Prediction of short-term success of electrical cardioversion

Lazar Vasović[1*] and Ana Jakovljević[1]

[1] Faculty of Mathematics, University of Belgrade, Studentski trg 16,
11000 Belgrade, Serbia
mi16099@alas.matf.bg.ac.rs

Abstract

Electrical cardioversion is a medical technique that uses synchronized electrical shocks to restore normal heart rhythm in people with persistent arrhythmia. This kind of heart rate problem is usually associated with a disease called atrial fibrillation.

The aim of this paper was to create a classification model that accurately predicts whether the procedure will be successful in the short term (immediate outcome), based on data on the clinical picture, other indications, and drug therapy prescribed to patients undergoing it. Dataset, consisting of 147 unique instances, was obtained from the Pacemaker Center of the Clinical Center of Serbia and pertains to patients with electrical cardioversion performed from 2014 to 2019. It is noticeably imbalanced with concern to target class; Successful procedures were marked as class True (130 – 88.4%) and unsuccessful as class False (17 – 11.6%).

The focus was on Bayesian networks, a well-known probabilistic graphical model. They, however, proved inferior to other methods of classification and machine learning, such as the random forest classifier and artificial neural networks. Fitted models were compared by their accuracy, sensitivity (recall), specificity, and other relevant metrics. Extra attention was given to data preprocessing and exploratory analysis, as well as predictor (feature) importance.

Experiments included one Bayesian network structure with fair predictive values and some other similarly successful models. A voting ensemble made of multilayer perceptron (sklearn MLPClassifier) and complement naïve Bayes (sklearn ComplementNB) turned out the best, with full 100% specificity, recall of the important class False, while maintaining a relatively high F_1 score of 63% on the same class. Accuracy was 87%, while balanced accuracy was 92%. Precision on class False was not that good 46%, but it was still the best compromise.

Results also give insight into predictor importance, such as that extracted from the decision tree classifier, which marked patient age, heart rate, and total duration of the indicated heart disease as the most significant ones. Other methods of statistical analysis and machine learning were also used to identify important features, including a dendrogram generated by clustering attributes based on their correlations as the similarity measure.

Keywords: cardioversion, atrial fibrillation (AF), voting ensemble, Bayesian network