Intelligent analysis in predicting outcome of
out-of-hospital cardiac arrest

Miljenko Krizmaric\textsuperscript{a}, Mateja Verlic\textsuperscript{b,*, Gregor Stiglic\textsuperscript{a,b}, Stefek Grmec\textsuperscript{c,d,e}, Peter Kokola\textsuperscript{a,b}}

\textsuperscript{a} Faculty of Health Sciences, University of Maribor, Zitna 15, Maribor, Slovenia
\textsuperscript{b} Faculty of Electrical Engineering and Computer Science, University of Maribor, Smetanova 17, Maribor, Slovenia
\textsuperscript{c} Faculty of Medicine, University of Ljubljana, Vrazov trg 2, Ljubljana, Slovenia
\textsuperscript{d} Faculty of Medicine, University of Maribor, Slomskov trg 15, Maribor, Slovenia
\textsuperscript{e} Centre for Emergency Medicine Maribor, Ulica talcev 9, Maribor, Slovenia

\section*{ARTICLE INFO}

Article history:
Received 22 December 2008
Accepted 21 February 2009

Keywords:
Machine learning
Cardiac arrest
Out-of-hospital resuscitation

\section*{ABSTRACT}

The prognosis among patients who suffer out-of-hospital cardiac arrest is poor. Higher survival rates have been observed only in patients with ventricular fibrillation who were fortunate enough to have basic and advanced life support initiated early after cardiac arrest. The ability to predict outcomes of cardiac arrest would be useful for resuscitation chains. Levels of EtCO\textsubscript{2} in expired air from lungs during cardiopulmonary resuscitation may serve as a non-invasive predictor of successful resuscitation and survival from cardiac arrest. Six different supervised learning classification techniques were used and evaluated. It has been shown that machine learning methods can provide an efficient way to detect important prognostic factors upon which further emergency unit actions are based.

\section*{1. Introduction}

Sudden out-of-hospital death is among the most frequent contributors to the mortality in many countries. Often it has no premonitory symptoms and it may be unwitnessed. Because it involves cardiac arrest, such deaths are usually attributed to the cardiovascular diseases. Out-of-hospital cardiac arrest (OHCA) is a cessation of cardiac mechanical activity as confirmed by the absence of signs of circulation \cite{1} in out-of-hospital settings. OHCA is a significant health issue in many countries, because the mortality from cardiovascular diseases (CVD) is among highest in many countries and unfortunately, Slovenia is not an exception.

The mortality from CVD in Slovenia is among the highest in Eur-A group of countries \cite{2}. World Health Organization - Regional Office for Europe reports that large excess mortality in Slovenia is due to diseases of the pulmonary circulation and other heart diseases. Slovenia is also above Eur-A average for CVD resulting in death \cite{2}. The good news is that cardiac arrest is potentially reversible if treated properly and early enough, but the bad news is that the prognosis among patients who suffer from OHCA is relatively poor and that increase in survival rate is a necessity \cite{3–6}.

Successful resuscitation depends on the effectiveness of the chain of survival (Fig. 1), which is a globally endorsed response model developed by the American Heart Association \cite{7,8} and also supported by European Guidelines for Resuscitation \cite{9}. The links in the chain are: (1) early access to care, (2) early cardiopulmonary resuscitation (CPR), (3) early defibrillation, and (4) early advanced cardiac life support (ACLS).

The first link in the chain (1) depends heavily on the presence of witnesses during cardiac arrest. Witnesses should request emergency assistance immediately and without hesitation. In witnessed cases of cardiac arrest Emergency Medical
ARTICLE IN PRESS

2 COMPUTER METHODS AND PROGRAMS IN BIOMEDICINE XXX (2009) XXX-XXX

Services (EMS) may arrive sooner in comparison to the unwitnessed cardiac arrests. In second link (2), CPR should be maintained until EMS personnel arrive. Early CPR can help sustain life during ventricular fibrillation (VF) of the heart. Rescue breathing (e.g., mouth-to-mouth, breathing bag) and chest compressions can help oxygenated blood to flow to the person’s brain and heart, until defibrillation attempts occur, which may then restore normal heart pumping. In third link (3) defibrillation can be performed by trained layperson, if an automated external defibrillator (AED) is available; otherwise patient has to wait for EMS personnel to arrive. In some cases untrained bystander can be guided by machine itself or via telephone until EMS arrival. The last link (4) is in the domain of EMS personnel and it includes administration of cardiac drugs, endotracheal intubation, defibrillation and attachment of monitoring devices.

During emergency intervention several prognostic factors are observed as recommended in Utstein style definitions, reporting templates and guidelines [1]. These reporting templates include factors, which cover all four links in the chain of survival. For example, parameters include information whether cardiac arrest was witnessed and how long it took for emergency unit to arrive (first link), if bystander CPR was performed (second link), whether defibrillation attempts have been made (third link), and whether a patient needed assistance by ventilation (fourth link).

In this study we analyzed the impact of several prognostic factors on return of spontaneous circulation (ROSC) and the outcome of the resuscitation (whether patients survived or not until discharge from hospital) using several machine learning methods and techniques.

Related study was already done on a smaller dataset using decision trees [10,11] to support the results obtained from statistical analysis of the dataset. However, in this study we approached the analysis from the machine learning perspective first to discover possible relations between prognostic factors, which can be then statistically analyzed and either supported or rejected. Furthermore, we used several machine learning techniques to obtain best results possible.

Our subgoal was also to investigate the role of cardiac drug Vasopressin in predicting ROSC and survival of patients with OHCA to see whether our results comply with the research findings from other countries.

This paper is organized as follows: Section 2 introduces prolonged study of out-of-hospital cardiac arrest and the prognostic factors included in the analysis. In Section 3 methods and techniques used in intelligent analysis are briefly described. Results and discussion are presented in Section 4 and Section 5 concludes the paper.

2. Prolonged study of out-of-hospital cardiac arrest

In our analysis we used the dataset collected during observational study in Maribor, Slovenia, in the period from January 2001 to December 2005. In our previous work [10] only the patients from 2001 to 2004 were included in the analysis. Dataset includes all cases classified as OHCA and dispatched to prehospital emergency unit. Data was collected according to the Utstein criteria and guidelines [4]. The study population was composed of adults over 18 years with OHCA. In a
Fig. 2 – Percentage of patients with cardiac and non-cardiac aetiology.

Fig. 3 – Graph of resuscitation success for male patients in our study.

Fig. 4 – Graph of resuscitation success for female patients in our study.

Table 1 – Important characteristics of the study population.

<table>
<thead>
<tr>
<th></th>
<th>Died in the field ROSC was not obtained</th>
<th>Survived to discharge from hospital</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sex (M/F)</td>
<td>150/44</td>
<td>111/59</td>
</tr>
<tr>
<td>Age (years)</td>
<td>60.77 ± 13.27</td>
<td>59.55 ± 12.05</td>
</tr>
<tr>
<td>Witnessed arrest (Y/N)</td>
<td>138/56</td>
<td>160/10</td>
</tr>
<tr>
<td>Bystander CPR (Y/N)</td>
<td>24/170</td>
<td>88/82</td>
</tr>
<tr>
<td>Initial etCO2 (kPa)</td>
<td>2.18 ± 2.28</td>
<td>3.094 ± 1.413</td>
</tr>
<tr>
<td>Final etCO2 (kPa)</td>
<td>0.995 ± 0.344</td>
<td>3.913 ± 1.126</td>
</tr>
<tr>
<td>Arrival time (min)</td>
<td>10.93 ± 4.33</td>
<td>5.92 ± 2.951</td>
</tr>
</tbody>
</table>

2.1. Prognostic factors

Before describing the potential prognostic factors for predicting successful or unsuccessful resuscitation the condition of cardiac arrest has to be briefly described. A layperson usually falsely considers that heart attack and cardiac arrest are terms for the same condition, although these two conditions may differ in various aspects and may have very different consequences. However, it is true that a person suffering a heart attack is more likely to develop abnormal heart rhythms and sudden cardiac arrest. Heart attack is a result of blocked blood flow to the heart muscle. In most cases it starts slowly and is accompanied by several warning signs, like discomfort in chest and other areas of the upper body, or maybe even nausea, sweating and shortness of breath. In comparison to heart attack a cardiac arrest happens immediately and without warning signs. It is defined as the cessation of cardiac mechanical activity confirmed by the absence of signs of circulation [4]. The heart beating stops, the victim collapses due to the lack of consciousness and the breathing also stops. If a victim of cardiac arrest is not provided with appropriate medical assistance soon after the cardiac arrest victim dies.

2.1.1. Age and sex

Several cardiovascular diseases are related to aging and according to World Health Organization’s report on Slovenia mortality due to cardiovascular causes starts to increase for both sexes in their forties. Because WHO’s report shows differences in mortality for males and females, we can assume that sex may also be an important risk factor for cardiac arrest and resuscitation interventions.

2.1.2. ECG (heart rhythm)

Electrocardiogram (ECG) refers to the first monitored heart rhythm after cardiac arrest. The rhythm is usually analyzed by person interpreting results from monitor or defibrillation device. If the defibrillation attempt was needed to restore heart rhythm this rhythm is defined as shockable; otherwise it is considered non-shockable. Shockable rhythms are further divided into ventricular fibrillation (VF) and pulseless ventricular tachycardia (VT), while non-shockable ones are asystole rhythm and pulseless electrical activity (PEA). The principal difference in the management of these two

prolonged study the resuscitation was attempted in a group of 477 patients, among which 69.39% were male and 30.61% female patients. The percentage of patients with presumed cardiac aetiology versus non-cardiac is illustrated in Fig. 2. ROSC was obtained in 283 (59.33%) cases, while 170 (35.64%) patients survived until discharge from hospital.

Fig. 3 illustrates the success of resuscitation attempts for male patients and Fig. 4 shows the success of resuscitation attempts for female patients. Table 1 summarizes other important characteristics of the study population.
groups of arrhythmias is the need for attempted defibrillation in those patients with VF/VT. Subsequent actions, including chest compressions, airway management, ventilation, venous access, administration of adrenaline, and the identification and correction of reversible factors, are common to both groups [3].

2.1.3. Witnessed (cardiac arrest)
As our intuition suggests witnessing cardiac arrest is considered as the most important moment in the chain of survival, because all further actions in the resuscitation depend on it. If a witness is present at the moment of victim’s collapse prompt action without a minute of hesitation can save the life of the victim. Usually the witness is a layperson, coincidental bystander; in rare cases a witness is also a medically trained person.

2.1.4. Bystander CPR
Bystander CPR is an attempt to restore victim’s spontaneous circulation by a combination of rescue breathing and chest compressions. Bystander CPR is CPR performed by a layperson or by a medically trained person, who is not a part of an organized emergency response system involved in resuscitation of a victim. Early CPR is important because it can prevent severe brain damages due to the lack of blood flow to the brain.

2.1.5. Cardiac and non-cardiac etiology
The etiology of a cardiac arrest is another potential factor for predicting the success of resuscitation. An arrest is presumed to be of cardiac etiology unless it is known or likely to have been caused by trauma, submersion, drug overdose, asphyxia, exsanguination, or any other non-cardiac cause as best determined by rescuers.

2.2. Ventilation
If an OHCA victim does not breathe assistance by ventilation may be needed. Patient’s lungs can be inflated by rescue breathing with or without special bagmask or any other device.

2.3. Measuring end-tidal CO₂
The significance of capnometry monitoring during CPR was first noted by Kalenda [12]. Kalenda showed the importance of monitoring pulmonary perfusion by means of the capnogram as a continuous guide to the cardiac output achieved by cardiac massage and resuscitation. During CPR the partial pressure of end-tidal carbon dioxide (PetCO₂) correlates with cardiac output and, consequently we can assume, it has a prognostic value in CPR. End-tidal CO₂ monitoring has potential as a non-invasive indicator of cardiac output during resuscitation and can serve as a prognostic indicator for resuscitation.

2.4. Vasopressin
During advanced cardiac life support different cardiovascular drugs can be given to the OHCA victim to regulate heart rhythm. Vasopressin is one of the drugs which have been used in cases of OHCA in Europe for some years now. Recent studies report quite substantial increase in survival to hospital discharge in patients with OHCA and vasopressin supremacy over epinephrine in the treatment of asystole heart rhythm [13,14].

2.5. Time of arrival
Potential prognostic factor time of arrival is time needed for emergency medical service unit to arrive to victim measured in minutes. Because human brain can sustain severe damages if the blood supply is interrupted for more than several minutes, every minute counts in the struggle for successful resuscitation of the OHCA patient. Time of arrival can be critical prognostic factor, especially in sub-urban or remote areas without their own EMS unit. In such cases EMS unit has to arrive from nearest hospital which can be quite far away.

2.6. Defibrillation
Electrical defibrillation is a passage of current at high voltage to attempt to restore a normal regular rhythm. In most cases it is attempted by trained medical personnel by means of automated (semi-automated) or manual external defibrillator, but recently fully automated external defibrillators have been installed in many larger public places, like airports and shopping centers. These automated external defibrillators can guide a layperson through whole procedure of defibrillation. In the Utstein report only the presence of defibrillation attempt is considered, but not the type of the defibrillation device. Please note that in our study in cases where patients initially displayed non-shockable hearth rhythm but which later changed to shockable, defibrillation attempts in dataset are indicated with label Y-P.

2.7. ROSC and Survive
The first step in a resuscitation attempt is to return spontaneous circulation (ROSC) and then long-term survival of the patient or at least to the discharge from the hospital. If during resuscitation attempt OHCA victim starts to breathe, and it should not be just few occasional gasps, cough, or move, a bystander can assume the spontaneous circulation is returned, but this does not necessarily imply the victim will survive. On the other hand, if spontaneous circulation is not returned at any time, the result of resuscitation attempts is known—victim dies. Medical personnel may consider additional signs of ROCS such as evidence of palpable pulse or measurable blood pressure. Assisted circulation should not be considered as ROSC until spontaneous circulation is restored. By consensus and for the purposes of the Utstein report template, the phrase “any ROSC” represents a brief (more than 30 s) restoration of spontaneous circulation that provides evidence of more than an occasional gasp, occasional fleeting palpable pulse, or arterial waveform [1].

3. Intelligent analysis
Machine learning approach to data analysis can uncover relationships between potential prognostic factors, which can be further analyzed by different statistical tests for significance.
Among a variety of machine learning methods and techniques six different classification methods were chosen and applied. All of them use supervised learning to build a classification model. In all cases n-fold cross-validation was used to partition the dataset and evaluate methods.

3.1. Cross-validation (CV)

In supervised learning data is separated into two different sets—a learning set and a testing set. Learning set is used for building a model, while testing set is used to evaluate the model. In some cases a problem of overfitting can occur—model fits data too well and it is not general enough anymore. In this case additional measures have to be taken. One of the solutions is the use of a resampling technique called n-fold cross-validation (CV) [15]. In n-fold cross-validation data is split into n non-overlapping subsets of equal or approximately equal size. Afterwards the model is trained n times and each time one of the subsets is left out from training to be used as a test set and to estimate prediction error. Final error of CV is given by the average of the n estimates of prediction error. CV is especially useful for small datasets to estimate the generalization error of the model, because we can use entire training data to learn a classifier and to evaluate its performance. The decision of how many folds are needed depends on the dataset. 10 folds are a common choice for n-fold cross-validation, especially if the dataset is not too large or too sparse. 10-fold cross-validation was used in all of the experiments in this study.

3.2. Methodology

Six different machine learning techniques were used during intelligent analysis of the data: Decision trees, k-nearest neighbors, Naïve Bayes, Neural networks, Support Vector Machine and Random forests. Each of the techniques will be briefly described.

3.2.1. Decision trees

Decision trees are among the most popular classifiers, especially in analyzing data from medical domain, probably mostly due to their interpretation power. A decision tree is hierarchical structure of internal nodes, branches and external nodes (or leaves) classifying data into known classes. It is induced from iterative splitting of the data from the training set according to predefined splitting criteria. Most often the splitting criteria are information gain and information gain ratio implemented in ID3, C4.5 and C5.0 algorithms. In our experimentation J48 [16], decision trees in Weka toolkit [17] were used.

3.2.2. k-nearest neighbors

The k-nearest neighbor method is one of the statistical machine learning methods [18,19]. It is based on a relatively simple idea of assigning each instance a label according to the majority class of k items in the neighborhood. Because it is one of the most commonly used classification methods, we used it as a reference. Weka toolkit implementation of k-nearest neighbors is an Instance-Based (IBk) learner with fixed neighborhood [20].

3.2.3. Naïve Bayes

The Naïve Bayes classifier [21] is the simplest of probabilistic classifiers applying Bayesian theorem. It is based on independent feature model and it requires only a small amount of data for training phase. Its success and stability in the medical domain was the main reason for selecting and using it in our study. Weka implementation of Naïve Bayes is using numeric estimator precision values based on the analysis of data [22].

3.2.4. Neural networks

Neural networks [23,24] belong to well known and frequently used classification models. They are formed by a large number of interconnected nodes. Many layers of relatively simple nodes can effectively capture much more complex relationships between input and output. Nodes form layers and commonly neural network consists of input layer, one or more hidden layers and output layer. Nodes in input layer correspond to predictor variables and they are connected to every node in a hidden layer. Nodes in hidden layers can be connected to nodes in other hidden layers or to output layer, which correspond to one or more response variables. Neural network can learn through changing weights of the connections. Multilayer perceptron (MLP), an implementation of neural networks in Weka, uses backpropagation to classify instances.

3.2.5. Support Vector Machine (SVM)

Support Vector Machine (SVM) [25] is a learning algorithm widely used in data mining and bioinformatics. SVM belongs to the group of generalized linear classifiers. The goal of SVM is to construct a hyperplane based on the support vectors nearest to the hyperplane, where support vectors are sets of transformed attributes. A SVM analysis finds the hyperplane, oriented in such a way, that the margin between support vectors in maximized. We used Weka implementation of sequential minimal optimization (SMO) [26,27] which is a fast method to train SVMs.

3.2.6. Random forests

Recently, ensemble approach to classification has gained on popularity. Combining classifiers into ensembles can diminish some of the disadvantages of individual classifier and thus improve classification accuracy. Random forests (RF) is an advanced ensemble classification technique [28] based on bagging [29] and enhanced by a combination of bootstrapping and random feature selection for decision tree building. Random decision trees created this way are grown by selecting the feature to split on at each node from randomly selected number of nodes. RF is a robust technique, it works well even for noisy data and is considered as one of the most competitive and robust methods that can be compared to bagging or boosting [30].

3.3. Experimental settings

Because of the logical correlation between ROSC and survival, namely if spontaneous circulation is not returned patient does not survive, it was necessary to limit its effect on the analysis. Therefore we decided to apply machine learning methods to dataset without survive attribute for prediction of ROSC.
Table 2 – Subsets of original dataset used in ROSC analysis.

<table>
<thead>
<tr>
<th>Dataset subset</th>
<th>Description (cases included)</th>
</tr>
</thead>
<tbody>
<tr>
<td>All</td>
<td>All patients</td>
</tr>
<tr>
<td>Cardio</td>
<td>Cases of patients with presumed cardiac etiology only</td>
</tr>
<tr>
<td>Non-cardio</td>
<td>Cases of patients with presumed non-cardiac etiology only</td>
</tr>
<tr>
<td>ECG_cat</td>
<td>All cases but with only two categories for heart rhythm</td>
</tr>
<tr>
<td>Non-shock</td>
<td>Cases of patients with non-shockable ECG only</td>
</tr>
<tr>
<td>Shock</td>
<td>Cases of patients with shockable ECG only</td>
</tr>
<tr>
<td>Female</td>
<td>Cases of female patients only</td>
</tr>
<tr>
<td>Male</td>
<td>Cases of male patients only</td>
</tr>
</tbody>
</table>

Table 3 – Subsets of original dataset used in Survive analysis.

<table>
<thead>
<tr>
<th>Dataset subset</th>
<th>Description (cases included)</th>
</tr>
</thead>
<tbody>
<tr>
<td>All</td>
<td>All patients</td>
</tr>
<tr>
<td>ECG_cat</td>
<td>All cases but with only two categories (instead of four) for heart rhythm</td>
</tr>
<tr>
<td>Non-shock</td>
<td>Cases of patients with non-shockable ECG only</td>
</tr>
<tr>
<td>Shock</td>
<td>Cases of patients with shockable ECG only</td>
</tr>
<tr>
<td>ROSC_Y_all</td>
<td>All patients with ROSC</td>
</tr>
<tr>
<td>ROSC_Y_cardio</td>
<td>All patients with ROSC and presumed cardiac etiology</td>
</tr>
<tr>
<td>ROSC_Y_non-cardio</td>
<td>All patients with ROSC and presumed non-cardiac etiology</td>
</tr>
<tr>
<td>ROSC_Y_female</td>
<td>All female patients with ROSC</td>
</tr>
<tr>
<td>ROSC_Y_male</td>
<td>All male patients with ROSC</td>
</tr>
</tbody>
</table>

Results of intelligent analysis are described in two following subsections. First subsection includes results from ROSC analysis and second subsection includes results from Survive analysis.

4. Results and discussion

4.1. Results of ROSC analysis

ROSC analysis focused on the importance of attribute to the return of spontaneous circulation. To exclude logical relation between ROSC and survival from OHCA the attribute Survival was omitted from the analysis. We included the most interesting decision trees to illustrate the importance of different prognostic factors for different data subsets.

Fig. 5 shows a decision tree build from records of all patients with OHCA. In this case classification accuracy of 86.58% was obtained by using 10-fold CV. By observing top three nodes, we can assume that initial end-tidal CO2 defibrillation attempt and the drug vasopressin play an important role in ROSC. In cases where initial value of end-tidal CO2 was below 1.3 kPa, the spontaneous circulation did not return and OHCA victims died. In other cases, when defibrillation was attempted (Y, Y-P), patients had greater chance to display signs of ROSC.

As it can be observed in Fig. 6, initial value of end-tidal CO2 is especially important for patients with shockable heart rhythm. If the initial value of end-tidal CO2 was above 1.3 kPa most of the patients regained signs of ROSC. However, we have...
to consider the fact that this type of decision tree is prone to overfitting; decision tree in Fig. 6 can also represent a typical case of overfitting because it contains single attribute.

In Fig. 7, decision tree was build from records of patients with presumed cardiac etiology. In this case we obtained 10-fold CV classification accuracy of 90.44%. Similar to previous decision trees of other subsets, the value of initial end-tidal CO2 is in the root node. Other important factors higher in the tree hierarchy are defibrillation attempts, bystander CPR and witnessed arrest, which are also important links in the chain

Fig. 7 – Decision tree for patients with cardiac etiology, with minimum of 3 objects in a leaf and with 10-fold CV classification accuracy of 90.44%.

Fig. 8 – Decision tree for male patients with minimum 7 objects in a leaf and with 10-fold CV classification accuracy of 87.62%.

Fig. 9 – Decision tree for female patients with minimum 7 objects in a leaf and with 10-fold CV classification accuracy of 86.99%.
of survival illustrated in Fig. 1. We can also notice possible differences between male and female patients if they were not given Vasopressin.

Figs. 8 and 9 show decision trees for male and female patients respectively. For both subsets it seems that initial end-tidal CO2 is the most important attribute. It can be observed that defibrillation attempts are more important for male patients than for female patients, while etiology is relatively important in the decision trees for both subsets.

Furthermore, for female patients it can be observed, that if initial end-tidal CO2 was above 1.3 kPa, most of the patients regained signs of ROSC, while in male patients return of ROSC signs was also dependant on defibrillation attempts.

Results of classification accuracy comparison of different classification methods are illustrated in Fig. 10. Rotation forests outperformed other classification methods in every subset and on the whole dataset in the prediction of ROSC.

Unfortunately, it is not possible to interpret classification results from rotation forests in such intuitive way as it is possible for decision trees and therefore we cannot extract prognostic factors that contributed most to this classification. Fortunately enough, in all subsets but one (non-shockable rhythms) decision trees still perform with high classification accuracy and better than other classification methods with the exception of rotation forests and therefore can be used to extract interesting relations between prognostic factors.

### 4.2. Results of Survive analysis

In study by Kuisma and Määttä [31] from Helsinki on importance of early access to out-of-hospital patients early access...
Fig. 12 – Decision tree built using a group of patients with non-shockable heart rhythm (10-fold CV classification accuracy: 83.4906%).

has turned out to be the weakest link in the chain of survival and it should receive major attention in the near future. Our study shows similar results with some additional interesting observations.

The following results show decision trees where attribute describing survival (Y, N) was used as decision class. Collected data was analyzed by dividing patients into two groups. First group included patients with shockable heart rhythm (Fig. 11), while other group consisted of patients with non-shockable rhythm (Fig. 12). It can be seen that in both groups there are two significant attributes, i.e., Bystander CPR and time of arrival. On the other hand the first decision tree (Fig. 11) shows another important attribute, i.e., VAZO which was placed in the topmost node of the decision tree. This means that VAZO plays a significant role in respect to survival in a group of patients with non-shockable heart rhythm. Another interesting interaction between VAZO and time of arrival attributes is shown in Fig. 13 where a group of female patients with ROSC was analyzed. In a very abstract interpretation of this tree it can be said that using VAZO, allows longer arrival times.

Fig. 13 – Decision tree for a group of female patients with ROSC (10-fold CV accuracy: 81.3725%).

Fig. 14 – Classification accuracy comparison of different classification methods used for subsets in analysis using Survive as decision class).
outcome of survival. All tests were done using 10-fold cross-validation using default Weka settings. In contrast to results of classification comparison results from ROSC analysis, rotation forests still performed well but did not outperform other methods. Unfortunately, the accuracy of decision trees for subsets was not as high as in ROSC analysis, i.e., in most subsets method of k-nearest neighbors (IBk), naive Bayes, SVM (SMO), and neural networks (MLP) classified better. However, decision trees interestingly performed better on the whole dataset and this can be probably attributed to the presence of attribute ROSC in the dataset—if ROSC did not return, OHCA victim did not survive.

5. Conclusion and future plans

This paper demonstrates an extensive study using machine learning techniques for prediction of ROSC and survival of victims with cardiac arrest in the out-of-hospital settings and it gives some possible directions for further investigation with classical statistical methods. Relatively high classification accuracies suggest that arrival time, witnessed arrest, bystander CPR, initial EtCO2 and final EtCO2 are positively related with ROSC and with survival on admission and they can serve as prognostic factors for predicting the outcome. Results of ROSC analysis show that decision trees still represent one of the most appropriate ways of machine learning analysis due to their power of interpretability.

We plan to collect even more OHCA data and analyze it with pure statistical approach and also with intelligent analysis approach to discover important factors that would increase the rate of survival. Moreover, it would be also very interesting to further explore gender differences in observed prognostic factors, for example, the effect of vasopressin on male and female victims of OHCA. Every option that increases the chances of survival is worth exploring.

Conflict of interest

The authors declare they have no conflicts of interest regarding this study.

References


