

2019
Том 2

Министерство науки и высшего образования
Российской Федерации
ФГБОУ ВО «Владивостокский государственный
университет экономики и сервиса»

17–19 апреля
2019 г.

ИНТЕЛЛЕКТУАЛЬНЫЙ ПОТЕНЦИАЛ ВУЗОВ – НА РАЗВИТИЕ ДАЛЬНЕВОСТОЧНОГО РЕГИОНА РОССИИ И СТРАН АТР

ИНТЕЛЛЕКТУАЛЬНЫЙ ПОТЕНЦИАЛ ВУЗОВ –

НА РАЗВИТИЕ ДАЛЬНЕВОСТОЧНОГО РЕГИОНА РОССИИ И СТРАН АТР

Материалы XXI Международной
научно-практической
конференции студентов,
аспирантов и молодых ученых

В пяти томах

Том 2



Для функционирования организации предполагается зарегистрировать ИП, количество сотрудников варьируется от 4 до 6 человек. Форма налогообложения – УСН (6% от дохода), с указанием следующих видов деятельности:

47.91 — Торговля розничная по почте или по информационно-коммуникационной сети Интернет Включает в себя:

47.91.1 — Торговля розничная по почте;

47.91.2 — Торговля розничная, осуществляется непосредственно при помощи информационно-коммуникационной сети Интернет. [3]

Таким образом, были разработаны сайт (его серверная составляющая и клиентская), продумана схема работы предприятия и выбран вариант регистрации юридического лица.

1. Python-scripts, база знаний. Плюсы и минусы Django [Электронный ресурс] / python-scripts.com URL: <https://python-scripts.com/django-obzor>

2. Django project, сайт автора фреймверка. Getting started with Django [Электронный ресурс] / djangoproject.com URL: <https://www.djangoproject.com/start/>

3. Центр Регистрации и Сопровождения Бизнеса, Код ОКВЭД 47.91 [Электронный ресурс] / www.regfile.ru URL: <https://www.regfile.ru/okved2/razdel-g/47/47.9/47.91.html>

Рубрика: Машинное обучение

УДК 004

ПОСТРОЕНИЕ МОДЕЛИ СЛУЧАЙНОГО ЛЕСА ДЛЯ ДИАГНОСТИКИ ИШЕМИЧЕСКОЙ БОЛЕЗНИ СЕРДЦА

С.Г. Чиркбая

бакалавр

Е.Д. Емцева

преподаватель

Владивостокский государственный университет экономики и сервиса
Владивосток, Россия

В современном мире особенно часто люди подвергаются сердечным заболеваниям. Они являются одними из основных причин смертности, временной и стойкой утраты трудоспособности во всем мире. В структуре смертности сердечно-сосудистые заболевания стоят на первом месте, из них на долю ишемической болезни сердца приходится около 40%. На данный момент не существует «идеального» способа диагностики ИБС. Коронарография – один из самых достоверных способов, но он является крайне опасным для жизни, люди нуждаются в альтернативе.

Ключевые слова и словосочетания: Ишемическая болезнь сердца, коронарография, модели случайного леса.

CONSTRUCTION OF A RANDOM FOREST MODEL FOR THE DIAGNOSIS OF CORONARY HEART DISEASE

In the modern world, it is especially common for people to undergo heart disease. They are among the leading causes of death, temporary and permanent disability worldwide. In the structure of mortality, cardiovascular diseases are in the first place, of which the share of coronary heart disease accounts for about 40%. At the moment, there is no "ideal" way to diagnose IHD. Coronary angiography is one of the most reliable ways, but it is extremely life-threatening, people need an alternative.

Keywords: coronary heart disease, coronary angiography, Random Forest.

Целью данной работы является построение моделей, позволяющих диагностировать ИБС на основе различных показателей медицинского обследования и индивидуальных характеристик пациента.

Для достижения поставленной цели необходимо выполнить следующие задачи:

– изучить возможность применения модели Random Forest в диагностике ишемической болезни сердца (ИБС),

– создать модели, представляющие альтернативную коронарографии возможность диагностики ИБС.

Нам были предоставлены данные по диагностике ишемической болезни сердца и других результатов медицинского обследования пациентов городской клинической больницы города Владивостока.

База данных содержит более 200 показателей для более 300 пациентов.

ИБС – ишемическая болезнь сердца (стенокардия напряжения, нестабильная стенокардия, инфаркт миокарда) – является следствием сужения и закупорки основных артерий сердца атеросклеротическими бляшками. Со временем их становится все больше и больше, и когда просвет сосуда перекрывается на 50% и более возникает затруднение тока крови. Вследствие этого уменьшается доставка кислорода и питательных веществ к мышце сердца, развивается кислородное голодание (гипоксия), что ведет к ишемии миокарда.

В структуре смертности сердечно-сосудистые заболевания стоят на первом месте, из них на долю ИБС приходится около 40%.

Среди факторов, увеличивающих риск развития ИБС можно выделить контролируемые и неконтролируемые.

Факторы, на которые человек не может оказывать влияние:

- возраст/пол: мужчины старше 45 лет, женщины старше 55 лет
- раннее развитие ИБС у близких родственников: – у мужчин до 55 лет, – у женщин до 65 лет
- Факторы, которые человек может контролировать
- повышенное АД,
- курение,
- повышенный уровень холестерина в крови,
- повышенный уровень глюкозы в крови (диабет),
- избыточное потребление алкоголя,
- сидячий образ жизни,
- избыточный вес,
- стрессовые ситуации.

К большому сожалению врачей, и пациентов, лечение этой «болезни века» лекарственными препаратами недостаточно эффективно. Ведь таблетки не в силах «растворить» бляшку внутри сосуда сердца. Требуется аортокоронарное шунтирование – тяжелейшая хирургическая операция. Волнения близких, долгие месяцы восстановления, пугающее слово «инвалидность».

Коронарография остается «золотым стандартом» в диагностике ИБС, так как на сегодняшний день это единственный из существующих методов исследований, позволяющий определить точные детали анатомического строения всего коронарного русла.

Это – инвазивный метод исследования. Изображения артерий, питающих сердце – коронарных артерий, достигается введением в них контрастного вещества через специальные катетеры. Катетеры подводятся к сосудам сердца через небольшой прокол в области запястья или, при необходимости, через паховую область. Коронарография предоставляет врачу всю необходимую информацию о состоянии коронарных артерий.

Вообще проблемы при выполнении коронарографии встречаются очень редко, не более чем 1 случай на 100 000 исследований. Однако учитывая, что ангиографическое исследование само по себе является травматичным, риски возникновения тех или иных осложнений всегда присутствуют.

Опасность:

- кровотечение из области пункции артерии,
- нарушение ритма сердца,
- образование тромбов и гематом,
- инфекция и воспаление раны в области пункции,
- аллергическая реакция на контрастный препарат,
- инсульт,
- инфаркт миокарда,
- перфорация кровеносного сосуда (образование отверстия в сосуде) или его разрыв,
- кессонная болезнь (попадание в просвет кровеносного сосуда воздуха, что может быть крайне опасным для жизни осложнением),
- смерть.

Так же нельзя не учесть тот факт, что данная диагностика является крайне дорогостоящей. И это несправедливо, ведь многие люди, просто не состоянны оплатить данную операцию и помочь себе.

В данной работе в качестве оцениваемой диагностической модели использовалась модель случайного леса.

Случайный лес – один из алгоритмов машинного обучения, придуманный Лео Брейманом и Адель Катлер еще в прошлом веке.

Дерево принятия решений – средство поддержки принятия решений, которое использует древовидный граф или модель принятия решений, а также возможные последствия их работы, включая вероятность наступления события, затраты ресурсов и полезность.

Наши данные были разделены на обучающую и тестовую подвыборки. В качестве зависимой переменной в данной работе является результат диагностирования ИБС методом коронарографии ($y=0$, если поставлен диагноз ИБС, $y=1$, в противном случае). Предикторами выступают различные факторы медицинского обследования и индивидуальных характеристик пациентов. Объясняющие переменные в этой модели могут быть как категориальные, так и количественные. Для построения моделей использована программа RStudio [4].

Метод случайного леса осуществляет бутстреп-агрегирование некоторого множества деревьев решений. Алгоритм случайного леса обычно хорошо прогнозирует, но не дает возможности получения простой наглядной формулы. Для описания внутреннего процесса построения леса используют меры важности регрессоров на основе среднего падения индекса Джинни. При каждом ветвлении на дереве падает индекс Джинни. Для каждой переменной считается суммарное падение индекса Джинни, вызванное ветвлением на базе этой переменной. Посчитав среднее падение Джинни по всем деревьям, получают меру важности.

Для построения леса использовали функцию randomForest(). На данный момент нами получена модель со следующими предикторами: эритроциты, тромбоциты, гемоглобин, AST, ALT, лейкоциты, ЭКГ.

С помощью функций importance() и varImpPlot() оценим (рисунок1) и отобразим (рисунок 2) важность предикторов.

	meanDecreaseGini
эритроциты	17.75296
тромбоциты	15.70998
гемоглобин	15.81446
AST	12.24568
ALT	14.10025
лейкоциты	17.75341
ЭКГ	11.95470

Рис.1. Индекс Джинни

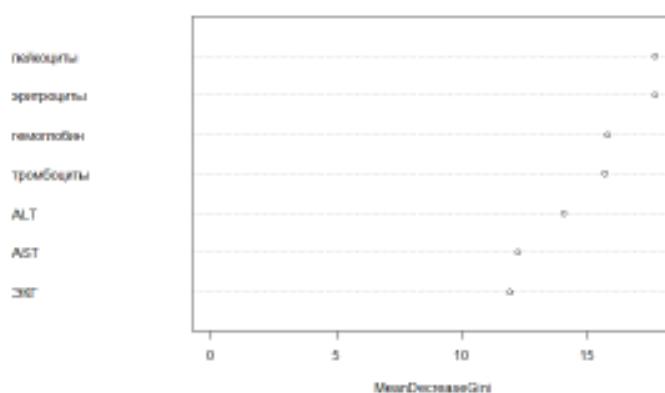


Рис. 2. Степень важности предикторов

В результате получены предсказанные значения по тестовой выборке:

0 0 0 0 0 0 0 1 0 0 0 1 0 1 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0
0 0 0 1 1 0 1 0 0 0 1 0 0 0 0
0 1 0 0 0 0 0 0 0 .

В то время, когда фактические значения по тестовой выборки:

0	0	0	0	0	1	0	0	0	1	1	0	1	1	0	0	0	0	0	1	1	0	0	0	1	0	1	0	1	1	0	
1	0	1	0	1	0	1	1	1	0	0	0	1	0	0	0	0	1	1	0	0	0	0	0	0	1	0	0	0	0	0	0

Доля верно классифицированных наблюдений по тестовой выборке составляет около 67%.

Мы понимаем тот факт, что полученная модель не идеальна, мы продолжаем поиск более качественных моделей.

Эритроциты – самая многочисленная разновидность клеток крови человека. Главная функция эритроцитов – перенос кислорода и оксида углерода. Помимо этого, красные кровяные тельца участвуют в транспорте питательных веществ, иммунных реакциях и помогают поддерживать кислотно-щелочного равновесие.

Тромбоциты – это форменные элементы крови, маленькие кровяные пластинки – образуются в красном костном мозге, материалом служит плазма его клеток

АЛТ преобладает в печени, а АСТ – в миокарде, следовательно, при инфаркте миокарда или гепатите обнаружится повышенная активность в крови какого-либо данного фермента.

Повышение АСТ при одновременном росте отношения АСТ/АЛТ свидетельствует о поражении сердца, и можно с уверенностью говорить об инфаркте миокарда. Если же наоборот, то это говорит о поражении печени.

1. Осложнения коронарографии, риски: основные проблемы после реконструкции сосудов через руку – гематома и прочее [Электронный ресурс]. URL: <http://cardiobook.ru/oslozhneniya-koronarografii/>

2. Ишемическая болезнь сердца – Российский кардиологический научно-производственный комплекс [Электронный ресурс]. URL: <https://www.rusintervention.ru/пациентам/> заболевания/ибс/

3. Кабаков, Р.И. R в действии. Анализ и визуализация данных в программе R / Р.И. Кабаков; пер. с англ. П.А. Волковой. М.: ДМК Пресс, 2014. 588 с.

4. Официальный сайт проекта R. [Электронный ресурс]: URL: <https://cran.r-project.org/>