Journal of Visual Languages and Computing 25 (2014) 858–867

Contents lists available at ScienceDirect

Journal of Visual Languages and Computing

journal homepage: www.elsevier.com/locate/jvlc

Short Paper

A decision support system for Evidence Based Medicine ☆

Giuseppe Polese

Department of Management & Information Technology, University of Salerno, Italy

ARTICLE INFO

Article history:
Received 25 September 2014
Accepted 30 September 2014
Available online 19 October 2014

Keywords:
EBM
Decision Support Systems
Data mining
Data warehousing

ABSTRACT

We present a decision support system to let medical doctors analyze important clinical data, like patients medical history, diagnosis, or therapy, in order to detect common patterns of knowledge useful in the diagnosis process. The underlying approach mainly exploits case-based reasoning (CBR), which is useful to extract knowledge from previously experienced cases. In particular, we used sequence data mining to detect common patterns in patients histories and to highlight the effects of medical practices, based on evidence.

We also exploited data warehousing techniques, such OLAP queries to let medical doctor analyze diagnosis along several measures, and recent visual data integration approaches and tools to effectively support the complex task of integrating and reconciling data from different medical data sources. In addition, due to massive presence of textual information within the clinical records of many hospitals, text mining techniques have been devised. In particular, we performed lexical analysis of free text in order to extract discriminatory terms and to derive encoded information. Finally, the system provides user friendly mechanisms to manage the protection of confidential medical data.

System validation has been performed, mainly focusing on usability issues, by running experiments based on a large database from a primary public hospital.

1. Introduction

Evidence based medicine (EBM) is the conscientious, explicit, and judicious use of current best evidence in making decisions about the care of individual patients [28]. The goal of evidence-based medicine is to complement the existing clinical decision making processes with the best available research evidence and patient values. Furthermore, EBM provides a direction and rationale for clinicians to manage their patients, and a means to integrate clinical expertise with the best available research evidence. It has been mainly used in the evaluation of clinical therapy effectiveness, especially due to the success of Randomized Controlled Trials (RCTs).

An effective application of EBM requires five important steps: (i) problem definition, (ii) best available evidences identification, (iii) critical evaluation, (iv) identified evidences and patient values integration, (v) whole process evaluation. However, EBM requires performing large scale trials, which require high expenses, prodigious labor, and sophisticated infrastructures, so they cannot be performed frequently [1]. Secondly, prediction of the outcome is required at the time of planning, and optimization of all the possible factors is not easy. Thirdly, a combinatorial explosion problem may be encountered when all drugs are used in a given clinical situation. Fourthly, the different backgrounds of patients can make results controversial, even when almost identical large-scale trials are carried out [1,2]. Lastly and importantly, large-scale trials cannot produce or even predict the new treatment.

To tackle the complexity of EBM processes, decision support systems (DSS) are being increasingly used. In this

☆ This paper has been recommended for acceptance by Shi Kho Chang.

E-mail address: gpolese@unisa.it

http://dx.doi.org/10.1016/j.jvlc.2014.09.013
1045-926X/

paper we present a decision support system for EBM, which exploits sequence data mining, and advanced data integration and data warehousing techniques [16,32,34], together with text mining techniques to process textual information within legacy clinical records [6,29,30]. Moreover, the system embeds modules enabling the visual specification of data privacy policies [15], since in this context it is vital protecting confidential medical data. The proposed system can be used to support decisions about the candidate therapies that can applied to a new patient, by providing medical doctors with information regarding the treatments of past patients, based on the analysis of a large number of patients records.

The system is the final output of a publicly funded research project developed in cooperation between University of Salerno and Healthware, a company specialized in the production of health care information systems. In particular, the system has been integrated within an existing health care information system produced by Healthware, namely the *Healthware NetCare*, which was conceived to support medical staff in the achievement of day-to-day medical operations. Finally, we present a user study to validate the usability of the current system prototype.

The paper is organized as follows. In Section 2 we introduce some principles of data mining and data warehousing in the context of EBM, whereas related works are presented in Section 3. In Section 4 we discuss the approaches underlying the proposed system. The latter is described in Section 5. A user study is presented in Section 6. Finally, conclusions are given in Section 7.

2. Data mining and data warehousing for EBM

In the context of EBM, data mining and data warehousing provide tools to acquire medical data, to extract relevant information from them, and to make this knowledge available to all the people involved in health care.

Decision support systems for EBM can vary in their scope. The simplest systems are fed by data concerning diseases and best practice guidelines to support care delivery. In addition, more sophisticated systems include various clinical internal and external data sources to support further decision making in the area of business management, staff management, and so forth. Relevant data sources in clinical decision support systems for evidence-based medicine purposes are:

- Evidence-based guidelines (in the form of rules).
- Clinical data (patient data, pharmaceutical data, medical treatments, length of stay).
- Administrative data (staff skills, overtime, nursing care hours, staff sick leave).
- Financial data (treatment costs, drug costs, staff salaries, accounting, cost-effectiveness studies).
- Organizational data (room occupation, facilities, equipment).

One of the most relevant applications used to extract knowledge from medical data sources is data mining. In particular, one of the most frequently encountered instances of data mining in the context of EBM is data clustering. The latter involves grouping the data into classes or clusters so that objects within a cluster have high similarity, while objects from different clusters are dissimilar.

The data mining approach underlying the proposed system is based on Sequence Clustering, which enabled us to detect interesting and hidden knowledge on clinical activities.

The most relevant application fields for data warehousing in the area of evidence-based medicine are:

1. The generation process of the evidence-based guidelines.
2. The clinicians at the point of care delivery, by making evidence-based rules available.
3. The monitoring of clinical treatment pathways.
4. The administrative and management tasks, by providing evidence-based knowledge, as well as diverse organizational and financial data.

The medical data warehouse we have built contains data from patients medical records as well as evidence-based guidelines. They are prepared and offered to be queried and analyzed at will.

Often, clinical management is interested in finding out which treatments and medications led to more rapid and cheaper patient convalescence. Data mining and OLAP analytical functions support business decision makers in creating the most effective business strategies satisfying both patient expectations and financial potential.

2.1. Visual data mining

Visual data mining exploits data visualization techniques to help humans in the identification of possible patterns and structures in complex data. It can be seen as a hypothesis-generating process: the user generates a hypothesis about relationships and patterns in the data [10,18].

Visual data mining has several advantages over the automatic data mining methods. It leads to a faster result with a higher degree of human confidence in the findings, because it is intuitive and requires less understanding of complex mathematical and computational background than automatic data mining. This makes visual data mining suitable for decision support in EBM, since it increases the participation of the medical doctor to the decision process. Moreover, visual data mining it is effective when little is known about the data and the exploration goals are vague, since these can be adjusted during the exploration process. It can provide a qualitative overview of the data and it can allow unexpectedly detected phenomena to be pointed out and explored using further quantitative analysis [19].

The visual data mining process starts by forming the criteria about the visualizations to choose and the attributes to display. These criteria are formulated according to the exploration task. The user recognizes patterns in open visualizations and selects a subset of items s/he is interested in. The result of this selection is a restriction of the