Expert system for predicting unstable angina based on Bayesian networks ☆

Joan Vila-Francés [a,*], Juan Sanchís [b], Emilio Soria-Olivas [a], Antonio José Serrano [a], Marcelino Martínez-Sober [a], Clara Bonanad [b], Silvia Ventura [b]

[a] Intelligent Data Analysis Laboratory, University of Valencia, Avd Universitat s/n, 46100 Burjassot, Valencia, Spain
[b] Servicio de Cardiología, Hospital Clinic Universitari de Valencia, Spain

ARTICLE INFO

ABSTRACT

The use of computer-based clinical decision support (CDS) tools is growing significantly in recent years. These tools help reduce waiting lists, minimise patient risks and, at the same time, optimise the cost health resources. In this paper, we present a CDS application that predicts the probability of having unstable angina based on clinical data. Due to the characteristics of the variables (mostly binary) a Bayesian network model was chosen to support the system. Bayesian-network model was constructed using a population of 1164 patients, and subsequently was validated with a population of 103 patients. The validation results, with a negative predictive value (NPV) of 91%, demonstrate its applicability to help clinicians. The final model was implemented as a web application that is currently been validated by clinician specialists.

1. Introduction

The use of computer-based clinical decision support (CDS) tools is growing in recent years due to different reasons (Steyerberg, 2009):

- Helps the clinicians making decisions, thus reducing the clinical errors.
- Improves the time to get a diagnostic reducing waiting time.
- Optimizes health resources reducing unnecessary medical tests.

Those advantages have expanded the use of these tools in clinical practice in the form of web services, desktop programs or applications for mobile phones and tablets. The use of these tools in different clinical areas, supported by the increasing amount of patient data and the ability to analyse and process it through Big Data techniques, is foreseen as a huge boost in health care (Manyika et al., 2011). Currently, one of the areas that demands more resources is cardiology. Nowadays cardiovascular diseases are the main cause of death in the developed world (Escaned et al., 2008). Modern lifestyle that leads us to many stressful situations, poor diet and little exercise have made heart disease the cause of many deaths. One of the symptoms of possible heart failure angina is characterised by severe chest pain. When suffering from that pain for more than 15 min, it is highly recommended requiring clinical attention, as it can be the initial phase of a myocardial infarction. This pain occurs when the demand for oxygen by the heart muscle is not served (because there is an interruption of the blood supply to a part of the heart muscle). This pain is in fact one of the most frequent causes of admission in the Emergency Services of the hospitals. Some of these patients suffer an acute coronary syndrome that is diagnosed by electrocardiogram (ECG) findings or alteration of biomarkers of myocardial damage (troponin). However, in some patients the ECG is nonspecific and troponin is normal. This population suffers form a chest pain of uncertain origin. They are mostly low-risk patients without any heart disease, but we can not reject an acute coronary syndrome in some of them. The key challenge is to identify those patients at risk for suffering an acute coronary syndrome with normal troponin (unstable angina), within a population that is generally at low risk. The tools currently available in the emergency room of a hospital do not work, because the ECG is nonspecific, and troponin is normal. As a result, the final decision of admission or release is postponed until a treadmill stress test is carried out (usually the next morning) Sanchis et al. (2006). This strategy is suboptimal because the patient must wait for several hours, many patients can not run on the treadmill, and sometimes the results are inconclusive.

This article proposes the development of a Clinical Decision Support System, for being use in the emergency units of a hospital in order to determine the probability of unstable angina within

24 h of patient entry into the hospital. The inputs of the system are the clinical data that are routinely collected in the emergency room of a hospital. Almost all of those data have a binary response (e.g. – patient gender, smoker, etc.). With this kind of inputs, the most appropriate machine learning models are decision trees and Bayesian networks (Alpaydin, 2009). The results obtained from decision trees were not good enough (deep trees were needed, and therefore generalisation was bad). Moreover, a system whose parameters could be updated continuously when new information was available was required, that is why Bayesian networks were used. The advantages of these models for using them in clinical problems are Lucas et al. (2004):

- The Bayesian model can be interpreted by the clinician, as the relationships among variables are clearly represented by a graph (directed graph in our model).
- The clinician can provide expertise knowledge to establish new relationships between variables that might not have been reflected in the case of using an automatic learning algorithm for the Bayesian network structure.
- Adding new knowledge is a straight forward process, which can be automated by updating the frequency tables of each input variable.
- It is not necessary to know the values of all inputs to the model to obtain a valid output. Thus, if the inputs are obtained in a sequential way, as it happens in an emergency unit, or some information about the patient is missing, the system may be updating the probability as soon as new information of clinical tests is available.

Once the Bayesian network was validated, it was implemented into a Decision Support System which allowed clinicians to access it remotely, in an easy and simple way. The obvious solution was to implement the expert system as a web application. This way it is possible to centralise the data at a single location while access can be granted from any computer without requiring dedicated software (only a web browser and an Internet connection). Moreover, it is possible to implement a user-based management to control the people that use the tool and access the data. The rest of this paper is organised as follows. Section 2 explains the Bayesian models used. Section 3 discusses the data used and results obtained. Section 4 explains the Web tool development. Finally, Section 5 summarizes the conclusions of the present work.

2. Bayesian networks

A Bayesian network (BN) is a probabilistic graphical model composed of two different parts: on one hand is the graphical structure (directed acyclic graph) that defines the relationship between variables and, on the other hand, the probabilities established between these variables (Koller and Friedman, 2009; Korb and Nicholson, 2011). The elements of a Bayesian network are as follows Rusell and Norvig (2009):

- A set of variables (continuous or discrete) forming the network nodes.
- A set of directed links that connect a pair of nodes. If there is a relationship with direction $X \rightarrow Y$ is said that X is the parent of Y.

The network fulfils the following facts:

- Each node X_i is associated with a conditional probability function $P(X_i|Parents(X_i))$ that takes as input a particular set of values for the node's parent variables, and gives the probability of the variable represented by the node X_i.

- The graph has no directed cycles.

The knowledge is reflected by the relationships established in the graph nodes, and the conditional probabilities of the variables represented in each node. Those probabilities are estimated using the dataset. In this paper, the most widely used Bayesian classifiers have been applied: Naïve Bayes, FAN (Forest Augmented Network) and TAN (Tree Augmented Network). The following subsections explain each of these approaches.

2.1. Naïve Bayes

The hypothesis of this approach is to assume that the predictive variables are conditionally independent given the variable to classify. Although this assumption is quite restrictive, this classifier is one of the most used, and several studies show that the results are as good as the ones obtained with other techniques (neural networks, decision trees etc.) Korb and Nicholson (2011). The big problem with this model arises when the conditional independence among predictive variables is not fulfilled. This happens when predictive variables are redundant or are highly correlated with each other. Assuming the conditional independence, the graph of a Naïve Bayes is shown in Fig. 1.

In Naïve Bayes approach, the conditional probability $P(class|X_1, X_2, \ldots, X_N)$ is factorized as $P(class|X_1, X_2, \ldots, X_N) = \prod_{i=1}^{N} P(class|X_i)$. This factorisation is more easily obtained from experimental data, in addition to being easier to analyze and, later on, to obtain inferences from available information.

2.2. Tree Augmented Network, TAN

This structure is an extension of a Bayesian classifier in which each variable is allowed to have another parent outside the class node. The idea is to build a Bayesian network tree for all predictive variables and complete the model with a Naïve Bayes. TAN algorithm forms a tree with the predictive variables and then add edges to the class node. Fig. 2 is a conceptual illustration of how the model works. It is seen that each predictive variable, X_i, can have up to two parents.

2.3. Forest Augmented Network, FAN

An important limitation that TAN model may have, *a priori*, is that some arcs of the tree, formed between the descriptive variables, can introduce noise into the classification if such relationships do not exist. This model proposes the formation of disjoint trees with predictive variables. So this approach creates a tree structure with all the variables. Obviously many of the dependencies are enforced by the method of construction and do not exist. These dependencies are discarded during the process of creating the tree; an edge is discarded if it is independent (for example, using a statistical χ^2 test.). Fig. 3 shows a model built from two disjoint tree predictors.
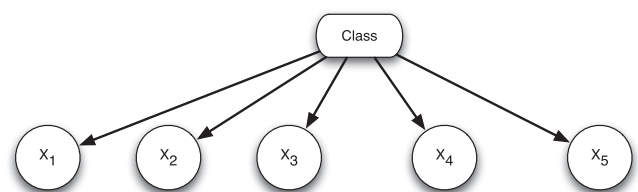

Fig. 1. Structure of a Naïve network.