Expert Systems with Applications xxx (2013) xxx–xxx

Contents lists available at SciVerse ScienceDirect

Expert Systems with Applications

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

Research on clinical decision support systems development for atrophic gastritis screening

Q1 Sergei Parshutin [*,1], Arnis Kirshners

Q2 *Riga Technical University, Institute of Information Technology, Decision Support Systems Group, 1 Kalku Str., LV-1658 Riga, Latvia*

ARTICLE INFO

Keywords:
Decision support
Clinical decision support systems
Data mining
Atrophic gastritis

ABSTRACT

The paper presents a pilot research on the application of clinical decision support systems in a atrophic gastritis screening task. Two different DSS learning strategies have been tested – a standalone classifier and classifier ensemble application. Such classification algorithms as C4.5, CART, JRip and Naive Bayes were used as base classifiers. The classifiers were evaluated on the respondent medical data from an inquiry form, containing 28 attributes and 840 records. The dataset was preprocessed using simple methods in initial data analysis as well as more complex data mining methods for feature selection. The obtained results are summarized and discussed in order to summarize an information on what learning strategies are more applicable to the present dataset and should be studied in more detail in primary research.

1. Introduction

Cancer is the worldwide problem in social health and one of the leading causes of death. Nevertheless it is known that most of cancer types are treatable. Referencing the World Health Organization data, at least 40% of all local cancer types are treatable and can be prevented, avoiding the risk factors, common not only for cancer, but also for the most chronic diseases. These risk factors are well known and the most important of them are smoking, alcohol and other pernicious habits, activity shortage, adiposis (excessive weight) and different infectious agents. New medical technologies, new medicaments, vaccines, screening systems are continuously developed and introduced, all aimed at the identification and treatment of cancer at initial stages and the improvement of life quality and life length for patients with cancer.

Even though globally the gastric cancer incidence is declining and in many Western countries the disease is not considered among the major health issues any more, globally the cancer of the stomach is still continuing to be an important healthcare problem. Gastric cancer is remaining the second leading cause of mortality worldwide within the group of malignant diseases after the lung cancer, and is accounting for almost 10% of cancer related deaths. Among men gastric cancer is the second (after lung cancer), but among women – the third leading (after breast and lung) cause Q5 of cancer-related deaths (Su et al., 2007; WHO, 2013).

Gastric cancer is a very challenging malignancy given that it presents late, has complex pathogenetic mechanisms with multiple carcinogenic processes implicated, and is only moderately sensitive to chemotherapy and radiation. Gastric cancer presents mostly in an advanced stage and is lethal unless diagnosed early (Crew & Neugut, 2006; Miranda, Abelha, Santos, Machado, & Neves, 2009; Varadhachary & Ajani, 2005;).

The present paper discusses a possibility of application of CDSS – Clinical Decision Support Systems in order to give an expert additional information on probable disease; an atrophic gastritis in our case. Section 2 gives a look into CDSS, defines the main objectives of the system and reveals methods used in the pilot research. Section 3 presents the system evaluation results, which are summarized and discussed in Section 4.

2. Clinical decision support system

Clinical decision support systems (CDSS) are computer systems designed to support clinician decision making about specific patients at the point of time these decisions are made. Decision support systems have been incorporated in healthcare information systems for a long time, but these systems have usually supported retrospective analyses of financial and administrative data (Sauter, 2011). Recently, sophisticated data mining approaches have been proposed for similar retrospective analyses of both administrative and clinical data (Intarajak & Kang, 2009; Sauter, 2011), bringing new possibilities in designing efficient clinical decision support systems.

Barner and La Lande in Sauter (2011) propose a simple CDSS classification as knowledge-based and nonknowledge-based

Q3 * Corresponding author. Tel.: +371 67089530; fax: +371 67089678.
Q1 *E-mail addresses:* sergei.parshutin@rtu.lv (S. Parshutin), arnis.kirsners@rtu.lv (A. Kirshners).

[1] DSSG homepage: http://dssg.cs.rtu.lv.

0957-4174/$ - see front matter © 2013 Published by Elsevier Ltd.
http://dx.doi.org/10.1016/j.eswa.2013.05.011

systems. The primary difference pointed out was the ability to generate and apply rules in knowledge-based systems with comparison to the nonknowledge-bases systems that use such methods as artificial neural networks and genetic algorithms (Lisboa & Taktak, 2006; Romero, Valdes, & Barton, 2007; Velikova, Lucas, Ferreira, Samulski, & Karssemeijer, 2008). This CDSS classification is still applicable, but it should be pointed out that modern data mining approaches provide a wide range of possibilities to generate rules using ANN and GA (Galinina & Parshutin, 2011; Quteishat & Lim, 2007). With respect to this classification the knowledge-based CDSS has been chosen for the present pilot research.

2.1. Concepts and objectives

The main objective of a clinical decision support system is to give an expert additional information in order to support an accurate diagnosis. The simplest way to reach this target considers solving the classification task. From this point of view, at least two types of CDSS can be defined – a binary DSS and a multi-class DSS. The multi-class DSS is oriented towards a multi-class classification task; it receives a number of symptoms as an input parameters and returns the most probable disease with such symptoms. The binary DSS is a subtype of a multi-class DSS; it is oriented towards a specific disease, receives input information the same way and gives an answer whether a patient has this specific disease or not.

As our pilot research focuses on a specific disease – atrophic gastritis, thus the application of binary CDSS structure was our choice. The main objectives of CDSS, as we see it, are ability to automatically process the input data – symptoms, inquiry data etc.; and return an information whether the input data is close to positive or negative result. To do so, the CDSS must have a knowledge base, manually constructed or created using a built-in learning algorithm.

2.2. Methods

Medical informatics research has employed traditional statistical methods, such as logistic regression and discriminant analysis, support vector machines, as well as learning methods, such as decision trees, neural networks and case-based reasoning (Chun, Kim, Hahm, Park, & Chun, 2008; Flouris & Duffy, 2006; Miranda et al., 2009; Lee, Lin, & Lee, 2006; Su et al., 2007; Vercellis, 2009). These methods are used individually or combined with other data mining methods.

There are two learning strategies compared in the present research. The first one consists in building a base classifier using a single classification algorithm. The most popular classification algorithms that are widely used are decision trees – C4.5 and CART; rule induction algorithms based on a Ripper algorithm and probabilistic algorithms as NB. Each of these algorithms will be evaluated and compared with others. An output of this learning strategy will be the prediction on the most probable class for a specific descriptive data. An additional information can be obtained if a user-friendly algorithms are user, such as decision trees, giving a possibility to visualize the rules in the knowledge base.

Another learning strategy applies classifier ensembles. This technique constructs a set of base classifiers from training data and performs classification by taking a vote on the predictions made by each base classifier (Tan, Steinbach, & Kumar, 2006). The most popular techniques in this area are bagging and boosting.

Bagging (bootstrap aggregating) is a technique that repeatedly samples (with replacement) from a dataset according to a uniform probability distribution. Each bootstrap sample has the same size as the original data. Each bootstrap is used to built a base classifier. Classification is made by taking a majority vote among the predictions made by each base classifier (Tan et al., 2006).

Boosting is an interactive procedure used to adaptively change the distribution of training examples so that the base classifiers will focus on examples that are hard to classify. Each training example is assigned a weight that may be adaptively changed at the end of each boosting round. Boosting can be applied to a single base classifier or combined with bagging (Tan et al., 2006).

This learning strategy additionally to the class prediction gives probabilities for each class, based on votes of each classifier in ensemble. Classifier efficiency rates, such as classification accuracy, false negative rate etc., can be used as weights, giving weighted class score. Eq. (1) can be used to calculate the score of the jth class. It is generalized equation where j is the class index, $w_{i,j}$ – coefficient of an ith of k classifier for the jth class and $L_{i,j}$ if the binary function, returning 1 if classifier predicts class j and 0 otherwise. The weight $w_{i,j}$ is linked to the class j because such coefficients as false negative rate differ for each class. In case if weight coefficient does not differ upon the class, such as classification accuracy, Eq. (2) can be used.

$$S_j = \frac{\sum_{i=1}^{k} w_{i,j} * L_{i,j}}{\sum_{i=1}^{k} w_{i,j}} \quad (1)$$

$$S_j = \frac{\sum_{i=1}^{k} w_i * L_{i,j}}{\sum_{i=1}^{k} w_i} \quad (2)$$

The final predicted class C will have the best score S', which is chosen using Eq. (3).

$$S' = \begin{cases} argmax(S_j), \ \forall c_j & w \to max \\ argmin(S_j), \ \forall c_j & w \to min \end{cases} \quad (3)$$

3. System evaluation

This section provides an information about the dataset used in experiments, describes the data preprocessing steps taken to prepare data (Section 3.1), gives a look into the experiments performed to evaluate the learning strategies (see Section 2.2) and shows the obtained results (Section 3.2).

3.1. Data description and preprocessing

Learning strategies were evaluated using respondent medical data, who filled the inquiry form. The initial dataset contained 28 descriptive attributes and a binary target attribute denoting positive or negative atrophic gastritis diagnosis. The target attribute values were obtained using the golden standard histological analysis of a respondent tissue examples. In total, the initial dataset contained 840 records, having almost equal class proportions – 430 (51%) positive records and 410 (49%) negative records. Fig. 1

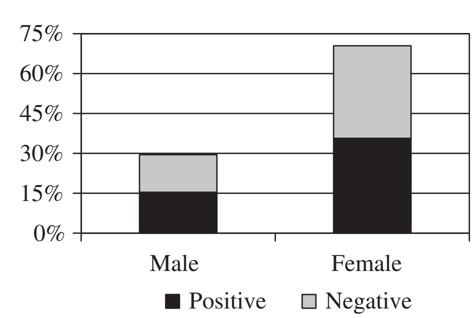

Fig. 1. Class absolute proportions for male and female patients.

Q1 Please cite this article in press as: Parshutin, S., & Kirshners, A. Research on clinical decision support systems development for atrophic gastritis screening. *Expert Systems with Applications* (2013), http://dx.doi.org/10.1016/j.eswa.2013.05.011