International Journal of Medical Informatics 94 (2016) 1–7

Contents lists available at ScienceDirect

International Journal of Medical Informatics

journal homepage: www.ijmijournal.com

Using machine learning to support healthcare professionals in making preauthorisation decisions

Flávio H.D. Araújo [a,*], André M. Santana [b], Pedro de A. Santos Neto [b]

[a] Campus Senador Helvídio Nunes de Barros, Federal University of Piauí, Picos, Piauí, Brazil
[b] Department of Computing, Federal University of Piauí, Teresina, Piauí, Brazil

ARTICLE INFO

Article history:
Received 25 March 2015
Received in revised form 13 June 2016
Accepted 14 June 2016

Keywords:
Classification
Data mining
Ensemble
Prepaid health plans
Machine learning

ABSTRACT

Background: Preauthorisation is a control mechanism that is used by Health Insurance Providers (HIPs) to minimise wastage of resources through the denial of the procedures that were unduly requested. However, an efficient preauthorisation process requires the round-the-clock presence of a preauthorisation reviewer, which increases the operating expenses of the HIP. In this context, the aim of this study was to learn the preauthorisation process using the dental set from an existing database of a non-profit HIP.
Methods: Pre-processing data techniques as filtering algorithms, random under-sample and imputation were used to mitigate problems that arise from the selection of relevant attributes, class balancing and filling unknown data. The performance of classifiers Random Tree, Naive bayes, Support Vector Machine and Nearest Neighbor was evaluated according to kappa index and the best classifiers were combined by using ensembles.
Results: The number of attributes were reduced from 164 to 15 and also were created 12 new attributes from existing discrete data associated with the beneficiary's history. The final result was the development of a decision support mechanism that yielded hit rates above 96%.
Conclusions: It is possible to create a tool based on computational intelligence techniques to evaluate the requests of test/procedure with a high accuracy. This tool can be used to support the activities of the professionals and automatically evaluate less complex cases, like requests not involving risk to the life of patients.

1. Introduction

Although Brazil has a public and universal health system, its supplementary health market is one of the largest worldwide. The National Health Agency (Agência Nacional de Saúde)—ANS [2] reports that there are currently approximately 1500 health insurance providers (HIPs) in Brazil, totalling more than 50 million people covered (ANS 2013). However, ANS also reported that most HIPs are in a very difficult financial situation. The costs of small companies exceed their revenue, mid-sized companies are at their operating limit, and only large companies have any financial stability.

HIPs have different control mechanisms for continuously verifying whether requests are being filed adequately and avoiding improper requests and even intentional fraud. Among these control mechanisms, preauthorisation (claim approval) stands out and can be defined in this context as conducting a technical review of the requested procedures/tests/treatments to determine the best option for the patient, thereby avoiding unnecessary requests.

We summarise the health preauthorisation process here. After a healthcare professional or an associated provider (hospital or clinic) requests a test or procedure, the HIP determines whether the request conforms with existing standards and if the clinical protocols that have been established in the contract between the parties are legal. If the analysis of the preauthorisation reviewer is positive, then the test/procedure is authorised; otherwise, the operator questions the petitioner about the merits of the request and does not authorise the request. Fig. 1 is a schematic of the preauthorisation process.

Health systems all over the world are facing considerable pressure to reduce costs while having to sustain or even improve the quality of health service delivery [16]. High cost is the primary obstacle to developing an efficient preauthorisation process. In small HIPs, for instance, few events per day occur that require preauthorisation. Thus, it very expensive to maintain a round-the-

* Corresponding author.
E-mail addresses: flavio86@ufpi.edu.br (F.H.D. Araújo), andremacedo@ufpi.edu.br (A.M. Santana), pasn@ufpi.edu.br (P. de A. Santos Neto).

http://dx.doi.org/10.1016/j.ijmedinf.2016.06.007
1386-5056/

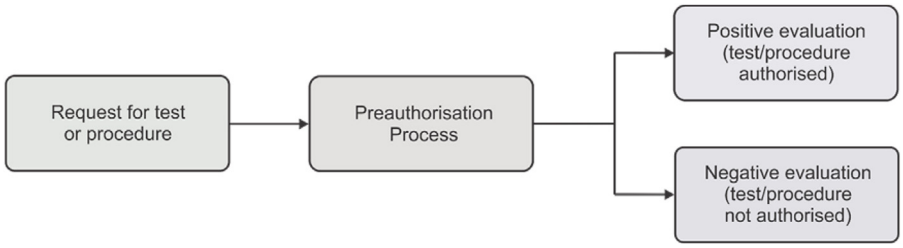

Fig. 1. Schematic of preauthorisation process.

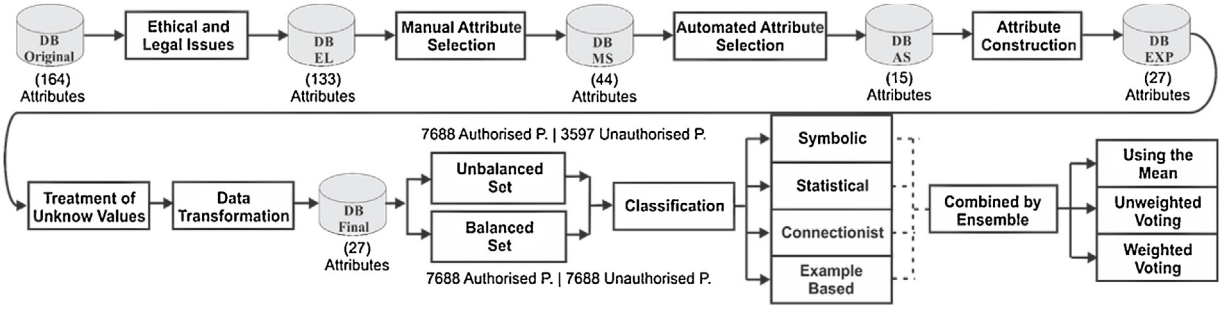

Fig. 2. Methodology for automated learning of the healthcare preauthorisation process.

clock preauthorisation reviewer to analyse requests. However, the absence of a preauthorisation reviewer can result in unnecessary procedures being requested, which places strong financial pressure on the HIP. An alternative that many companies often use is to forego active control while maintaining a strong audit process to block the payment of procedures/tests/treatments that have been performed on the basis of unsuitability or so-called disallowance. This procedure, however, results in significant burnout of the involved parties, creates disharmonious relationships and makes the provision of health services a highly antagonistic process. That is, this process is not conducive to a strong and lasting relationship among the parties.

The preauthorisation process critically affects patients' health because the denial of a procedure could potentially result in a patient's death. Thus, the preauthorisation mechanism must be carefully considered to prevent problems.

The preauthorisation process in most HIPs is facilitated by computer technology, i.e., there is a data record of orders and the results of requests (authorised/unauthorised). Therefore, supervised learning techniques can be used to learn the preauthorisation process by analysing the data that are stored in HIP databases (DBs). However, most HIP DBs have problems such as redundant and missing values that reduce the quality of the available information. Thus, pre-processing techniques are required to improve data quality before a learning process can be performed. In this study, we used earlier studies in which pre-processing techniques were used to improve data followed by an automated learning process as references. These studies were compared in Table 1.

A literature review showed that numerous studies in the medical field have used pre-processing techniques combined with machine learning algorithms to facilitate the decision-making process. However, data from HIPs were used to facilitate decision-making in only three studies. The objectives of the study of Barros et al. [4] and Martins et al. [15] were different from those of the present study: the aforementioned authors used information about the HIP to find production rules to associated patients with certain groups of diseases, whereas our objective was to automate the learning of the preauthorisation process. However, Araújo et al. [3] had a similar objective to that of the present study, i.e., the automated learning of the preauthorisation process. However, Araújo et al. [3] used medical data from a HIP to learn medical preauthorisation, whereas we used dental data to learn dental preauthorisation.

In the present study, a mechanism for supporting healthcare preauthorisation was developed to assist healthcare professionals in decision-making, whereby a potential decision was generated by learning the HIP database. Thus, this process is an important tool for assisting healthcare professionals.

This paper is structured as follows: all of the steps for learning the preauthorisation process are described in Section 2; the results obtained using the classifiers are presented in Section 3; a discussion about how to use the proposed methodology and the limitations of this work are presented in Section 4; conclusions and suggestions for future studies are described in Section 5.

2. Methods

Pre-processing techniques were used in this study to improve the quality of the data in a HIP database. These data were then combined with machine learning algorithms to learn the healthcare preauthorisation process. In this study, we used dental data, but the original version of this technique was performed using medical data [3]. Fig. 2 shows all of the steps that were performed in this study. Details on the execution of each step are given below.

First, the HIP's analysts generated a DB (DB-Original) that contained all of the data that were related to the dental procedures. This DB had 164 attributes; however, for ethical and legal reasons, all of the data that identified individuals, such as IDs, SSNs, dates of birth, addresses and phone numbers, were removed from the DB. In addition, an artificial key was generated for the DB to identify all of the anonymous individuals. The resulting database from which the attributes that identified individuals were eliminated was called DB-EL.

The DB-EL comprised 133 attributes of which many attributes were irrelevant and in a format that was incompatible with the algorithms that were used during the mining phase. Therefore, the DB-EL was subjected to pre-processing to improve the data quality and decrease the amount of irrelevant information.

An attribute selection step was performed in the first step of pre-processing. This selection was performed by manual and auto-