

From Data Mining Rules to Medical Logical Modules and Medical Advices

Valentin Gomoia, Mihaela Vidaa, Raul Robua, Vasile Stoicu-Tivadar^a, Elena Bernadb, Oana Lupșe

^a Department of Automation and Applied Informatics, University "Politehnica" of Timișoara, Timișoara, România

^b Department of Obstetrics and Gynaecology, University of Medicine and Pharmacy of Timișoara, Timișoara, România

Abstract

Using data mining in collaboration with Clinical Decision Support Systems adds new knowledge as support for medical diagnosis. The current work presents a tool which translates data mining rules supporting generation of medical advices to Arden Syntax formalism. The developed system was tested with data related to 2326 births that took place in 2010 at the Bega Obstetrics – Gynaecology Hospital, Timișoara. Based on processing these data, 14 medical rules regarding the Apgar score were generated and then translated in Arden Syntax language.

Introduction

Clinical decision support (CDS) systems are developed in order to improve the medical decision process [1]. After a CDS is developed, in order to obtain medical recommendations, two main elements are needed: rules and medical data. Usually, medical rules are extracted from narrative medical guidelines. Alternatively, medical rules are obtained using data mining. In this paper we present a solution for the extraction of medical rules from databases using data mining and the translation of these rules in Arden Syntax language [1], in order to generate better medical advice.

Materials and Methods

This system is based on two major components: the rule extraction part (based on data mining) and the translator to Arden Syntax language (DB2A, Database to Arden Syntax). The rule extraction part is based on data mining. The translation part refers to the translation of "if-then" rules (in the format obtained by using data mining) to the Arden Syntax language. In order to obtain Arden Syntax rules from medical data the next steps are performed: analyzing medical data, applying data mining algorithms, generating medical rules, identifying elements of the rules, XML representation of the rule elements (transmitted to the translation engine), generating Arden Syntax medical rules.

The data source for discovering classification rules is represented by data regarding the 2326 births during 2010 at the Bega Obstetrics – Gynaecology Hospital, Timișoara, Romania. The algorithm used to discover the rules is Ant Colony Optimization from Ant Miner software [2]. After generating medical rules, the text file that contains the rules is transformed into XML format to serve as entry for the DB2A system.

The first step performed by the DB2A is extracting and categorizing the information from the XML file. The information from the XML file is categorized in five types (variable, value, operator, conclusion_variable, conclusion_value) representing different part of the rules. The elements of the arrays (tokens) are further processed in order to be integrated in different MLM categories (Maintenance, Library and Knowledge). Besides the information extracted from the rules, the Maintenance and the Library is generated based on the information related to the: database (to which data mining algorithms are applied), system, and applied data mining algorithms. The Knowledge slot is the core of the MLM and it contains the medical rules. Based on these rules a certain action is triggered. All the five arrays extracted from the XML file are used to generate medical rules

Conclusion

The presented system is able to generate medical rules in a standardized (Arden Syntax) manner, starting from datasets, with data mining technology. The conversion of data mining rules into MLM enables the use of the new medical rules in multiple medical units implementing Arden Syntax standard. The adaptation of the resulted MLMs to a local context can be achieved by modifying the data source addresses in the curly brackets from the data slot. The new medical rules can help in situations in which the medical guidelines cannot offer solutions. The new medical advices can also be used as an alternative to the existing recommendation. The most delicate issue regarding our system is represented by the acceptance of new medical rules by the medical staff. In this respect, each new MLM should be validated by medical experts. Testing the new medical advices in medical units represents the next step in improving their acceptance.

References

- [1] Latoszek-Berendsen A, Tange H, van den Herik HJ, Hasman A. From clinical practice guidelines to computer-interpretable guidelines. A literature overview. *Methods Inf Med.* 2010; 49(6):550-70.
- [2] Parpinelli RS, Lopes HS, Freitas, AA. Data mining with an ant colony optimization algorithm. *IEEE Transactions on Evolutionary Computation.* 2010; 6(4):321-32.

Address for correspondence

Vida Mihaela, Faculty of Automation and Computers, Bd. V. Parvan 2, 300223, Timișoara, Romania, email: mihaela.vida@aut.upt.ro