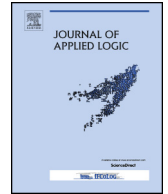




Contents lists available at ScienceDirect

Journal of Applied Logic

www.elsevier.com/locate/jal

Information retrieval from hospital information system: Increasing effectivity using swarm intelligence



Miroslav Bursa^{a,*}, Lenka Lhotska^a, Vaclav Chudacek^a, Jiri Spilka^a,
Petr Janku^{b,**}, Lukas Hruban^b

^a Dept. of Cybernetics, Faculty of Electrical Engineering, Czech Technical University in Prague, Czech Republic

^b Obstetrics and Gynaecology Clinic, University Hospital in Brno, Czech Republic

ARTICLE INFO

Article history:

Available online 18 November 2014

Keywords:

Swarm intelligence
Ant colony
Textual data mining
Medical record processing
Hospital information system

ABSTRACT

This paper details the process of mining information from a hospital information system that has been designed approximately 15 years ago. The information is distributed within database tables in large textual attributes with a free structure. Information retrieval from these information is necessary for complementing cardiotocography signals with additional information that is to be implemented in a decision support system.

The basic statistical overview (*n*-gram analysis) helped with the insight into data structure, however more sophisticated methods have to be used as human (and expert) processing of the whole data were out of consideration: over 620,000 text fields contained text reports in natural language with (many) typographical errors, duplicates, ambiguities, syntax errors and many (nonstandard) abbreviations.

There was a strong need to efficiently determine the overall structure of the database and discover information that is important from the clinical point of view. We have used three different methods: k-means, self-organizing map and a self-organizing approach inspired by ant-colonies that performed clustering of the records. The records were visualized and revealed the most prominent information structure(s) that were consulted with medical experts and served for further mining from the database.

The outcome of this task is a set of ordered or nominal attributes with a structural information that is available for rule discovery mining and automated processing for the research of asphyxia prediction during delivery. The proposed methodology has significantly reduced the processing time of loosely structured textual records for both IT and medical experts.

© 2014 Elsevier B.V. All rights reserved.

* Principal corresponding author.

** Corresponding author.

E-mail addresses: bursam@fel.cvut.cz (M. Bursa), pjanku@fnbrno.cz (P. Janku).

URL: <http://bio.felk.cvut.cz> (M. Bursa).

1. Introduction

The boom of computers, computational appliances, personalized electronics, high-speed networks, increasing storage capacity and data warehouses is present everywhere. An enormous amount of various data is transferred and stored, often mixed from different sources, containing different data types, unusual coding schemes, and seldom come without any errors (or noise) and omissions. Massively parallel distributed storage systems are used nowadays to provide computational nodes with data in reasonable time.

Even with rapidly increasing computational power of modern computers (and cloud technologies), the analysis of large datasets becomes very expensive, justifying the development of novel techniques.

A related phenomenon is that humans are getting flooded by the *big data* and are unable to draw correct conclusions. Especially in text processing, the influence of effective automated methods is crucial. Contrary to classical methods, nature-inspired methods offer many techniques, that can increase speed and robustness of classical methods.

By studying the behavioral and self-organizing patterns in nature, namely ant colonies, we have high chance to discover inspiring concepts for many successful metaheuristics. The high number of individuals and the decentralized approach to task coordination in the studied species indicates that ant colonies show a high degree of parallelism, self-organization and fault tolerance.

In this work we use the ant-colony paradigms for clustering similar literals together to ease the human task of structural discovery in a text.

1.1. Motivation and outcome

Grave asphyxia during childbirth can lead to severe complications, such as several brain damage of the neonate. If asphyxia could be detected in advance, Caesarian section might be indicated to speed up the whole process and reduce the time spent in the oxygen insufficiency. The detection is much more process complicated and the prediction is usually based on the cardiocography signal (CTG) that is being monitored during the delivery. Moreover, supplemental clinical information is needed for making the correct decision – such as Apgar score, biochemical markers (pH value, base excess, base deficit), delivery and neonatal descriptors, risk factors and others.

The outcome of this work is a set of representative attributes retrieved from the text attributes that significantly facilitates rule mining and automated information retrieval that are used for the design of decision support system for asphyxia prediction. The proposed methodology has an important benefit in reduction of the processing time of loosely structured textual records for both informatics and medical experts and allows efficient design and definition of further processing rules (regular expressions). By loosely structured we mean that the information is entered as a free text, the only formatting is performed by the newline characters mainly.

Apart from analyzing the CTG (and related data), the CTG signal database (together with clinical information retrieved with the help of this approach) has been accepted and is currently being published in order to be available for comparative studies.¹

1.2. Nature inspired methods

Nature inspired metaheuristics offer fast and robust solutions in many fields (graph algorithms, feature selection, optimization, clustering, feature selection, etc). Stochastic nature inspired metaheuristics have in-

¹ The open-access database [3] should already be freely available at the following link: <http://www.physionet.org/physiobank/database/ctu-uhb-ctgdb/>.

Download English Version:

<https://daneshyari.com/en/article/4663019>

Download Persian Version:

<https://daneshyari.com/article/4663019>

[Daneshyari.com](https://daneshyari.com)