

УДК 519.24:612.086.2

ПРОГНОЗИРОВАНИЕ ИСХОДА РЕНТГЕНЭНДОВАСКУЛЯРНЫХ ОПЕРАЦИЙ НА КРОНАРНЫХ АРТЕРИЯХ



В.В. Козловский
Магистрант БГУ



М.С. Абрамович
Заведующий НИЛ статистического анализа и моделирования НИИ прикладных проблем математики и информатики, кандидат физико-математических наук, доцент



В.И. Стельмашок
Заведующий отделом интервенционной кардиологии ГУ «РНПЦ «Кардиология», доктор медицинских наук, доцент

Белорусский государственный университет, Республика Беларусь.
Научно-исследовательский институт прикладных проблем математики и информатики Белорусского государственного университета, Республика Беларусь.
ГУ «РНПЦ «Кардиология», Республика Беларусь.
E-mail: fpm.kozlovskVV2@bsu.by.

В. В. Козловский

Окончил Белорусский государственный университет. Магистрант БГУ. Работает в должности инженера-программиста. Проводит научные исследования в области применения методов машинного обучения в эндоваскулярной кардиологии.

М. С. Абрамович

Заведующий НИЛ статистического анализа и моделирования НИИ прикладных проблем математики и информатики БГУ. Область научных интересов – методы и алгоритмы машинного обучения для диагностики заболеваний.

В. И. Стельмашок

Заведующий отделом интервенционной кардиологии ГУ «РНПЦ «Кардиология». Область научных интересов – разработка и внедрение новых способов рентгенэндоваскулярного лечения пациентов, страдающих сердечно-сосудистыми заболеваниями.

Аннотация. На обучающей выборке, состоящей из групп пациентов успешно прооперированных и прооперированных с перфорациями коронарных артерий пациентов, из всех клиничко-ангиографических признаков с использованием критерия хи-квадрат отобрано 13 информативных. Так как классы обучающей выборки несбалансированы, то для наблюдений класса обучающей выборки меньшего объема, назначались веса, которые определяли потери от неправильной классификации. В качестве методов классификации использовались логистическая регрессия, метод опорных векторов, деревья решений и случайный лес. Найдены оптимальные гиперпараметры для рассмотренных моделей. ROC-анализ показал достаточно высокую точность классификации пациентов.

Ключевые слова: Коронарные артерии, хронические тотальные окклюзии, перфорация стенки коронарного сосуда, обучающая выборка, несбалансированность классов, информативные признаки, логистическая регрессия, метод опорных векторов, случайный лес.

Введение.

Хорошо известно, что успешное восстановление кровотока в хронически окклюзированной коронарной артерии рентгенэндоваскулярными способами позволяет улучшить качество жизни и прогноз у оперируемых пациентов [1]. Вместе с тем проведение подобных вмешательств нередко ассоциируется с развитием осложнений, наиболее тяжелым из которых является перфорация стенки коронарного сосуда.

С целью оптимизации качества оказания медицинской помощи пациентам с хроническими окклюзионными поражениями коронарных артерий представляется целесообразным разработать систему прогнозирования риска развития перфораций в ходе выполнения оперативного вмешательства, что позволит предотвратить данные осложнения, а также снизить финансовые затраты на лечение вышеобозначенных лиц.

Построение решающих правил классификация успешно прооперированных и прооперированных с перфорациями коронарных артерий пациентов проводилось с использованием методов машинного обучения.

Материалы и методы.

Исследование являлось ретроспективным, одноцентровым и включало данные о 369 пациентах, у которых успешно проведены рентгенэндоваскулярными операции и 29 пациентах, у которых отмечалось развитие перфораций коронарных артерий. Необходимо отметить, что классы обучающей выборки имеют существенно различный объем (несбалансированны).

Из всех клинико-ангиографических признаков у пациентов, включенных в обучающую выборку, с использованием критерия хи-квадрат отобрано 13 информативных. Информативные признаки использовались при построении статистических решающих правил классификации. В качестве методов классификации использовались логистическая регрессия, метод опорных векторов, деревья решений и случайный лес [2, 3]. Подбор гиперпараметров алгоритмов происходил по сетке с использованием перекрестной проверки по отдельным объектам [2]. Проблема несбалансированности классов решалась путем назначения весов для наблюдений класса меньшего объема [3]. Веса определяли потери от неправильной классификации пациентов этого класса. Оценка эффективности алгоритмов классификации проводилась методом ROC-анализа по результатам перекрестной проверки по отдельным объектам [3].

Результаты.

В таблице 2 представлены найденные оптимальные гиперпараметры для рассмотренных моделей. В таблице 3 представлены значения показателя AUC (площадь под ROC-кривой) обученных моделей на перекрестной проверке по отдельным объектам. В таблице также представлены значения показателя AUC для обучающей выборки. На рисунке 1 показаны ROC-кривые для обученных моделей на перекрестной проверке по отдельным объектам. На рисунке 2 представлены ROC-кривые для обученных моделей на обучающей выборке.

Таблица 1. Гиперпараметры для рассмотренных моделей

Название модели	Гиперпараметры
Логистическая регрессия	Обратная величина регуляризации=0.0001 Веса классов='сбалансированные'
Случайный лес	Веса классов='сбалансированные' Количество деревьев=30 Максимальная глубина дерева=3
Дерево решений	Максимальная глубина дерева=5 Минимальное число экземпляров в листе дерева=5 Минимальное число экземпляров для разбиения вершины=50 Веса классов='сбалансированные'
Метод опорных векторов	Веса классов='сбалансированные' Обратная величина регуляризации=0.25

Таблица 2. Значения показателя AUC для обученных моделей

Название модели	Показатель AUC на перекрестной проверке по отдельным объектам	Показатель AUC на обучающей выборке
Логистическая регрессия	0.72	0.72
Метод опорных векторов	0.73	0.8
Дерево решений	0.7	0.79
Случайный лес	0.71	0.77

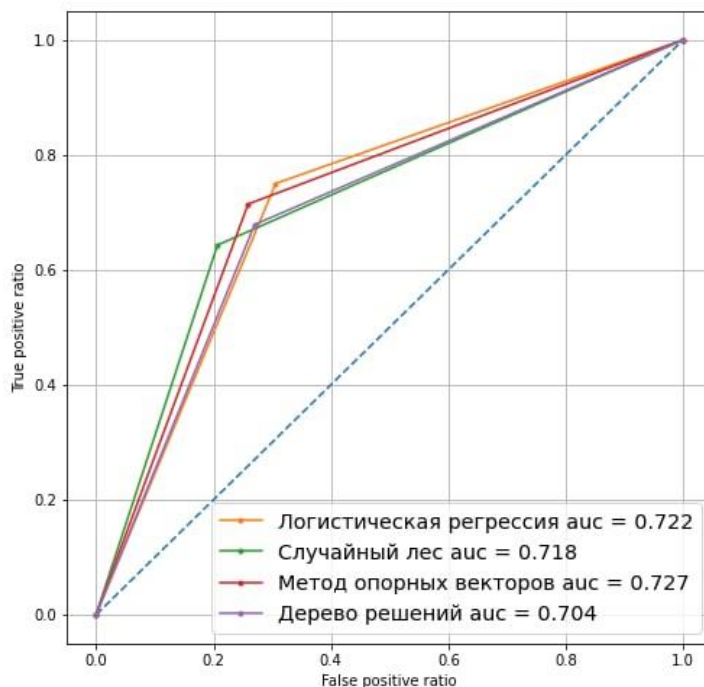


Рисунок 1. ROC-кривые для обученных моделей на перекрестной проверке по отдельным объектам

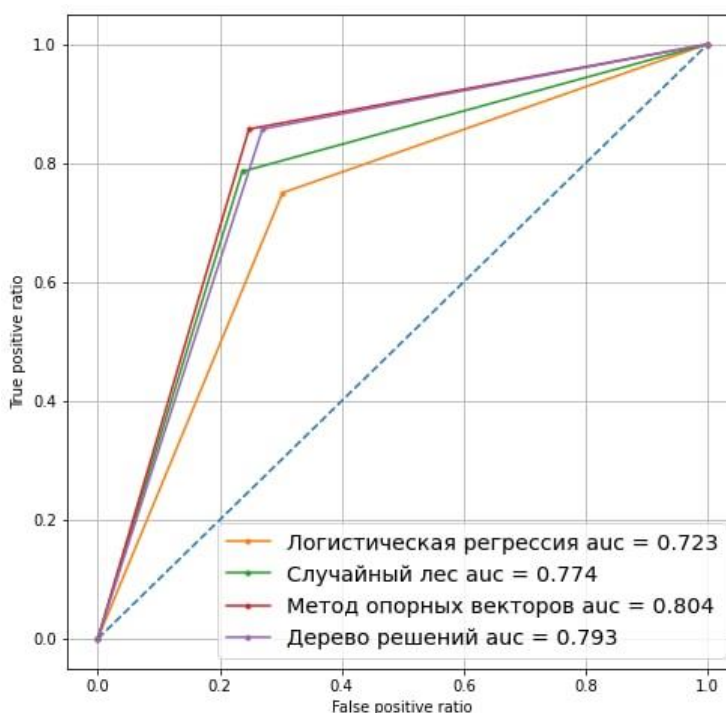


Рисунок 2. ROC-кривые для обученных моделей на обучающей выборке

Заключение.

Полученные результаты показывают, что даже в условиях существенно несбалансированных классов обучающей выборки удалось достичь высоких значений точности классификации успешно прооперированных и прооперированных с перфорациями коронарных артерий пациентов. Максимальная точность классификации пациентов достигнута с применением метода опорных векторов.

Список литературы

- [1] Patel Y., Depta J., De Martini T. Complications of chronic total occlusion percutaneous coronary intervention / Interv. Cardiol. – 2013. – Vol. 5, № 5. – P. 567–575.
[2] Hastie, T. The Elements of Statistical Learning / T. Hastie, R. Tibshirani, J. H Friedman. 2nd ed. New York, Springer Publ. – 2009. – 764 p.
[3] Harrington, P. Machine Learning in Action / P. Harrington. New York: Manning. – 2012. – 382 p.

PREDICTING OF THE OUTCOME OF X-RAY AND VASCULAR SURGERY ON THE CORONARY ARTERIES

V.V. KOZLOVSKY
Magistrate of the BSU

M.S. ABRAMOVICH,
PhD, Associate Professor
Head of the Research Laboratory
Statistical Analysis and Modeling of
Research Institute of Applied
Mathematics and Informatics Problems

V.I. STELMASHOK
PhD, MD, Associate Professor
Head of the Department of
Interventional Cardiology of
National Scientifical and
Practical Centre «Cardiology»

University, Republic of Belarus

*Research Institute for Applied Problems of Mathematics and Informatics of the Belarusian State University,
Republic of Belarus*

National Scientifical and Practical Centre «Cardiology», Republic of Belarus

E-mail: fpm.kozlovskVV2@bsu.by

Abstract. The training set consists of groups of patients who have been successfully operated on and have been operated on with perforation of the coronary arteries. Chi-squared test was used to select important features from all clinical and angiographic features. Class weight for the class with smaller number of samples was modified to increase loss from misclassification because of class imbalance in the training set. Logistic regression, support vector machine, decision tree and random forest were selected for classification. Hyperparameters are optimized for the considered models. ROC analysis showed a fairly high accuracy of patient classification.

Keywords: Coronary arteries, chronic total occlusions, coronary vessel wall perforation, training set, class imbalance, informative features, logistic regression, support vector machine, random forest.