiseti [80] implemented an EHR system on cloud networks. The proposed application achieved semantic interoperability through two-level approach; these are reference model, and the archetype model. The proposed application used the data integration engine to aggregate data from remote data sources and supports advanced security features. The cloud-based EHR systems are preferable than traditional client-server EHR systems because these systems have better interoperability, maintainability, scalability, accessibility, portability, and reduced costs. Future work requires a framework for cloud information integration for healthcare applications. In addition, it is necessary to improve the efficiency and accuracy of querying through the development of progressive sampling and data thinning approach.

LH et al. [81] developed a personal health records and health knowledge-sharing system. The system aimed to increase efficient gathering, comparing, and exchanging of healthcare knowledge. IHE Cross-Enterprise Document Sharing (XDS) is an architecture that was adopted to facilitate the process of sharing PHRs. This work suggested using registry ontology for supporting the health knowledge combination and sharing functions. This application is put on the schedule to be tested on a larger level later.

Bargen et al. [82] presented a personal intelligent system that achieves the international standards. This system was mainly designed for patients suffering from serious heart failure. One of the most commonly accepted standards for representation of information is HL7 that relies on an open-source service-oriented architecture approach and a system for controlling the processes of business. The proposed system is provided with sensors that provide a decision support system with data, continuously pay attention to several vital signs and send the data to help the communication with a healthcare provider and patient. This proposed system is a standards-based, flexible, and open system that facilitates personalized decision support. Additional development needed for implementation of components on sensor nodes and small devices.

2.6.1 Securing EHR Exchange on Clouds

In the recent healthcare environment, healthcare practitioners would prefer shifting EHR towards clouds. One of the biggest problems of cloud computing for adoption of EHR systems is the security challenges related to identity management, authentication, trust management, access control, and so on. Various studies have been conducted for addressing these security issues.

Wu et al. [83] presented a broker-based access control approach in order to support selective sharing of composite EHRs systems. The distributed instances of EHR can be combined from different fields of healthcare in the cloud network. Selective parts of composite EHRs can be shared with authorized providers by placing rules and policies of access control specified by patients. A prototype of the system being implemented which is proof-of-concept EHR sharing system. Results proved the efficiency and feasibility of that approach. The implemented system needs more extensive evaluation with health care data available in the real-world. It also needs to deal with policy composition issues [84] [85]. In addition, how to handle fine-grained delegation mechanism for EHR sharing in a cloud environment. Finally, they want to extend their healthcare ecosystem to apply their approach for EHR by using tablets and smartphones consumer devices.

Jin et al. [86] [87] presented a unified approach to patient-centric access control. Their approach supported selective composite EHRs sharing by using various levels of granularity. It accommodated collecting data and requirements that are necessary to protect privacy. The main difference between this approach and the previous one is that this approach supposes all healthcare providers. However, it is not used in cloud computing environment.

Zhang et al. [88] suggested a group of security requirements that is needed for e-Health cloud applications. It presented a reference model of EHR security to support EHRs sharing.

Jafari et al. [89] suggested a Digital Right Management (DRM) mechanism that was patient-centric for protecting the privacy of EHRs that were stored in the cloud on the basis of patient needs. These approaches are not fine-grained and not able to accommodate requirements of selective EHR sharing.

kukhun et al. [90] investigated an approach of mobile querying of distributed XML DBs in a pervasive health system. This approach is not based on cloud and does not support the requirements of EHR integration from various healthcare providers.

Li et al. [91] proposed a framework that handled the access control issues. This framework can achieve privacy of personal health records in the cloud environment by applying Attributes Based Encryption (ABE) techniques. This approach focuses on ensuring that EHRs had been shared by a selective group of users.

Durate et al. [92] implemented an extension application to EHRs called Stand-Alone EHR. This application ensured that critical information will be permanently available to achieve the quality related to treatments of patients when there is a failure of information systems. The project applied an Agency for Integration, Diffusion and Archive of Medical Information (AIDA) platform. It indicated that there was a combination and interoperability among all the hospital subsystems (HIS). Results showed that the stand-alone module had more advantages than traditional backup systems. It also achieved improvement in the EHR confidence level by health professionals. Their next step is to perform extension of this work for all services of the hospital and other health institutions; where AIDA platform is installed.

3. EHR CHALLENGES

There are many challenges of using EHRs that may exist. Therefore, there are many barriers that obstacle the development and adoption of EHRs. These challenges include financial, ethical, technical, social, and legal problems.

3.1 Financial Support

The availability of capital or financial support given by the government is considered the most critical factor that affects the EHRs development success [93].

3.2 Difficulty of Using the Technology

It is known that EHR system has many benefits; so many countries stimulated to implement it. However, the rate of using and adopting EHRs is still low to some degree, and this is because the shortage of experience and knowledge in using technology. Therefore, the complexity of using EHRs [94] is considered a barrier to implementing EHRs.

3.3 The Patients and Physicians Rejection

The third factor that cause patients and physicians reject to adopt EHRs is that health data related to patients may be penetrated and may not secure. Protection of EHRs [95] is an important problem that faces many countries that may result in lacking the confidence in using EHRs. Another reason for resistance is that there are many areas that are still using paperwork, and it was difficult for them to use and adapt the new technology.

3.4 Permanent Training

It was suggested to conduct periodical training for specific user groups like physicians to teach them how to deal with an error happened or incorrect action that the user may perform when using the system [96].

3.5 Expensive Costs for EHRs Installation

EHRs installation is regarded an expensive project. The high costs come from the installation of software and hardware devices like PC's and printers. Therefore, managers have to

consider the costs of upgrading equipment, maintaining software, or for keeping the professional staff [97].

3.6 Data Interoperability Challenge

Many healthcare providers use the health information of patients. Information is stored in health information systems in different formats, such as structured document-based storage, unstructured document storage, and relational DB tables. These different formats cause a technical problem that is data interoperability [98] in EHR installation.

3.7 System's Downtime Problem

System's downtime is a serious problem that not encourages the adoption of EHRs. This problem occurs especially in the emergency room. Thus, sufficient procedures are needed to be available to help users deal with the interruption that occur in the connectivity of the system [99].

3.8 Complicating Processes

Some users viewed that the health information system might have some parts that complicate the processes to be done. This reason makes them away from the application of EHRs.

3.9 The Problem of Secure Data Communication

Many providers in health institutions do not have enough reliable network infrastructures to secure data [100]. Therefore, securing the data communication is regarded a technical problem we need to deal with for successful use of EHRs.

3.10 Necessity of Adherence Study

The most critical issue we need to consider is that applying EHR is not sufficient to improve the quality of care. These records need to be used with systems such as a decision support system to help make more effective decisions [101].

On the other hand, there are some studies mentioned the challenges of implementing EHRs. For example, Castillo et al. [102] passed 68 articles that are published between January 1985 and December 2009. They listed some challenges of EHR adoption, such as a change of workflow, the technology features, the behavior of users, the available support and capabilities, and capacities for communication.

Devine et al. [103] agreed with Castillo et al.. They also mentioned that interests of users, the interaction of persons with change, previous knowledge, the skills of using a computer, and technical experience may be the challenges that prevent the adoption of EHRs in organizations.

4. CURRENT RESEARCH TOPICS & FUTURE TRENDS

In this section, we will take a closer look at the previously discussed studies. We will focus on identifying the limitation points in each study and the future research.

- Some researchers mentioned that there was a system for diagnosis of diseases using a new approach called phenotype term binding. This proposed system needs further research on automatic disease prediction that helps making decisions quicker [9].
- The system being designed for diseases diagnosis using time-to-event framework has some limitations. It did not handle clinical texts, and feature selection has more work

done manually. As well, it is not scalable as automatic schemes of feature selection. Finally, comparing guideline-based method and other statistical methods for selecting features is necessary [10].

- A previous study proposed applying EHRs for cardiovascular medicine. Its purpose is to list, analyze and evaluate the currently certified EHRs by certification body named CCHIT. It put testing criteria for EHRs to test their functionality, interoperability, and security issues. There are two points being observed in this study. The first point is that the quality assessment process may include some criteria that make a group of EHRs has an advantage than others. Secondly, adding more criteria can produce a different result in quality assessment classification process. A future research requires analyzing the specification to recognize the best EHR for cardiovascular. In addition, security and privacy requirements also need to be addressed [43].
- Another research proposed using a rough set technique based on granular computing framework for classifying the effect of new drugs for Hepatitis C Virus (HCV) treatment. The limitation here was the processing time of data reduction. In the future, new hybrid approaches with larger datasets are required to be compared with the proposed approach and other techniques for reaching higher accuracy as possible [48].
- There is an effort done by researchers who proposed the previous study. They also proposed a hybrid rough genetic algorithm model. They utilized the rough set advantages and genetic algorithms for classifying the effects of new medication for patients with HCV. The time taken for processing of data reduction is also a limitation. Thus, it is suggested to apply new hybrid approaches with larger datasets and compared them with the proposed approach and other techniques [49].
- There is a system that proposed a method for early detection of Adverse Drug Reactions (ADR). The proposed technique named Poisson method. The detection here was through mining EHRs. It was discovered that poison method applicable here detected only single cause-effects relationships. The future directions are that proposed method needs more extension for detecting required causes and effects associations in other longitudinal data types and deal with multiple causes and effects. Finally, it is required to select observation methods automatically [50].
- The EHR system implemented for kidney disease prevention proposed using archetypes. These archetypes allow representation of information and rules included in clinical practice guidelines (CPGs). CPGs are integrated into EHR for chronic kidney disease prevention. The limitation here is that it is not easy to formalize CPGs prevention and choosing structure type is a problem. Therefore, more work needed on these issues and validation phase required to complete that work [51].
- There is a proposed EHR system for monitoring chronic kidney illness. The EHR researchers used data from EHRs to develop chronic kidney disease registry from social security and kidney data system. However, data associated with dialysis and dissolution might not be immediately available. In addition, the value of the

personal health records (PHRs) depends on patient's interest of their health condition. Thus, dependence on health records should involve patient awareness and education about provider education to improve EHRs efficient use [52].

- An effort proposed building personalized systems to track the health of chronic patients. These systems took into consideration more than one of chronic conditions. However, this work is restricted by resources available, existing infrastructure, restricted time, and funding. Therefore, it is necessary to modify the application functionalities completely. As well, performing validity and usability test for efficiency and bugs of the application algorithm. It also needs to be applied on a larger scale [54].
- Other researchers proposed a meaningful integration of EHRs for facilitating collaboration among care providers to improve the care provided to patients. It minimizes the increased costs of healthcare and achieves service continuity. In future research, it is necessary to evaluate patient-centered care outcomes by applying technological products based on EHR data [59].
- A cloud-based EHR system was developed to achieve semantic interoperability through the two-level approach. These are reference model, and the archetype model. A framework for cloud information integration is required later. In addition, improving the efficiency and accuracy of querying through the development of progressive sampling and data thinning approach is also needed [80].
- A personal intelligent system was implemented to fulfill the international standards. This system was mainly designed for patients suffering from serious heart failure. It is a standards-based, flexible, and open system that facilitates personalized decision support. Additional development is needed for implementation of components on sensor nodes and small devices [82].
- There is a research that concentrated on access control issues of EMRs in clouds. It presented a broker-based access control approach for selective sharing problem of composite EHRs systems. Distributed sources of EHRs can be collected from different health care domains in the cloud environment. The implemented system needs more extensive evaluation with health care data available in the real-world. We need also to deal with policy composition issues. Handling fine-grained delegation mechanism for EHR sharing in a cloud environment is also a critical point. Finally, we want to extend our healthcare ecosystem to apply our approach for EHR by using tablets and smartphones consumer devices [83].
- Some researchers discussed an extended application to EHRs. This application guarantees that critical information will be permanently available to achieve the quality related to treatments of patients when there is a failure of information systems. The types of failures that may occur are crashes of database management, unavailability of system application servers, no electric power, and when there is no network. The next step is to extend this work to all services of the hospital and other health institutions; where AIDA platform is installed [92].

5. CONCLUSION

It is well known that technological development has led to the improvement of health services provided to patients and consequently digital systems appeared. Among modern applications of these digital systems are the EHRs that are used in many medical application areas. This paper highlighted the widespread applications of EHR in predicting and anticipating diseases, determining the most suitable medications for diseases, and monitoring the patient's health in the future. In addition, EHR plays an initial step in building a knowledge source for CDSS to make more accurate decisions. We also mentioned the necessity of sharing EHR data with different stakeholders to improve the care delivery and reduce costs related to health care services. Adequate policies are critical to assure EHR data security and privacy features as they represent key factors to EHR success. Lastly, we presented the barriers of EHR implementation, and the current research issues were also discussed.

6. REFERENCES

- [1] E. Docteur and H. Oxley, "Health-Care Systems: Lessons from the Reform Experience," *OECD Economics Department Working Papers*, no. 374, Dec. 2003.
- [2] G. Eysenbach, "What is e-health?," *Journal of medical Internet research*, vol. 3, no. 2, 2001.
- [3] K. M. Nazi, "The journey to e-health: VA healthcare network upstate New York (VISN 2)," *Journal of Medical Systems*, vol. 27, no. 1, pp. 35–45, 2003.
- [4] C. Pagliari, D. Sloan, P. Gregor, F. Sullivan, D. Detmer, J. P. Kahan, W. Oortwijn, and S. MacGillivray, "What is eHealth (4): a scoping exercise to map the field," *Journal of medical Internet research*, vol. 7, no. 1, 2005.
- [5] L. Yun and L. Xiang-sheng, "The data mining and knowledge discovery in biomedicine," in *Computer Science and Education (ICCSE), 2010 5th International Conference on*, 2010, pp. 1050–1052.
- [6] A. Begoyan, "An overview of interoperability standards for electronic health records," *USA: society for design and process science*, 2007.
- [7] M. N. Walsh, N. M. Albert, A. B. Curtis, M. Gheorghide, J. T. Heywood, Y. Liu, M. R. Mehra, C. M. O'Connor, D. Reynolds, C. W. Yancy, and others, "Lack of Association Between Electronic Health Record Systems and Improvement in Use of Evidence-Based Heart Failure Therapies in Outpatient Cardiology Practices," *Clinical cardiology*, vol. 35, no. 3, pp. 187–196, 2012.
- [8] S. Biafore, "Predictive solutions bring more power to decision makers.," *Health Management Technology*, vol. 20, no. 10, p. 12, 1999.
- [9] M. K. A. Sikder, A. N. Chy, and M. H. Seddiqui, "Electronic health record system for human disease prediction and healthcare improvement in Bangladesh," in *Informatics, Electronics & Vision (ICIEV), 2013 International Conference on*, 2013, pp. 1–5.
- [10] Y. Lin, H. Chen, R. Brown, S. Li, and H. Yang, "Time-to-Event Predictive Modeling for Chronic Conditions using Electronic Health Records," 2014.

- [11] M. Kumari and S. Godara, "Comparative Study of Data Mining Classification Methods in Cardiovascular Disease Prediction 1," 2011.
- [12] C. S. Dangare and S. S. Apte, "Improved study of heart disease prediction system using data mining classification techniques," *International Journal of Computer Applications*, vol. 47, no. 10, pp. 44–48, 2012.
- [13] N. Guru, A. Dahiya, and N. Rajpal, "Decision support system for heart disease diagnosis using neural network," *Delhi Business Review*, vol. 8, no. 1, pp. 99–101, 2007.
- [14] V. J. R. Krishniah, D. C. Sekar, and D. K. R. H. Rao, "Predicting the heart attack symptoms using biomedical data mining techniques," *The International Journal of Computer Science & Applications (TIJCSA) ISSN-2278-1080*, vol. 1, no. 3, pp. 10–18, 2012.
- [15] S. Palaniappan and R. Awang, "Intelligent heart disease prediction system using data mining techniques," in *Computer Systems and Applications, 2008. AICCSA 2008. IEEE/ACS International Conference on*, 2008, pp. 108–115.
- [16] N. A. Sundar, P. P. Latha, and M. R. Chandra, "Performance analysis of classification data mining techniques over heart disease database," *IJESAT International Journal of engineering science & advanced technology ISSN*, pp. 2250–3676, 2012.
- [17] M. Shahbaz, S. Faruq, M. Shaheen, and S. A. Masood, "Cancer Diagnosis Using Data Mining Technology," *Life Science Journal*, vol. 9, no. 1, pp. 308–313, 2012.
- [18] B. Prasad, P. K. Prasad, and Y. Sagar, "A Comparative Study of Machine Learning Algorithms as Expert Systems in Medical Diagnosis (Asthma)," in *Advances in computer science and information technology*, Springer, 2011, pp. 570–576.
- [19] K. Rajesh and V. Sangeetha, "Application of data mining methods and techniques for diabetes diagnosis," *International Journal of Engineering and Innovative Technology (IJEIT) Volume*, vol. 2, 2012.
- [20] D. S. Kumar, G. Sathyadevi, and S. Sivanesh, "Decision support system for medical diagnosis using data mining," *International Journal of Computer Science Issues*, vol. 8, no. 3, pp. 1694–814, 2011.
- [21] M. T. Khan, D. S. Qamar, and L. F. Massin, "A prototype of cancer/heart disease prediction model using data mining," *International Journal of Applied Engineering Research*, vol. 7, no. 11, pp. 1–6, 2012.
- [22] K.-M. Wang, B. Makond, W.-L. Wu, K. Wang, and Y. Lin, "Optimal data mining method for predicting breast cancer survivability," *International Journal of Innovative Management, Information & Production*, vol. 3, no. 2, pp. 28–33, 2012.
- [23] S. Chao and F. Wong, "An incremental decision tree learning methodology regarding attributes in medical data mining," in *Machine Learning and Cybernetics, 2009 International Conference on*, 2009, vol. 3, pp. 1694–1699.
- [24] S. Chao and F. Wong, "A Multi-Agent Learning Paradigm for Medical Data Mining Diagnostic Workbench," in *Data Mining and Multi-agent Integration*, Springer, 2009, pp. 177–186.
- [25] T. McCormick, C. Rudin, and D. Madigan, "A hierarchical model for association rule mining of sequential events: An approach to automated medical symptom prediction," 2011.
- [26] M. A. Jabbar, B. L. Deekshatulu, and P. Chandra, "Knowledge Discovery Using Associative Classification for Heart Disease Prediction," in *Intelligent Informatics*, vol. 182, A. Abraham and S. M. Thampi, Eds. Springer Berlin Heidelberg, 2013, pp. 29–39.
- [27] I. Ullah, "Data Mining Algorithms And Medical Sciences," *International Journal of Computer Science & Information Technology (IJCSIT)*, vol. 2, no. 6, 2010.
- [28] H. A. Sadeghi, "Paper: A Probabilistic Bayesian Classifier Approach For Breast Cancer Diagnosis And Prognosis."
- [29] D. Patil, D. Wadhai, M. Vijay, M. Gund, R. Biyani, S. Andhalkar, and B. Agrawal, "An Adaptive parameter free data mining approach for healthcare application," *arXiv preprint arXiv:1211.1788*, 2012.
- [30] A. Shukla, R. Tiwari, and P. Kaur, "Knowledge based approach for Diagnosis of Breast Cancer," in *Advance Computing Conference, 2009. IACC 2009. IEEE International*, 2009, pp. 6–12.
- [31] M. L. Jimenez, J. M. Santamarta, R. Barchino, L. Laita, L. M. Laita, L. A. González, and A. Asenjo, "Knowledge representation for diagnosis of care problems through an expert system: Model of the auto-care deficit situations," *Expert Systems with Applications*, vol. 34, no. 4, pp. 2847–2857, 2008.
- [32] M. G. Tsipouras, T. P. Exarchos, D. I. Fotiadis, A. P. Kotsia, K. V. Vakalis, K. K. Naka, and L. K. Michalis, "Automated diagnosis of coronary artery disease based on data mining and fuzzy modeling," *Information Technology in Biomedicine, IEEE Transactions on*, vol. 12, no. 4, pp. 447–458, 2008.
- [33] N. Bhatla and K. Jyoti, "A Novel Approach for heart disease diagnosis using Data Mining and Fuzzy logic," *International Journal of Computer Applications*, vol. 54, no. 17, pp. 16–21, 2012.
- [34] S. H. Ha and S. H. Joo, "A Hybrid Data Mining Method for the Medical Classification of Chest Pain," *International Journal of Computer and Information Engineering*, vol. 4, no. 1, pp. 33–38, 2010.
- [35] K. Kavitha, K. Ramakrishnan, and M. K. Singh, "Modeling and design of evolutionary neural network for heart disease detection," *International Journal of Computer Science Issues*, vol. 7, no. 5, pp. 272–283, 2010.
- [36] J. Soni, U. Ansari, D. Sharma, and S. Soni, "Predictive data mining for medical diagnosis: An overview of heart disease prediction," *International Journal of Computer Applications*, vol. 17, no. 8, pp. 43–48, 2011.
- [37] M. Shouman, T. Turner, and R. Stocker, "Integrating Naive Bayes and K-means clustering with different initial centroid selection methods in the diagnosis of heart disease patients," *CS & IT-CSCP*, pp. 125–137, 2012.

- [38] M. Shouman, T. Turner, and R. Stocker, "Applying k-nearest neighbour in diagnosing heart disease patients," *International Journal of Information and Education Technology*, vol. 2, no. 3, pp. 220–223, 2012.
- [39] S. B. Patil and Y. Kumaraswamy, "Intelligent and effective heart attack prediction system using data mining and artificial neural network," *European Journal of Scientific Research*, vol. 31, no. 4, pp. 642–656, 2009.
- [40] D. L. Hoyert, J. Xu, and others, "Deaths: preliminary data for 2011," *National vital statistics reports*, vol. 61, no. 6, pp. 1–51, 2012.
- [41] N. J. Pagidipati and T. A. Gaziano, "Estimating deaths from cardiovascular disease: a review of global methodologies of mortality measurement," *Circulation*, vol. 127, no. 6, pp. 749–756, 2013.
- [42] P. Jensen, L. Jensen, S. Brunak, and others, "Mining electronic health records: towards better research applications and clinical care," *IMIA Yearbook*, vol. 8, pp. 188–188, 2013.
- [43] S. Ouhbi, A. Idri, J. L. Fernandez-Aleman, A. Toval, and H. Benjelloun, "Electronic health records for cardiovascular medicine," in *Engineering in Medicine and Biology Society (EMBC), 2014 36th Annual International Conference of the IEEE*, 2014, pp. 1354–1357.
- [44] A. A. Aljumah, M. G. Ahamad, and M. K. Siddiqui, "Predictive Analysis on Hypertension Treatment Using Data Mining Approach in Saudi Arabia," 2011.
- [45] D.-Y. Yeh, C.-H. Cheng, and Y.-W. Chen, "A predictive model for cerebrovascular disease using data mining," *Expert Systems with Applications*, vol. 38, no. 7, pp. 8970–8977, 2011.
- [46] L. Duan, W. N. Street, and E. Xu, "Healthcare information systems: data mining methods in the creation of a clinical recommender system," *Enterprise Information Systems*, vol. 5, no. 2, pp. 169–181, 2011.
- [47] T. Sakthimurugan and S. Poonkuzhali, "An Effective Retrieval of Medical Record using Data Mining Techniques," *International Journal of Pharmaceutical Science and Health Care*, vol. 2, 2012.
- [48] F. A. Badria, M. M. Eissa, M. Elmogy, and M. Hashem, "Rough Based Granular Computing Approach for Making Treatment Decisions of Hepatitis C," in *23rd International Conference on Computer Theory and Applications ICCTA, Alexandria, Egypt*, 2013, pp. 29–31.
- [49] M. M. Eissa, M. Elmogy, M. Hashem, and F. A. Badria, "Hybrid rough genetic algorithm model for making treatment decisions of hepatitis C," in *Engineering and Technology (ICET), 2014 International Conference on*, 2014, pp. 1–8.
- [50] H. Z. Lo, W. Ding, and Z. Nazeri, "Mining Adverse Drug Reactions from Electronic Health Records," in *Data Mining Workshops (ICDMW), 2013 IEEE 13th International Conference on*, 2013, pp. 1137–1140.
- [51] D. Garcia, L. M. M. Cintho, and C. M. C. Moro, "Electronic health record to support Chronic Kidney Disease prevention—Integrating guidelines and archetypes," in *Biomedical and Health Informatics (BHI), 2014 IEEE-EMBS International Conference on*, 2014, pp. 193–196.
- [52] S. D. Navaneethan, S. E. Jolly, J. Sharp, A. Jain, J. D. Schold, M. J. Schreiber Jr, and J. V. Nally Jr, "Electronic health records: a new tool to combat chronic kidney disease?," *Clinical nephrology*, vol. 79, no. 3, p. 175, 2013.
- [53] C. K. Yamin, S. Emani, D. H. Williams, S. R. Lipsitz, A. S. Karson, J. S. Wald, and D. W. Bates, "The digital divide in adoption and use of a personal health record," *Archives of Internal Medicine*, vol. 171, no. 6, pp. 568–574, 2011.
- [54] I. Khan, "Personalized Electronic Health Record System for Monitoring Patients with Chronic Disease," in *Systems and Information Engineering Design Symposium (SIEDS), 2013 IEEE*, 2013, pp. 121–126.
- [55] S. Gupta, D. Kumar, and A. Sharma, "Data mining classification techniques applied for breast cancer diagnosis and prognosis," *Indian Journal of Computer Science and Engineering (IJCSE)*, vol. 2, no. 2, pp. 188–195, 2011.
- [56] D. Delen, "Analysis of cancer data: a data mining approach," *Expert Systems*, vol. 26, no. 1, pp. 100–112, 2009.
- [57] A. Osofisan, O. Adeyemo, B. Sawyerr, and O. Eweje, "Prediction of kidney failure using artificial neural networks," *European Journal of Scientific Research*, vol. 61, no. 4, p. 487, 2011.
- [58] S. G. Jacob and R. G. Ramani, "Efficient classifier for classification of prognostic breast cancer data through data mining techniques," in *Proceedings of the World Congress on Engineering and Computer Science*, 2012, vol. 1, pp. 24–26.
- [59] C. B. Noteboom, S. P. Motomy, S. Qureshi, and S. Sarnikar, "Meaningful Use of Electronic Health Records for Physician Collaboration: A Patient Centered Health Care Perspective," in *System Sciences (HICSS), 2014 47th Hawaii International Conference on*, 2014, pp. 656–666.
- [60] J. Corbin and A. Strauss, *Basics of qualitative research: Techniques and procedures for developing grounded theory*. Sage publications, 2014.
- [61] B. G. Glaser and J. Holton, "Remodeling grounded theory," *Historical Social Research/Historische Sozialforschung. Supplement*, pp. 47–68, 2007.
- [62] G. Walsham, *Interpreting information systems in organizations*. John Wiley & Sons, Inc., 1993.
- [63] K. E. Keick, "Cosmos vs. chaos: Sense and nonsense in electronic contexts," *Organizational Dynamics*, vol. 14, no. 2, pp. 51–64, 1985.
- [64] S. Qureshi, M. Liu, and D. Vogel, "A grounded theory analysis of e-collaboration effects for distributed project management," in *System Sciences, 2005. HICSS'05. Proceedings of the 38th Annual Hawaii International Conference on*, 2005, p. 264c–264c.
- [65] S. Qureshi, "Adaptation in distributed projects: Collaborative processes in digital natives and digital immigrants," in *System Sciences, 2006. HICSS'06*.

Proceedings of the 39th Annual Hawaii International Conference on, 2006, vol. 8, p. 202c–202c.

- [66] S. Qureshi and D. Vogel, “Adaptiveness in virtual teams: Organisational challenges and research directions,” *Group Decision and Negotiation*, vol. 10, no. 1, pp. 27–46, 2001.
- [67] S. El-Sappagh, M. Elmogy, A. Riad, H. Zaghlol, and F. A. Badria, “EHR Data Preparation for Case Based Reasoning Construction,” in *Advanced Machine Learning Technologies and Applications*, Springer, 2014, pp. 483–497.
- [68] M. M. Richter and R. Weber, *Case-Based Reasoning*. Springer, 2013.
- [69] X. Blanco, S. Rodriguez, J. M. Corchado, and C. Zato, “Case-Based Reasoning Applied to Medical Diagnosis and Treatment,” in *Distributed Computing and Artificial Intelligence*, Springer, 2013, pp. 137–146.
- [70] S. Begum, M. U. Ahmed, P. Funk, N. Xiong, and M. Folke, “Case-based reasoning systems in the health sciences: a survey of recent trends and developments,” *Systems, Man, and Cybernetics, Part C: Applications and Reviews, IEEE Transactions on*, vol. 41, no. 4, pp. 421–434, 2011.
- [71] N. G. Weiskopf and C. Weng, “Methods and dimensions of electronic health record data quality assessment: enabling reuse for clinical research,” *Journal of the American Medical Informatics Association*, vol. 20, no. 1, pp. 144–151, 2013.
- [72] X. Xie, L. Lin, and S. Zhong, “Handling missing values and unmatched features in a CBR system for hydro-generator design,” *Computer-Aided Design*, vol. 45, no. 6, pp. 963–976, 2013.
- [73] S. Guessoum, M. T. Laskri, and J. Lieber, “RespiDiag: A case-based reasoning system for the diagnosis of chronic obstructive pulmonary disease,” *Expert Systems with Applications*, vol. 41, no. 2, pp. 267–273, 2014.
- [74] N. Esfandiari, M. R. Babavalian, A.-M. E. Moghadam, and V. K. Tabar, “Knowledge discovery in medicine: Current issue and future trend,” *Expert Systems with Applications*, vol. 41, no. 9, pp. 4434–4463, 2014.
- [75] J. Han, J. C. Rodriguez, and M. Beheshti, “Diabetes data analysis and prediction model discovery using rapidminer,” in *Future Generation Communication and Networking, 2008. FGCN’08. Second International Conference on*, 2008, vol. 3, pp. 96–99.
- [76] A. Pla, B. López, P. Gay, and C. Pous, “eXiT* CBR. v2: Distributed case-based reasoning tool for medical prognosis,” *Decision Support Systems*, vol. 54, no. 3, pp. 1499–1510, 2013.
- [77] P. Andritsos, I. Jurisica, and J. I. Glasgow, “Case-Based Reasoning for Biomedical Informatics and Medicine,” in *Springer Handbook of Bio-/Neuroinformatics*, Springer, 2014, pp. 207–221.
- [78] R. Jagannathan and S. Petrovic, “Dealing with missing values in a clinical case-based reasoning system,” in *Computer Science and Information Technology, 2009. ICCSIT 2009. 2nd IEEE International Conference on*, 2009, pp. 120–124.
- [79] L. C. Molina, L. Belanche, and À. Nebot, “Feature selection algorithms: A survey and experimental evaluation,” in *Data Mining, 2002. ICDM 2003. Proceedings. 2002 IEEE International Conference on*, 2002, pp. 306–313.
- [80] A. Bahga and V. K. Madiseti, “A cloud-based approach for interoperable electronic health records (EHRs),” *Biomedical and Health Informatics, IEEE Journal of*, vol. 17, no. 5, pp. 894–906, 2013.
- [81] L.-H. Lee, Y.-T. Chou, E.-W. Huang, and D.-M. Liou, “Design of a personal health record and health knowledge sharing system using IHE-XDS and OWL,” *Journal of medical systems*, vol. 37, no. 2, pp. 1–12, 2013.
- [82] T. von Bargen, M. Gietzelt, M. Britten, B. Song, K.-H. Wolf, M. Kohlmann, M. Marschollek, and R. Haux, “Design and Implementation of the Standards-Based Personal Intelligent Self-Management System (PICS),” in *Data and Knowledge for Medical Decision Support: Proceedings of the EFMI Special Topic Conference, 17-19 April 2013, Prague, Czech Republic*, 2013, vol. 186, p. 135.
- [83] R. Wu, G.-J. Ahn, and H. Hu, “Secure sharing of electronic health records in clouds,” in *CollaborateCom*, 2012, pp. 711–718.
- [84] H. Hu and G. Ahn, “Enabling verification and conformance testing for access control model,” in *Proceedings of the 13th ACM symposium on Access control models and technologies*, 2008, pp. 195–204.
- [85] H. Hu, G.-J. Ahn, and K. Kulkarni, “Anomaly discovery and resolution in web access control policies,” in *Proceedings of the 16th ACM symposium on Access control models and technologies*, 2011, pp. 165–174.
- [86] J. Jin, G.-J. Ahn, H. Hu, M. J. Covington, and X. Zhang, “Patient-centric authorization framework for sharing electronic health records,” in *Proceedings of the 14th ACM symposium on Access control models and technologies*, 2009, pp. 125–134.
- [87] J. Jin, G.-J. Ahn, H. Hu, M. J. Covington, and X. Zhang, “Patient-centric authorization framework for electronic healthcare services,” *computers & security*, vol. 30, no. 2, pp. 116–127, 2011.
- [88] R. Zhang and L. Liu, “Security models and requirements for healthcare application clouds,” in *Cloud Computing (CLOUD), 2010 IEEE 3rd International Conference on*, 2010, pp. 268–275.
- [89] M. Jafari, R. Safavi-Naini, and N. P. Sheppard, “A rights management approach to protection of privacy in a cloud of electronic health records,” in *Proceedings of the 11th annual ACM workshop on Digital rights management*, 2011, pp. 23–30.
- [90] D. Al Kukhun and F. Sedes, “Adaptive solutions for access control within pervasive healthcare systems,” in *Smart Homes and Health Telematics*, Springer, 2008, pp. 42–53.
- [91] M. Li, S. Yu, K. Ren, and W. Lou, “Securing personal health records in cloud computing: Patient-centric and fine-grained data access control in multi-owner settings,” in *Security and Privacy in Communication Networks*, Springer, 2010, pp. 89–106.

- [92] J. Duarte, G. Pontes, M. Salazar, M. Santos, A. Abelha, and J. Machado, "Stand-alone electronic health record," in *Industrial Engineering and Engineering Management (IEEM), 2013 IEEE International Conference on*, 2013, pp. 1052–1056.
- [93] J. Hook, E. Grant, and A. Samarth, "Health Information Technology and Health Information Exchange Implementation in Rural and Underserved Areas: Findings from the AHRQ Health IT Portfolio (AHRQ Publication No. 10-0047-EF)," *Rockville, MD: Agency for Healthcare Research and Quality*, 2010.
- [94] J. G. Anderson, "Social, ethical and legal barriers to e-health," *International journal of medical informatics*, vol. 76, no. 5, pp. 480–483, 2007.
- [95] K. T. Win, W. Susilo, and Y. Mu, "Personal health record systems and their security protection," *Journal of medical systems*, vol. 30, no. 4, pp. 309–315, 2006.
- [96] D. Charles, M. Furukawa, and M. Hufstader, "Electronic health record systems and intent to attest to meaningful use among non-federal acute care hospitals in the United States: 2008-2011," *ONC Data Brief*, vol. 1, pp. 1–7, 2012.
- [97] R. N. Charette, "EHRs: electronic health records or exceptional hidden risks?," *Commun. ACM*, vol. 49, no. 6, p. 120, 2006.
- [98] M. Eichelberg, T. Aden, J. Riesmeier, A. Dogac, and G. B. Laleci, "A survey and analysis of electronic healthcare record standards," *Acm Computing Surveys (Csur)*, vol. 37, no. 4, pp. 277–315, 2005.
- [99] N. M. Lorenzi, L. L. Novak, J. B. Weiss, C. S. Gadd, and K. M. Unertl, "Crossing the implementation chasm: a proposal for bold action," *Journal of the American Medical Informatics Association*, vol. 15, no. 3, pp. 290–296, 2008.
- [100] R. Dantu, H. Oosterwijk, P. Kolan, and H. Husna, "Securing medical networks," *Network Security*, vol. 2007, no. 6, pp. 13–16, 2007.
- [101] L. Zhou, C. S. Soran, C. A. Jenter, L. A. Volk, Ej. Orav, D. W. Bates, and S. R. Simon, "The relationship between electronic health record use and quality of care over time," *Journal of the American Medical Informatics Association*, vol. 16, no. 4, pp. 457–464, 2009.
- [102] V. H. Castillo, A. I. Martinez-Garcia, and J. Pulido, "A knowledge-based taxonomy of critical factors for adopting electronic health record systems by physicians: a systematic literature review," *BMC medical informatics and decision making*, vol. 10, no. 1, p. 60, 2010.
- [103] E. B. Devine, E. C. Williams, D. P. Martin, D. F. Sittig, P. Tarczy-Hornoch, T. H. Payne, and S. D. Sullivan, "Prescriber and staff perceptions of an electronic prescribing system in primary care: a qualitative assessment," *BMC medical informatics and decision making*, vol. 10, no. 1, p. 72, 2010.