Knowledge Discovery for Clinical Decision Support System in Patient Records

Dev Raj Budhathoki
Department of Computer Science
East Carolina University
Greenville, North Carolina, USA.
budhathokid16@ecualumni.ecu.edu

Kamran Sartipi
Department of Computer Science
East Carolina University
Greenville, North Carolina, USA.
sartipik16@ecu.edu

ABSTRACT

Knowledge discovery from the patients' health records is a challenging task for the medical specialists. The knowledge generated from the patients' records can assist physicians to make an effective decision and recommend more precise diagnosis. This provides the basis for decision-making process with the recommendation for patient diagnosis and expertise advice by retrieving the information from the knowledgebase. This research aims at utilizing data mining techniques to discover patterns and relationships among diagnosis and corresponding symptoms. The extracted patterns are used to assist the physician to determine a precise diagnosis with respect to the patient's context. We consider graph database - Neo4j to develop a knowledgebase that stores the extracted medical knowledge in the ontological format of patterns and relationships. The knowledgebase in clinical decision support system provides effective recommendations for next possible symptoms and diagnosis. In addition, we integrate the expert knowledge with our knowledgebase and explore the features of graph visualization, with more detailed information of patterns and the connection of associated patterns.

KEYWORDS

Knowledge Extraction, Maximum Association Rules, Neo4j Graph Database, Cypher Query Language, Clinical Decision System.

ACM Reference Format:
Dev Raj Budhathoki and Kamran Sartipi. 2020. Knowledge Discovery for Clinical Decision Support System in Patient Records. In *Proceedings of The 33rd IEEE International Symposium on Computer-Based Medical Systems (IEEE CBMS),* Dev Budhathoki and Kamran Sartipi (Eds.). ACM, New York, NY, USA, Article 4, 11 pages. https://doi.org/10.475/123_4

1 INTRODUCTION

Electronic health records (EHR) maintain medical and historical patient data which are growing tremendously due to the availability of advanced information technology in the healthcare domain. Several records and data elements are generated from a single patient during their diagnosis, treatment, and discharge [13] which are recorded in their EHR.

IEEE CBMS, July 2020, Mayo Clinic, Rochester, MN, USA

ACM ISBN 123-4567-24-567/08/06.
https://doi.org/10.475/123_4

Huge capitals are invested in order to provide effective healthcare for a patient, as the statistics show global EHR market share was estimated at USD 20.55 billion in 2016 [34]. Any technological effort to enhance the state of healthcare services to the public will cause huge cost savings for the governments and the public.

In the healthcare domain, quality services with appropriate diagnosis and effective treatment at an affordable cost are challenging. A simple mistake in diagnosis may cause a serious damage with disastrous outcomes. A non-intrusive automated decision support system can assist the healthcare personnel to make correct decisions during the examination of a patient which makes the diagnostic process more objective and reliable. Determining the patterns and relationships from the large dataset generates the knowledge that guides the personnel during the decision-making process.

The analysis of hidden patterns and trends within the data provides a better understanding of disease progression and management. EHR data are in a raw format and include medical history, vital signs, progress notes, diagnoses, medications, immunization dates, allergies, lab data and imaging report [35]. Knowledge discovery from such an EHR is a challenging task as a large amount and a variety of patient data must be preprocessed (anonymized, cleaned, encoded) to be ready for analysis. Next, big data analytics (data mining, machine learning, and statistical analysis) will be applied to create predictive models, associations, and sequential patterns to provide knowledge-driven and accurate intelligent decision support systems.

A Clinical Decision Support System (CDSS) assists medical practitioners to retrieve medical knowledge for making effective decisions with the current condition of a patient and to improve the practitioner's medical practice [7]. A CDSS interacts efficiently with medical practitioners, searches for patterns and connections in the medical knowledgebase, integrates them with knowledge from the domain experts, and provides recommendations for patient diagnosis and treatment, as well as warnings and alerts on special cases.

In this paper, we first apply concept lattice analysis to extract useful patterns from the patient data in the form of frequent itemsets (a maximal group of symptoms shared among a maximal group of diseases) that could help the physicians to explore the associated diagnose as concepts of the lattice. However, we are mainly interested in using these concept-nodes to develop a navigable knowledgebase that could assist physicians during the decision-making process to recommend more appropriate diagnosis that best suits the patient's context. Where the context is the current condition of a particular patient in terms of symptoms and particular historical diseases in the past. The proposed knowledgebase integrates the physician's