Application of Bayesian networks in emergency medicine

Miljenko Križmarić*, Matej Mertik**

*Faculty of Health Sciences
University of Maribor
Žitna 15, 2000 Maribor, Slovenia,

**Faculty of information sciences
Novi trg 5, 8000 Novo mesto, Slovenia

miljenko.krizmaric@uni-mb.si
matej.mertik@fis.si

Abstract. *There are large collections and sets of data about the patients, diagnosis, treatment procedures etc... in the field of medicine today. Using data mining techniques in those cases provide a statistical and logical analysis of the data looking for patterns that can aid by decision making and prediction.*

In this paper we presents study of using Bayesian network (BN) in the domain of emergency medicine where BN are especially appropriate because of their symbolic representation, handling of uncertainity, where different scenarios are possible by given evidences. We show a use of BN in the case of study out-of-hospital cardiac arrest in emergency medicine, where we use the BN as tool for prediction of return of spontaneous circulation (ROSC) and survival of hospital discharge.

Based on experiences we conclude the paper with discussion of some future applications of using Bayesian networks on human patient simulators in emergency medicine.

Keywords: Data mining, machine learning, Bayesian networks, emergency medicine, human patient simulators

1. Introduction

Data mining focuses on sorting through large amounts of data and picking out relevant information. It is usually used by business intelligence organizations, and financial analysts, but is increasingly being used in the sciences to extract information from the enormous data sets generated by modern experimental and observational methods. It is defined as the science of extracting useful information from large data sets or databases [1].

The medicine today uses large collections and sets of data about the patients, diagnosis, treatment procedures etc... Using data mining techniques in those cases provide a statistical and logical analysis of the data looking for patterns that can aid decision making. However, there is enormous list of data mining methods which are all trying to extract useful information, but for getting valuable information, we should choose the proper methods or their combination.

In this paper we presents a study of using Bayesian network as data mining tool for decision support in the domain of emergency medicine. We shortly introduce well known and used data mining methods in general that are used in the domain of medicine and in particular explain the method of Bayesian network, where we show important properties that are especially appropriate for applications in the emergency medicine. We show a real application of using Bayesian network on the dataset of emergency medicine that was provided by Centre for Emergency Medicine in Maribor. Based on the network important relationship between parameters are showed and can be simulated whereby experts are provided by the causes and types of cardiac arrests and different procedures needed for saving a patient. We conclude the paper with discussion and explanation of some future applications of using Bayesian networks on human patient simulators in education processes of emergency medicine.

2. Data mining methods in medicine

In general, data mining algorithms represents machine learning methods which can be classified as supervised learning or unsupervised learning. In supervised learning, training examples consist of input/output pair patterns. Learning algorithms aim to predict output values of new examples based on their input values. In unsupervised learning, training examples contain only the input patterns and no explicit target output is associated with each input. The unsupervised learning algorithms need to use the input values to discover meaningful association or patterns.

Many successful machine learning systems have been developed over the past three decades in the computer science and statistics communities. Chen and Chau [2] categorized five major paradigms of machine learning research; probabilistic and statistical models, symbolic learning and rule induction, neural networks, evolution-based models, and analytic learning and fuzzy logic.

Probabilistic and statistical analysis techniques and models have the longest history and strongest theoretical foundation for data analysis. Popular statistical techniques, such as regression analysis, discriminant analysis, time series analysis, principal component analysis, and multi-dimensional scaling, are widely used in medical data especially in biomedical data analysis. Popular probabilistic models represents Bayesian model, which simple variation called Naive Bayesian model, that assumes mutually independently of features within each class, has been adopted in different domains [3]. Popular and wide used technique in recent years is the support vector machines (SVMs) which is based on statistical learning theory that tries to find a hyperplane to best separate two or multiple classes [4]. SVM is suitable for various medical classification problems, medical diagnosis based on patient indicators.

Symbolic learning can be classified according to its underlying learning strategy where learning from examples appears to be most promising symbolic learning approach for knowledge discovery and data mining [5]. It is implemented by applying an algorithm that attempts to induce a general concept description that best describes the different classes of the training examples. Numerous algorithms have been developed, each using one or more different techniques to identify patterns that are useful in generating a concept description. Most widely used symbolic learning technique become Quinlan's ID3 decision-tree building algorithm and its variations such as C4.5 [6] where given a set of objects a decision tree that attempts to classify all the given objects correctly is produced. At each step, the algorithm finds the attribute that best divides the objects into the different classes by minimizing entropy (information uncertainty). By the end of the procedure the results can be represented by a decision tree or a set of production rules. Although not as powerful as SVM or neural networks (in terms of classification accuracy), symbolic learning techniques are computationally efficient and their results are easy to interpret. This ability is invaluable in the medicine domain where data mining results in a way can be understandable to physicians and patients. Powerful machine learning techniques such as SVM and neural networks often suffer because they are treated as a "black-box."

Neural networks are following paradigms in machine learning, where they attempt to achieve human-like performance by modeling the human nervous system. A neural network is a graph of many active nodes (neurons) that are connected with each other by weighted links (synapses). Knowledge is learned and remembered by a network of interconnected neurons, weighted synapses, and threshold logic units, where based on training examples, learning algorithms can be used to adjust the connection weights in the network such that it can predict or classify unknown examples correctly. However neural networks often suffer because they are treated as a "black-box", while knowledge by symbolic descriptions such as decision trees and production rules is represented in understandable way.

Evolution-based algorithms rely on analogies to natural processes and Darwinian survival of the fittest. Fogel [7] identifies three categories of evolution-based algorithms: genetic algorithms, evolution strategies, and evolutionary programming. Among these, genetic algorithms are the most popular and have been successfully applied to various optimization problems. Genetic algorithms were developed based on the principle of genetics [8].

Analytic learning represents knowledge as logical rules and performs reasoning on such rules to search for proofs. Proofs can be compiled into more complex rules to solve similar problems with a smaller number of searches required. For example, Samuelson and Rayner [8] used analytic learning to represent grammatical rules that improve the speed of a parsing system. The analytic learning systems were improved with fuzzy systems and fuzzy logic due to the hard computing rules which have clear distinction between values and classes, what is not a case in the real world.

Not at last hybrid approaches combine mentioned different approaches to over claim some disabilities of the single methods. Although the results of such combination (ensembles) are often not understandable as classical symbolic methods they provide better classification/prediction accuracies.

In the following section we presents the Bayesian network model which fit between symbolic and probabilistic methods and is gaining more and more importance in the recent years.

3. Bayesian network

A Bayesian network is a probabilistic graphical model that represents a set of variables and their probabilistic independences. For example, a Bayesian network could represent the probabilistic relationships between diseases and symptoms. Given symptoms, the network can be used to compute the probabilities of the presence of various diseases.

Formally, Bayesian networks are directed acyclic graphs whose nodes represent variables, and whose arcs encode conditional independences between the variables. Nodes can represent any kind of variable, be it a measured parameter, a latent variable or a hypothesis. They are not restricted to representing random variables, which represents another "Bayesian" aspect of a Bayesian network. Efficient algorithms exist that perform inference and learning in Bayesian networks.

Bayesian Network approach merges and supersedes existing approaches coming from machine learning and data mining, both symbolic and statistical ones. Bayesian Networks are rigorously justified, provide a distributed knowledge representation, and are as understandable as a rule base. They deal particularly well with uncertainty, and they can be manually generated by consultation of an expert, or inductively built by machine learning. Therefore Bayesian networks are used for modeling knowledge in different domains as bioinformatics (gene regulatory networks, protein structure), medicine, document classification, image processing, data fusion, decision support systems, engineering, and law[9].

For example, the network can be used to find out updated knowledge of the state of a subset of variables when other variables (the evidence variables) are observed. This process of computing the posterior distribution of variables given evidence is called probabilistic inference. The posterior gives a universal sufficient statistic for detection applications, when one wants to choose values for the variable subset which minimize some expected loss function, for instance the probability of decision error. A Bayesian network can thus be considered a mechanism for automatically applying Bayes' theorem to complex problems.

In the following section we represent the application of Bayesian network in emergency medicine.

4. Emergency medicine - study out-of-hospital cardiac arrest

In this paper the study data was collected for patients with out-of-hospital cardiac arrest in Emergency centre Maribor. Dataset from the study of Grmec et al [10] was used for application of the Bayesian network model. The complete data was collected for patients with out-of-hospital cardiac arrest in Maribor. Advanced cardiac life support was initiated in 737 patients among which 495 were male and 242 female. Unfortunately, only 170 patients survived and were discharged alive from hospital. Table 1 summarizes important characteristics of the dataset.

Table 1: dataset characteristics

Number of instances:	737
Number of attributes:	7
Number of missing values:	0

Following parameters were used in the study: gender, age, cause, initial rhythm, bystander, return of spontaneous circulation and survival. Below some detailed description of parameters are presented.

- ●Cause describe the etiology of cardiac arrest, which can be primary cardiac or non-cardiac (Card, Ncard).
- ●Initial rhythm as the first cardiac rhythm present when a monitor or defibrillator is attached to a patient (ECG) after a cardiac arrest. Rhythm is described as asystole (ASY), ventrical fibrillation (VF), pulseless ventrical tachycardia (VT) and pulseless electrical activity (PEA).
- ●Bystander - many victims of sudden cardiac arrest (SCA) can survive if bystanders act immediately while VF is still present, but successful resuscitation is unlikely once the rhythm has deteriorated to asystole (ASY). The optimum treatment for VF cardiac arrest is immediate bystander CPR (combined chest compression and rescue breathing) plus electrical defibrillation.
- ●Return of spontaneous circulation (ROSC) can be described as successful resuscitation on the field. Signs of ROSC include breathing, coughing or movement, but for healthcare personnel they also may include evidence of a palpable pulse or a measurable blood pressure.

5. Bayesian network in emergency medicine

The Bayesian network was manually generated by consultation of an expert. It was trained on the presented dataset. Figure 1 presents manually generated Bayesian network which was trained at presented dataset. The GeNIe (Graphical Network Interface) software package was used for this purposes [11].

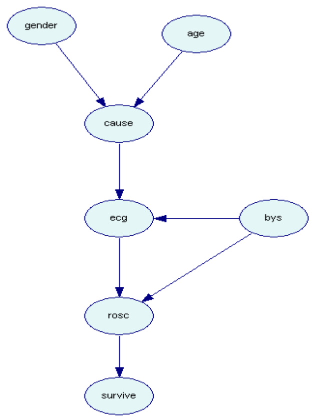

Figure 1: Bayesian network structure.

The Bayseian network on the figure above consists from conditional probability tables and prior probabilities where conditionally probability tables (hidden and defined in every node) are calculated from every node by given parents (nodes that directly depends upon).

We can see that expert define three variables with apriori probabilities: gender, age and bystanders. The variables of cause, initial rhythm, return of spontaneous circulation and survive are variables that depends directly from their nodes (parents) above. We can see that the cause depends from age and gender for example and initial ECG rhythm from cause and bystanders.

6. Experimental settings

One among useful properties of Bayesian networks as data mining models is modeling different scenarios by given the evidence (inference), what is not possible by other symbolic approaches as decision trees. Given evidence about the state of a variable, or set of variables, the state of other variables can be inferred. For example, to find the probability of non-cardiac cause when gender is a male under 45 years, that is to find p(cause = non-cardiac | gender = M, age = below_45), it is necessary to set the evidence of gender and age at appropriate values and then recalculate the network where Bayes rule is then applied for posterior probability.
At figure 2 we can see the trained Bayesian network for the presented case.

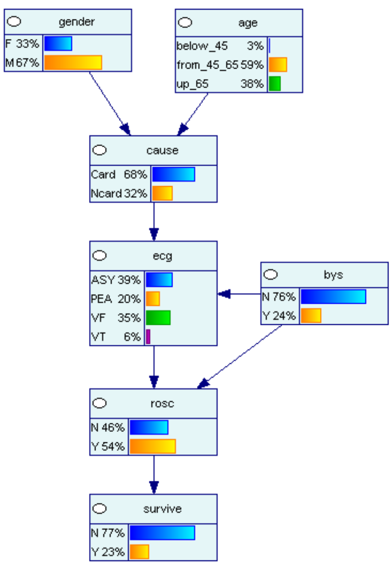

Figure 2: Bayesian network for out-of-hospital cardiac arrest trained from the data with apriori probabilities

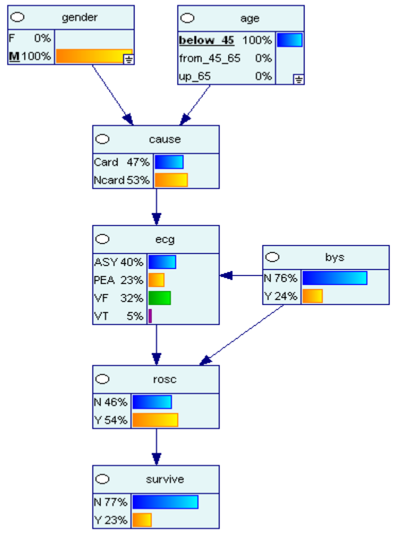

Figure 3: Bayesian network which was recalculated considering the evidence of gender and age

On the figure above the Bayesian network is presented which was recalculated considering the evidence of gender and age as was explained above. We can see that for the male that are under 45 years old the cause of non-cardiac etiology is non-cardiac in 53%. This represent interesting view and knowledge modeled by Bayesian network. Similar others sets of evidences can be adjusted by expert for investigating different scenarios (modeling predictions for scenarios).

7. Results and discussion

In this paper we presented a use of data mining method of Bayesian network, which was defined by expert and later trained on the real data from the study out-of-hospital cardiac arrest. As mentioned in previous section, prediction can be generated by different scenarios which can be inferred from the network by the expert. In this section we present some of the rules that were inferred from Bayesian network and as such used for prediction. The rules are represented below. On the left site of the rule we can see evidences, on the right their is posterior probability recalculated from the scenario modeled by the expert.

IF ECG = VT **AND** BYS = YES **THEN** ROSC = YES (P = 0.96)

IF ECG = ASY **AND** BYS= NO **THEN** ROSC = YES (P = 0.4)

IF ROSC = YES **THEN** SURVIVE = YES (P = 0.42)

IF ROSC = NO **THEN** SURVIVE = YES (P = 0)

IF CAUSE = CARD **AND** BYS = YES **THEN** ROSC = YES (P = 0.78) **AND** ECG = VF (P = 0.46)

We can see from the rules above that prediction for the probability of survival (discharged from hospital) is 0.42 by the evidence of presence of return of spontaneous circulation influences (ROSC). When ROSC is not present the patient unfortunately dies on the field. Those rules can be used for decision support for termination of resuscitation (TOR). In the BN we can see also the probabilities for other parameters as ECG and ROSC are.

Bayesian network is a useful data mining model in the domain of medicine, especially because of its symbolic presentation and ability of model different scenarios based on the real case data. In our study we present a case where Bayesian network was manually generated by consultation of the expert, however it can be also inductively built by machine learning algorithm.

We showed that the use of the network is appropriate and useful in the domain of emergency medicine, where we have intention to develop further applications of the Bayesian network, which will also include inductive building of network by machine learning algorithms.

The Bayesian network as a tool is especially appropriate for the applications on patient simulators in education processes of emergency medicine, where different scenarios for the students at the human patient simulator are prepared by the experts for emergency medicine. Within the help of Bayesian network the real case data can be used and integrated for preparation of different scenarios what are future application, where we will use the model in emergency medicine at Centre for Emergency Medicine in Maribor.

9. References

[1] Hand, H. Mannila, P. Smyth (2001). “Principles of Data Mining”, MIT Press, Cambridge, MA.

[2] Chen, H. and Chau, M (2004). “Web Mining: Machine Learning for Web Applications,”Annual Review of Information Science and Technology, 38, 289-329.

[3] Kononenko, I. (1993). “Inductive and Bayesian Learning in Medical Diagnosis,” Applied Artificial Intelligence, 7, 317-337, 1993.

[4] Vapnik, V. (1998). Statistical Learning Theory, Wiley, Chichester, GB, 1998.

[5] Cohen, P. R. and Feigenbaum, E. A. (1982). The Handbook of Artificial Intelligence: Volume III, Reading, MA: Addison-Wesley.

[6] Quinlan, J. R. (1993). C4.5: Programs for Machine Learning, Los Altos, CA: Morgan Kaufmann.

[7] Fogel, D. B. (1994). “An Introduction to Simulated Evolutionary Optimization,” IEEE Transactions on Neural Networks, 5, 3-14.

[8] Holland, J. H. (1975). Adaptation in Natural and Artificial Systems, Ann Arbor, MI: University of Michigan Press.

[9] O. Pourret, P. Naim and B. Marcot (2008). Bayesian Networks: A Practical Guide to Applications. Chichester, UK: Wiley.

[10] Grmec, Š., et al. (2007), Utstein style analysis of out-of-hospital cardiac arrest – Bystander CPR and end expired carbon dioxide. 72: 404-414.

[11] http://genie.sis.pitt.edu/about.html