Chapter 100

Cooperation between Expert Knowledge and Data Mining Discovered Knowledge

Fernando Alonso
Universidad Politécnica de Madrid, Spain

Aurora Pérez
Universidad Politécnica de Madrid, Spain

Loïc Martínez
Universidad Politécnica de Madrid, Spain

Juan Pedro Valente
Universidad Politécnica de Madrid, Spain

ABSTRACT

Although expert elicited knowledge and data mining discovered knowledge appear to be completely opposite and competing solutions to the same problems, they are actually complementary concepts. Besides, together they maximize their individual qualities. This chapter highlights how each one profits from the other and illustrates their cooperation in existing systems developed in the medical domain. The authors have identified different types of cooperation that combine elicitation and data mining for knowledge acquisition, use expert knowledge to enact the knowledge discovery, use discovered knowledge to validate expert knowledge, and use discovered knowledge to improve the usability of an expert system. The chapter also describes their experience in combining expert and discovered knowledge in the development of a system for processing medical isokinetics data.

INTRODUCTION

Expert knowledge and data mining discovered knowledge are two powerful instruments for producing knowledge-based systems. Together, expert knowledge and data mining discovered knowledge, maximize their individual qualities.

Looking at the discipline of medicine, medical expert systems are a challenging field, now requiring cooperation between both types of knowledge.

From the late 1970s to the early 1990s medical expert systems to aid diagnosis were developed on the basis of rules or models entered and validated by the expert, and data input by the physician after examining the patient and the patient's medical record (Davis, Buchanan, & Shortliffe, 1977;

Shortliffe et al., 1981). The system automatically suggested a diagnosis and helped the physician to make the final decision. These systems were useful because they automatically processed complex decision trees that provided a differential diagnosis. They were not, however, without weaknesses. For example, most of the time experts simplified the knowledge model entered in the system, as they were unable to express all their knowledge as rules.

More recently, automatic knowledge acquisition, discovery and maintenance have attracted a lot of interest. The manual definition of the model by the expert has been overtaken by the automatic discovery of models and reference patterns based on Data Mining (DM) techniques, i.e. automatic medical knowledge elicitation (discovery) from the comprehensive and contextual interpretation of patient data.

The fact is that these two approaches, knowledge discovery and knowledge elicitation from experts, complement rather than oppose each other. Applied together, they can output more realistic and efficient models and patterns: DM techniques can be used to support the tasks involved in Expert System (ES) or Knowledge-Based System (KBS) development, and expert knowledge can be used to improve the knowledge discovery process. The aim of this chapter is to highlight a series of relevant examples of the cooperation between expert knowledge and discovered knowledge and describe the characteristics and possibilities of this cooperation in the medical field. The authors also include an overview of current trends in this field of research. Finally the chapter describes their experience in this field, presenting a long-term project that integrates expert knowledge and DM-discovered knowledge to process isokinetics data.

BACKGROUND

The joint application of knowledge discovered through DM and heuristic knowledge elicited beforehand from the expert has the goal of optimizing and maximizing the performance of the resulting KBS.

To illustrate this cooperation, cases where knowledge is acquired through both expert knowledge acquisition and data mining and where expert knowledge is applied to improve knowledge discovery is presented. Cases where DM techniques are used to aid ES development with respect to aspects like ES validation, improving system usability and efficiency or validating the local application of generic knowledge is also highlighted. This section finalizes with a summary of related work and with an enumeration of major current trends in the field of combining ES and DM.

Knowledge Acquisition: Combining Elicitation and Data Mining

Acquisition of the knowledge to develop a KBS through elicitation from an expert and knowledge discovery is an applicable case of mutual cooperation taking place in the medical field. A rule-based expert system for diagnosing pulmonary tuberculosis, called TUBERDIAG, described by Phuong et al. (1999), presents this type of cooperation. Tuberculosis is still a frequent and very dangerous disease in developing countries. Pulmonary Tuberculosis (PT) is the most frequent form because it spreads easily. Even though it is relatively easy to treat, there are many PT patients because it is often difficult to diagnose. For this reason, the TUBERDIAG system was developed. This system can act like a high-ranking medical consultant specialized in PT diagnosis.

The system's most important components are: knowledge acquisition subsystem, knowledge base, reasoning engine, and explanation subsystem. The knowledge acquisition subsystem uses two sources of rules:

1. **Rules from Experts:** Most rules are formed by doctors. To form these rules, all

22 more pages are available in the full version of this document, which may be purchased using the "Add to Cart" button on the publisher's webpage: www.igi-global.com/chapter/cooperation-between-expert-knowledge-data/73529

Related Content

An Approach for Retrieving Faster Query Results From Data Warehouse Using Synonymous Materialized Queries
Sonali Ashish Chakrabortyand Jyotika Doshi (2021). *International Journal of Data Warehousing and Mining (pp. 85-105).*
www.irma-international.org/article/an-approach-for-retrieving-faster-query-results-from-data-warehouse-using-synonymous-materialized-queries/276766

Efficient and Effective Aggregate Keyword Search on Relational Databases
Luping Li, Stephen Petschulat, Guanting Tang, Jian Peiand Wo-Shun Luk (2012). *International Journal of Data Warehousing and Mining (pp. 41-81).*
www.irma-international.org/article/efficient-effective-aggregate-keyword-search/74755

Cloud-Based Intelligent DSS Design for Emergency Professionals
Shah J. Miah (2013). *Data Mining: Concepts, Methodologies, Tools, and Applications (pp. 991-1003).*
www.irma-international.org/chapter/cloud-based-intelligent-dss-design/73480

Conceptual Model and Design of Semantic Trajectory Data Warehouse
Michael Mireku Kwakye (2020). *International Journal of Data Warehousing and Mining (pp. 108-131).*
www.irma-international.org/article/conceptual-model-and-design-of-semantic-trajectory-data-warehouse/256165

Visual Data Mining for Collaborative Filtering: A State-of-the-Art Survey
Marenglen Biba, Narasimha Rao Vajjhalaand Lediona Nishani (2017). *Collaborative Filtering Using Data Mining and Analysis (pp. 217-235).*
www.irma-international.org/chapter/visual-data-mining-for-collaborative-filtering/159505