PREDICTING THE SUCCESS OF ANTEGRADE CHRONIC TOTAL OCCLUSION RECANALIZATION USING MACHINE LEARNING

M.S. Abramovich[1], V.V. Kozlovsky[1], V.I. Stelmashok[2], O.L. Polonetsky[2]
[1]*Research Institute for Applied Problems of Mathematics and Informatics*
[2]*Republican Scientific and Practical Centre of Cardiology*
Minsk, BELARUS
e-mail: `Abramovichms@bsu.by`

Abstract

The problem of predicting the success of antegrade coronary arteries chronic total occlusion recanalization, based on X-ray anatomical and clinical markers is considered. The results of comparative analysis of the prediction accuracy based on various machine learning algorithms are presented.

Keywords: data science, machine learning, recanalization

1 Introduction

Chronic total occlusion (CTO) of the coronary arteries is a frequently detectable type of coronary lesion in patients with ischemic heart disease [1]. Successful restoration of antegrade blood flow by interventional techniques allows to improve the quality of life [2]. However, the indications for interventional CTO recanalization are based on the clinical signs, without taking into account the technical complexity of this procedure and the risk of possible failure of percutaneous coronary intervention.

The most important stage of CTO correction by interventional methods is the CTO crossing by coronary wire; majority of unsuccessfully performed operations are due to the inability to perform this manipulation. Minimization the frequency of unsuccessful CTO crossing by coronary wire will optimize the treatment quality in this group of patients as well as reduce the risk of possible complications.

Analysis of the literature has shown the absence of specific prognostic scales to predict the success of antergrade CTO crossing by coronary wire. This fact as well as the significance of the identified problem mean that it is critically important to develop a system for predicting the success of antegrade coronary arteries CTO recanalization.

Various methods of machine learning such as logistic regression, decision tree and random forest to predicting the success of antegrade chronic total occlusion recanalization are used. Optimal parameters for methods were found and the maximal accuracy was obtained on these parameters.

2 Materials and methods

The present study was retrospective, single-center, and included data from 395 patients whom attempts of coronary arteries CTO recanalization were performed during the period from 2009 to 2018. Based on the success of the coronary wire passage through the CTO zone all of the above-mentioned patients were divided into the 2 groups: patients with successful (n_1=292) and failed (n_2=103) CTO recanalization.

77 clinical and X-ray anatomical markers are analysed.

First, the sample is divided: train data with 264 patients and test data with 131 patients. Then logistic regression with $l1$ regularization is trained on test data.

Feature selection was made using LASSO method [3]. Importance features are ones with nonzero weights. The process repeats until the number of features decreases.

Various methods of machine learning such as logistic regression, decision tree and random forest to predicting the success of antegrade chronic total occlusion recanalization are used.

Selection of hyperparameters was made on a grid using cross validation. For evaluating each set of hyperparameters train data is divided into 4 folds. Then model trains on 3 folds and the last fold is used for testing. The process repeats 4 times; finally, each part is used for testing. As a result, we got performance evaluation for given set of hyperparameters with uniform using of data.

In table 1 grid parameters of methods are presented.

Table 1: Grid parameters of methods

Method	Parameter	Values	Description
Logistic regression	penalty	’l2’, ’l1’	used to specify regularization
	C	0.8, 0.9, 1, 1.1, 1.2, 1.3, 1.4	inverse of regularization strength
Decision tree	max_depth	2, 4, . . . , 18	the maximum depth of the tree
	criterion	’gini’, ’entropy’	function to measure the quality of split
	min_samples_split	2, 4, 6, 8	the minimum number of samples to split inner node
	mmin_samples_leaf	2, 4, 6, 8	the minimum number of samples in leaf node
Random forest	n_estimators	10, 60, 110, 160, 210	number of trees in the forest
	max_depth	2, 3, 4, 5	the maximum depth of each tree
	criterion	‘gini’, ’entropy’	function to measure the quality of split
	min_samples_split	2, 3, 4	the minimum number of samples to split inner node
	min_samples_leaf	2, 3, 4	the minimum number of samples in leaf node

3 Results

In table 2 importance features for classification of patients with successful and failed CTO recanalization are presented. Features are ranked on a 0 – 100 scale in terms of their potential importance.

Table 2: Features importance

Feature	Importance
What coronary artery segment?	100
Blunt CTO stump	71
Soft tapered tip wires as first choice	61
Level of CTO complexity	48
Middle stiffness non-tapered tip wires as a first choice	44
Wire toughening and support	41
Number of wires used for CTO recanalization	35
Is the tortuosity in the CTO zone?	27
If yes, number of side branches	21

In table 3 results of the grid search for machine learning methods and the accuracy patients classification with successful and failed CTO recanalization on train sample are presented.

Table 3: Accuracy patients with successful and failed CTO recanalization

Method	Best parameters	Results on cross validation
Logistic regression	penalty: 'l2' C: 0.8	0.7803
Decision tree	criterion: 'entropy' max_depth: 4 min_samples_leaf: 6 min_samples_split: 2	0.7771
Random forest	n_estimators: 60 criterion: 'gini' max_depth: 2 min_samples_leaf: 3 min_samples_split: 3	0.7924

In figure 1 ROC curves and AUC values for logistic regression, decision tree and random forest on test sample are presented. Logistic regression shows the best result on the test sample.

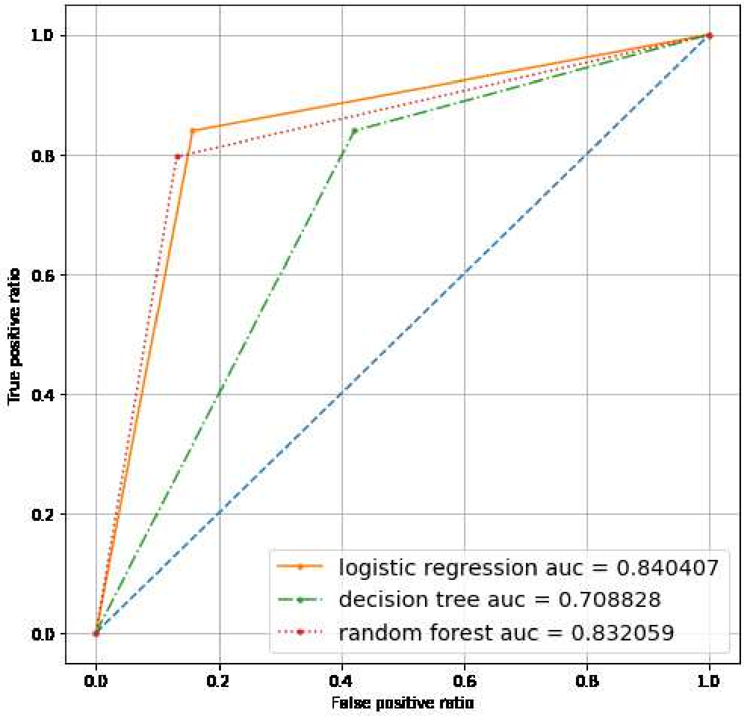

Figure 1: ROC curves

References

[1] Di Mario C., Werner G., Sianos G., Galassi A., Büttner J., Dudek D., Chevalier B., Lefevre T., Schofer J., Koolen J., Sievert H., Reimers B., Fajadet J., Colombo A., Gershlick A., Serruys P., Reifart N. (2007). Perspective in the recanalization of Chronic Total Occlusions (CTO): consensus document from the EuroCTO Club. *EuroIntervention*. Vol. **3(1)**, pp. 33-43.

[2] Werner G., Martin-Yuste V., Hildick-Smith D., Boudou N., Sianos G., Gelev V., Rumoroso J.R., Erglis A., Christiansen E.H., Escaned J., di Mario C., Hovasse T., Teruel L., Bufe A., Lauer B., Bogaerts K., Goicolea J., Spratt J.C., Gershlick A.H., Galassi A.R., Louvard Y. (2018). A randomized multicentre trial to compare revascularization with optimal medical therapy for the treatment of chronic total coronary occlusions. *Eur Heart J*. Vol. **39(26)**, pp. 2484-2493.

[3] Harrington P. (2012) *Machine Learning in Action*. Manning. New York.