

Intelligence Based Recommendation System for Healthcare: a Patient centred Framework

M. Ambika

Dr. K. Latha

Abstract

Healthcare is a major and rapidly thriving industry with the prime goal of imparting health awareness and health education to the general population. With the growing integration of the Internet in all aspects of our society, many fast-growing organizations are pursuing their vision of becoming trusted providers of health information. A healthcare system needs to be able to analyze a large amount of patient data to uncover insights that improve prediction, early detection and treatment of diseases of public health. The majority of services have a positive impact on people, but some services still fail to strike much of a chord in healthcare services. We have accordingly analyzed the causes behind the lack of success, with the goal of illuminating how services can be improved. We have presented an intelligence based recommendation approach to the healthcare system. The primary reasons for the lack of healthcare services fall the following categories: information quality, trustworthy, authentication, and privacy concerns. The main objective of this research is to provide a framework for an effective healthcare system by addressing these issues.

Keywords- Healthcare; Recommendation system; Classification; Opinion mining; Privacy preservation.

I. Introduction

According to World Health Organization [1], "A good health system should deliver quality services to all individuals, when and where needed." The nature of service differ from country to country, but in all cases oblige an improved health care efficiency, reliable information on which base to take decisions and policies, safety and social benefits [2]. In the recent year's healthcare industries in India has witnessed a momentous growth with the aim to impart good-quality health care services at the lower cost compared to other developed countries. Moreover, now a day's diseases are likelihood to increase by many risk factors and hence avoidances and effective treatment of them are undoubtedly crucial public health and health economics issues in the 21st century. Many public and private sectors have initiated to contribute towards various aspects of health care.

The rapid advancement in World Wide Web, made it conceivable to provide healthcare services through online. Thus many healthcare portals are available in India and

Research Scholar, Department of CSE, Anna University (BIT Campus), Tiruchirappalli, Tamil Nadu, India. Email: mani.ambika@gmail.com

Assistant Professor, Department of CSE, Anna University (BIT Campus), Tiruchirappalli, Tamil Nadu, India. Email: erklatha@gmail.com

worldwide. Most of them provide better services, but some of them never really took off, since there is no awareness among the people about their necessity and the importance. Even the famous organizations have got doomed in health care services, since failed to strike much of a chord. These reasons induced us to analysis the causes, for the success and failure of the health care services and to throw light on the improvisation of health care services.

An extensive literature was done from the end-user perspective to analyze the factors for success and failure of the various health care services. Though, the services offer several features, still lack in the success. The primary reasons come under the following categories. Information qualities in terms of accuracy and relevance are poor, since information was provided by end users without having sufficient medicinal knowledge. Low trust and high risk as the information are not trustworthy and authenticated. Furthermore, people do not trust the system with their sensitive health information, since privacy risks are considered too eminent [3-6].

The main goal of this research is to provide a solution for the rectification of the above enounced issues, for the betterment and benefit to the people and the society. This paper is structured as follows: In section II, reveals the profound and extreme investigation of the review on the research and expounds the proposed work, section III, throw lights on the different techniques used as a part of the proposed work. Finally, we concluded the paper in section IV and outlined some of the future work.

II. Review Of Research

A. Health Care Services

The healthcare is the dynamic research area which requires the usage of several fields like data mining, web mining, machine learning, natural language processing, web personalization, recommender system and security. Indeed, healthcare analytics is a huge and rapidly growing market, recently predicted to reach \$8 billion by 2017.

The Internet plays an essential part to impart and distribute knowledge among worldwide individuals. A survey says that almost 80% of the users have explored the web for health-related topics, and around 117 million grown-ups search for healthcare information [7, 8]. Thus many fast-growing organizations are pursuing their vision to provide healthcare service using web portals. As the number of people seeking for health education online increases rapidly, the demand for healthcare portals surges [9, 10]. Several healthcare portals are available in India and around the world. Seth Emont (2011) done a detailed review on various healthcare portals and elaborates their functions, patient level measures and outcomes.

Healthcare portals must provide relevant, credible and timely information, since misplaced truths may persuade a serious health decision [11]. An effective health care

system should also be easy to use and understand. Hence it is necessary to ensure and evaluate the effectiveness of such health care portals. A National survey has designed to provide meaningful data on how consumers understand, trust, and use the health care information available to them [12-14]. Moreover, an extensive review was done among different group people to find important healthcare factors among access and quality, and cost [5]. As most health care systems are moving towards the patient centered approach [6], it is needed to evaluate such a system from the end user point of view and experiences (Davies, 2005). A detailed review has been done from the end-user perspective to determine factors for evaluation.

Thi et al (2014) identified 16 success factors for implementation of the health information systems based upon empirical studies. Ali Sunyaev (2010) has reviewed the evaluation criteria of a PHR system from end user perspective. A list of 25, end-user feature was elicited for successful implementation. Ton spil et al. (2014) had done a study adopting user perspective in conjunction with ground theory and validated qualitative approach. This approach employed PRIMA model, which incorporates of five areas of analysis; includes process, relevance, information needs, means and people, and attitude.

Marcel Landeweerd (2013) carried out an extensive literature study to provide an overview of current academic insights into the area of e-commerce user adoption. Lee's (2010) objective is to assess diverse interactive features used by healthcare portals to facilitate knowledge management and collaboration between the providers and users. Lee evaluated the portals under three mechanisms such as Knowledge Access Mechanism, Knowledge Creation Mechanism and Knowledge Transfer Mechanism. Muhammad Saeed (2009) has evaluated the healthcare portal based on usability test and post-test questionnaire techniques.

The world the most famous organization like Google, Microsoft, and Yahoo provides health care services. Google has propelled Google Health in the year 2008, which offered individuals the opportunity to manage and access their personal health and medicinal information [3]. They sought to decipher their successful consumer-centered approach from other fields to healthcare, so as to possess a real impact on the people. Similarly, Microsoft launches Microsoft HealthVault to offer health care service and Yahoo provides Yahoo Health. Though many organizations are providing these services, most services of them lack wide adoption among the users [6, 16, 19]. This induces us to do research in health care services.

B. Significance of the study

This research work analyzed the primary drives for the lack of success of the various health care services. Most existing system services lack the features like,

- Accurate prediction and diagnosing of disease
- Support for decision making and health follow up

- Personalized Health Information
- Behavior based predictions

Besides, another study from the end-user perspective was done to analyze the factors for failures like, relevance, information quality, perceived usefulness and usability, high-risk and low-trust, social and individual impact, and perceived compatibility. Some of the following services may not available or partially available, like secure messaging, notification, decision support, support groups, filing referral requests, quality comparison, localization, searching, education information, prescription refills, appointment scheduling, reminders and address book. The fundamental objective of this research is to provide a solution for the rectification of the above enounced issues, for the betterment and benefit to the people and the society.

The proposed work is to improvise the healthcare services, by focusing on the implication of the issues like, information quality, trustworthy, authentication, and privacy concern. The proposed work deals the issues under the three cases as follows,

Case -1: The primary reason for the lack of success lies in information needs. It portrays the information quality, the user like to experience from the system. Information quality influences perceived usefulness and perceived usability, both mediated by the trust. Besides, researchers often measure information quality in terms of relevance, accuracy, timeliness, completeness, and consistency. This research work tries to enhance the information quality by analyzing the Personal Health Record to provide the solutions as,

- Individual based personalized recommendation
- Prediction based on the behaviour and disease
- Predicting the patterns that forecast the incidence of disease in human races.

Case -2: Low trust and high risk emerge as two additional significant reasons for the failure of personal health records. Furthermore, people consider information will not be reliable, if it is provided without sufficient medicinal knowledge. Henceforth, the inclination of risk and safety potentially drive the adoption of trustworthiness. Consequently the proposed methodology has planned to ensure trustworthiness by providing sentiment analysis.

Case -3: In recent years, advances in technology have promoted an increment in the capability to store and record personal data about individuals. Privacy is a real concern when sensitive healthcare information about the patient is utilized for analysis. This research work proposes an approach for reconciliation and searching of healthcare information, in a secure and privacy manner. Clinical text De-Identification can be applied in order to maintain anonymization.

III. Methodology

This section elaborates the proposed approach and its various stages. The Figure.1 demonstrates the entire framework of the Intelligent based Recommendation System. It comprise of the following phases,

- Training Phase
- User profile generation
- Sentiment analysis
- Recommendation
- Privacy preservation

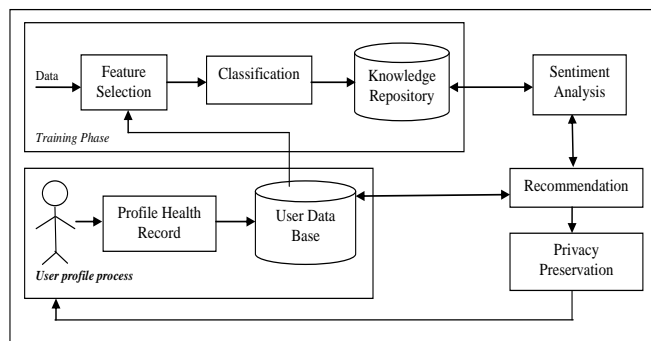


Figure 1: Framework of the Intelligence - Based Recommendation System.

C. Training Phase

In general there are several clinical tests required to detect even a basic disease like fever. To anticipate a disease like cancer, diabetes, hypertension, investigation must be carried on the various factors like age, race, family history, environment, other syndrome, past history of patient and so on. Therefore, many attributes will be required for the study and analysis. Moreover there is rapid increase in the amount of medicinal information being collected electronically in critical care. The one of the crucial steps is the collection of dataset.

The proposed investigation, have a plan to collect real time data from government hospitals, so as to enhance the effectiveness of the recommendation. The data set may contains various information about the patients, including demographics information, diagnoses, medication, and lab results, medical history, life style and environment etc. Pre-processing is to be done in order to improve the dataset and to achieve better predictions. This phase consists of the following process,

- Feature Selection
- Classification

Feature Selection: Feature Selection is very important in prediction as it evacuates unnecessary and useless data, which in turn increases the quality of prediction. The proposed approach selects the relevant features from the dataset using Principal

Component Analysis (PCA) technique. PCA will reduce higher dimensional data into lower dimension. It is a statistical procedure that uses linear dependencies of the original attributes in the dataset to form principal component. It compresses the number of attributes by identifying the strongest pattern in the data. PCA can attain reduction in attributes with lower noise than the original patterns.

Classification: There is an affluence of information acquired within the health care systems. However, there is a lack in realizing hidden relationships and trends in data. The significant aim of this research is to predict the people affected by various maladies. To analyze the medicinal data among the group of people data-mining techniques can be applied. Classification is the process of categorizing data records into one of the set of predefined classes. It is an imperative component of machine learning algorithms in order to extract rules and patterns of data that could be used for prediction. Some of the classifiers are K-Nearest Neighbour, Navie Bayes, Decision Tree, Artificial neural network, rule based classifier, SVM and so on. The proposed work, medical data are classified using SVM. Support vector machines (SVMs) are supervised learning models that analyze data and recognize patterns. The algorithm builds a model using the training set, which assign the new samples to the predefined categories. SVMs operate by finding a linear hyper-plane that separates the positive and negative examples with a maximum inter-class distance. SVM can efficiently perform both linear and non-linear classification [22].

D. User Profile Generation

This stage is utilized to create a user profile definition and user segments. Personal Health Record (PHR) will be generated by the user. The user profile has information that can distinguish one user from a multitude of other users. It is a patient centered approach, where the medical records and other information are accessed, maintained and managed by the patients [26, 27]. PHR has the potential to support decision making and enables the healthcare system towards the personalized medical model [28].

Tang (2006) suggests that PHR includes personal information, major illness, past medical records, family and social history, diagnoses, lifestyle, medications, allergies, immunizations, laboratory tests and appointments. This information can be from different sources, including patients, care givers, physicians, and electronic medical records. If the individual is a new user, then the data has to be pre-processed and pattern has to be extracted. In case of an existing user the profile can be dynamically updated.

E. Sentiment Analysis

In order to support the user based recommendation for the medicinal services, it is necessary to ensure the trustworthiness, as the information provided by the user with or without sufficient medical knowledge. The proposed method guarantees trustworthiness by using sentiment analysis.

An information gathering system mainly depends upon the views of the other people. The recent advancement in the social networking communities, provide an opportunity for the users to share and exchange their opinions and views. Sentiment analysis has provided an efficient and effective way to reveal people's opinion, which is critical for decision making in medical field. It intends to determine the attitude of various users with respect to given topic. The term "sentiment" used in reference to the automatic analysis of evaluative text and tracking of the predictive judgments. Natural language processing techniques can be used to extract sentiments for a topic.

F. Recommendation

The user profile and the extracted rules are combined to provide recommendations to the individuals. Personalized recommendation regarding the prevention and corrective measures, reasons for the causes of the disease, prediction based upon the behavior, health information can be provided. The proposed work uses collaborative based recommendation.

Collaborative filtering approach can be used to predict the interest of a user based on the profiles of the other user with similar preferences. Generally user can express the preferences in the form of rating. Based on the rating and user preferences the system provides personalized recommendation to the users. These recommendations can support decision making process and improve wellness.

G. Privacy Preservation

The health based recommendation system requires pooling of medicinal information in order to improve the quality. Subsequently, ensuring the privacy of an individual plays a vital role in clinical research. In the proposed approach, the integrity of medical information will be preserved while personal identity is effectively concealed.

The proposed approach uses De-Identification process in order to maintain anonymization. De-identification is the process used to prevent a person's identity from being connected with information. The common strategies used for in this process are, deleting or masking personal identities like name, unique identity number and so on, and suppressing or generalizing quasi-identifiers like date of birth and zip code and masking the identity of the highly sensitive data.

IV. Conclusion

Healthcare service acts as a gateway to support qualitative, credible and timely information sharing for patient care and treatment. In this study, we have presented an intelligence based recommendation approach to the healthcare system. There is no universal model to explain success and failure from the end-user perspective. The primary objective is to ensure the availability of the valuable information at the right time by

ensuring information quality, trustworthy and authentication. Thus the health care service can be extemporized and it would be within the reach of every citizen. In our future work, we have planned to implement our framework using the real-time data collected from the government hospitals in order to assist better diagnosis and treatment.

References

- [1] "Health Systems", Available from: http://www.who.int/topics/health_systems/en/, [accessed on 11 August 2014].
- [2] Richard Hillestad, James Bigelow, Anthony Bower, Federico Girosi, Robin Meili, et al. "Can Electronic Medical Record Systems Transform Health Care? Potential Health Benefits, Savings, and Costs." *Health Affairs*, vol. 24(5), 2005, pp.1103-17. [Available from: 10.1377/hlthaff.24.5.1103].
- [3] Ton Spil, Rich Klein. "Personal Health Records Success; Why Google Health failed and what does that mean for Microsoft HealthVault?." 47th Hawaii International Conference on System Science, 2014, pp.2818-27.
- [4] Davies E, Cleary P D. "Hearing the Patient's Voice? Factors Affecting the Use of Patient Survey Data in Quality Improvement." *Quality & Safety in Health Care*, vol. 14(6), 2005, pp.428-32.
- [5] Lisa Clemans-Cope, Bowen Garrett, Katherine Hempstead, Nathaniel Anderson. "Expectations for Health Care Quality, Access, and Costs in 2014", The Urban Institute. 2014.
- [6] Thi Thanh Hai Nguyen, Kaija Saranto, Tommi Tapanainen, Diana Ishmatova. "A Review of Health Information Technology Implementation Success Factors: Importance of Regulation and Finance". *Proceedings of the 47th Hawaii International Conference on System Science 2014*, pp. 2693-04.
- [7] Fox S, Fallows D. "Internet Health Resources". 2003.
- [8] Krane D. "Number of Cyberchondriacs - I.S. Adults Who Go Online for Health Information-Increases to Estimated 117 Million", *Healthcare News Harris Interactive* vol. 5(8), 2005.
- [9] Bernard S, Wisnicki H.J. "How health information on the Internet is evolving", *Ophthalmology Times*, vol. 25(23), 2000, pp.10.
- [10] Bliemel M, Hassanein K. "Consumer satisfaction with online health information retrieval: a model and empirical study", *E-Service Journal*, vol. 5, 2007, pp.53-83.
- [11] Theng Y, Soh E. "An Asian Study of Healthcare Web Portal: Implications for Health care Digital Libraries", Springer- Verlag Berlin Heidelberg, 2005, pp.52-61.
- [12] University of Chicago, "National survey examines perceptions of health care provider quality", *Science Daily*, 23 July 2014. [Available from: www.sciencedaily.com/releases/2014/07/140723110131.htm, accessed on 11 August 2014.]

- [13] Seror A A. "Case analysis of INFOMED: The Cuban National Health Care Telecommunications Network and Portal", *Journal of Medical Internet Research*, vol.8(1), 2006.
- [14] Ranerup, Agneta. "Constructing Consumers in Healthcare: A Comparative Study of National Public Portals in Scandinavia and the UK", *ECIS 2008 Proceedings*, pp.214.
- [15] Ali Sunyaev, Dmitry Chorny, Christian Mauro, Helmut Krcmar. "Evaluation Framework for Personal Health Records: Microsoft HealthVault vs. Google Health", *Proceedings of the 43rd Hawaii International Conference on System Sciences*, 2010, pp.1-10.
- [16] Marcel Landeweerd. "E-commerce user adoption of Google – The Success of Google Search, The Failure of Google Health and the Future of Google+", Thesis, University of Twente, 2013.
- [17] Lee C S, Goh DH, Chua A. "An analysis of knowledge management mechanisms in healthcare portals". *Journal of Librarianship and Information Science*, vol.42(1), 2010, pp.20-44.
- [18] Muhammad Saeed, Sami Ullah. "Usability Evaluation of a Health Web Portal", Master Thesis Computer Science, Thesis no: MSC-2009:16, Blekinge Institute of Technology, Sweden, 2009.
- [19] Marcel Landeweerd, Ton Spil, Richard Klein. "The Success of Google Search, the Failure of Google Health and the Future of Google Plus", *IFIP Advances in Information and Communication Technology*, Springer Berlin Heidelberg, vol.402, 2013, pp.221-39.
- [20] Huan Liu, Hiroshi Motoda. "Feature Selection for Knowledge Discovery and Data Mining". Kluwer Academic Publishers, 1998.
- [21] Ian T Jolliffe. "Principal Component Analysis". Springer, 2nd edition, 2002.
- [22] Nahla H Barakat, Andrew P Bradley. "Intelligible Support Vector Machines for diagnosis of Diabetes Mellitus", *IEEE Transactions on Information Technology in Biomedicine*, vol. 14(4), 2010, pp.1114-20.
- [23] Christopher J C Burges. "A Tutorial on Support Vector Machines for Pattern Recognition". *Data Mining and Knowledge Discovery*, Springer, vol. 2(2), 1998, pp.121-67.
- [24] Herron P. "Machine Learning for Medical Decision Support: Evaluating Diagnostic Performance of Machine Learning Classification Algorithms", *INLS 110, Data Mining*, Springer, 2004.
- [25] Eddy Mayoraz, Ethem Alpaydin. "Support vector machines for multi-class classification, Engineering Applications of Bio-Inspired Artificial Neural Networks", *Lecture Notes in Computer Science*, vol.1607, 1999, pp.833-42.
- [26] Ball E F, Carla Smith N F, Richard S Bakalar. "Personal Health Records: Empowering Consumers", *Journal of Healthcare Information Management*, vol. 21(1), 2007, pp.76-86.

- [27] Archer N, Fevrier-Thomas U, Lokker C, et al. "Personal health records: a scoping review", *Journal of the American Medical Informatics Association*, vol.18, 2011, pp.515-22.
- [28] Kaelber D C, Jha A K, Johnston D, et al. "A research agenda for personal health records (PHRs)", *Journal of the American Medical Informatics Association*, vol.15, 2008, pp.72-36.
- [29] Tang P C, Ash J S, Bates D W, et al. "Personal health records: definitions, benefits, and strategies for overcoming barriers to adoption", *Journal of the American Medical Informatics Association*, vol.13, 2006, pp.121-5.
- [30] Susan Gauch, Mirco Speretta, Aravind Chandramouli, Alessandro Micarelli, "User profiles for personalized information access". *The Adaptive Web, Lecture Notes in Computer Science*, vol.4321, 2007, pp.54-89.
- [31] Sarabjot Singh Anand, Bamshad Mobasher. "Intelligent Techniques for Web Personalization", *Lecturer Notes in Artificial Intelligence, Springer-Verlag Berlin Heidelberg*, vol.3169, 2005, pp.1-36.
- [32] Hu M, Liu B, "Mining and summarizing customer reviews", *Proceedings of the tenth ACM SIGKDD international conference on Knowledge discovery and data mining*, 2004, pp.168-77.
- [33] Ryu J S, Kim W Y, Kim K I, and Kim U M. "Mining opinions from messenger", *Proceedings of the 2nd International Conference on Interaction Sciences: Information Technology, Culture and Human, ICIS, AC*; 2009.pp.287-90.
- [34] Seyed-Ali Bahrainian, Andreas Dengel. "Sentiment Analysis using Sentiment Features", *IEEE/WIC/ACM International Conferences on Web Intelligence (WI) and Intelligent Agent Technology (IAT)*; 2013, pp.26-29.
- [35] Meisamshabanpoor, Mehregan Mahdavi. "Implementation of a Recommender System on Medical Recognition and Treatment", *International Journal of e-Education, e-Business, e-Management and e-Learning*, vol. 2(4), 2012, pp.315-18.
- [36] Martin Wiesner, Daniel Pfeifer. "Health Recommender Systems: Concepts, Requirements, Technical Basics and Challenges", *International Journal of Environmental Research and Public Health*, vol.11, 2014, pp.2580-07.
- [37] Fong A C M, Baoyao Zhou, Hui S C, Guan Y Hong, The Anh Do, "Web Content Recommender System based on Consumer Behavior Modeling", *IEEE Transactions on Consumer Electronics*, vol. 57(2), 2011, pp.962-69.
- [38] Manoj Swami, Manasi Kulkarni. "Understanding Web Personalization with Web Usage Mining and its Application: Recommender System", *International Journal of Emerging Technology and Advanced Engineering*, vol. 3(5), 2013, pp.726-30.
- [39] Fung B C M, Wang K, Chen R, Yu P S. "Privacy-preserving data publishing: A survey on recent developments", *ACM Computing Surveys*, vol. 42(4), 2010, pp.14:1-53.

- [40] Charu C. Aggarwal, Philip S. Yu. "A General Survey of Privacy- Preserving Data Mining Models and Algorithms". The Kluwer International Series on Advances in Database Systems, vol.34, 2008, pp.11-52.
- [41] Lei Chen, Ji-Jiang Yang, Qing Wang, Yu Niu. "A framework for privacy-preserving healthcare data sharing", 2012 IEEE 14th International Conference on e-Health Networking, Applications and Services (Healthcom), pp.341-46.