Data Mining in Healthcare Records: A Review Based on the Kind of Knowledge

Maria Sabastin S.
Department of Management
Amrita Vishwa Vidyapeetham, Bengaluru, Karnataka, India
s_mariasabastin@blr.amrita.edu, ssebastin@mites.amrita.edu

Akshay Anil Nair
Department of Management
Amrita Vishwa Vidyapeetham, Bengaluru, Karnataka, India
akshayanil99@gmail.com

Rohit R.
Department of Management
Amrita Vishwa Vidyapeetham, Bengaluru, Karnataka, India
rohitadoor@gmail.com

ABSTARCT

There is a tremendous amount of attention being focused on improving human health these days. The World Health Organization (WHO) statistics show that disease and mortality rate greatly depend on access to proper healthcare, which is not available to a vast majority of the global population. The healthcare environment is generally perceived as being 'information rich' yet 'knowledge poor'. The knowledge discovery in database (KDD) is alarmed with development of methods and techniques for making use of data. One of the most important step of the KDD is the data mining. Data mining is the process of pattern discovery and extraction where huge amount of data is involved. The analysis of medical records is a major challenge, considering they are generally presented in plain text, have a very specific technical vocabulary, and are nearly always unstructured. It is an interdisciplinary work that requires knowledge from several fields. The analysis may have several goals, such as assistance on clinical decision, classification of medical procedures, and to support hospital management decisions. There is a wealth of data available within the healthcare systems. However, there is a lack of effective analysis tools to discover hidden relationships and trends in data. Knowledge discovery and data mining have found numerous applications in business and scientific domain. Terabytes of data are generated every day in many organizations. To extract hidden predictive information from large volumes of data, data mining (DM) techniques are needed. Organizations are starting to realize the importance of data mining in their strategic planning and successful application of DM techniques can be an enormous payoff for the organizations. Valuable knowledge can be discovered from application of data mining techniques in healthcare system. In this study, we briefly examine the potential use of classification based data mining techniques such as Rule based, decision tree and Artificial Neural Network to massive volume of healthcare data. This paper also discusses the requirements and challenges of DM, and describes major DM techniques such as statistics, artificial intelligence, decision tree approach, genetic algorithm, and visualization. Knowledge discovery in databases is well-defined process consisting of several distinct steps. Data mining is the core step, which results in the discovery of hidden but useful knowledge from massive databases. A formal definition of Knowledge discovery in databases is given as follows: "Data mining is the non-trivial extraction of implicit previously unknown and potentially useful information about data". Data mining technology provides a user- oriented approach

to novel and hidden patterns in the data. The discovered knowledge can be used by the healthcare administrators to improve the quality of service. The discovered knowledge can also be used by the medical practitioners to reduce the number of adverse drug effect, to suggest less expensive therapeutically equivalent alternatives. With the digitalization of medical records and the large amount of medical data available, this is an area of wide research potential. It provides a comprehensive contextualization to all those who wish to perform an analytical work of medical records, enabling the identification of fruitful research.

Keywords Healthcare, Text Mining, Data Mining in Healthcare, Health Informatics, Medical Records

BIBLIOGRAPHY

Maria Sabastin S. Currently serves as Assistant Professor at the Department of Management, Amrita Vishwa Vidyapeetham, Bengaluru Campus. With the passion for teaching and research he started his career with Amrita School of Business, Bangalore campus, and now he is also pursuing his Ph.D. in the area of Information Technology and Management. His key area of interest is to explore what are the determinants to manage Information Technology in Health Care to increase their efficiency and improve their performance.

Akshay Anil Nair is graduate student of MBA-MS dual degree program from Department of Management, Amrita Vishwa Vidyapeetham, Bengaluru Campus.

Rohit R. is graduate student of MBA-MS dual degree program from Department of Management, Amrita Vishwa Vidyapeetham, Bengaluru Campus.