
Метод минимизации необоснованной антибиотикотерапии пациенто в с коронавирусной инфекцией

Алексеев Александр Николаевич

Понимание цели

В клинической практике, назначение антибиотиков требует наличие показаний: наличие доказанной или высоковероятное наличие бактериальной флоры, вызывающей инфекционный процесс.

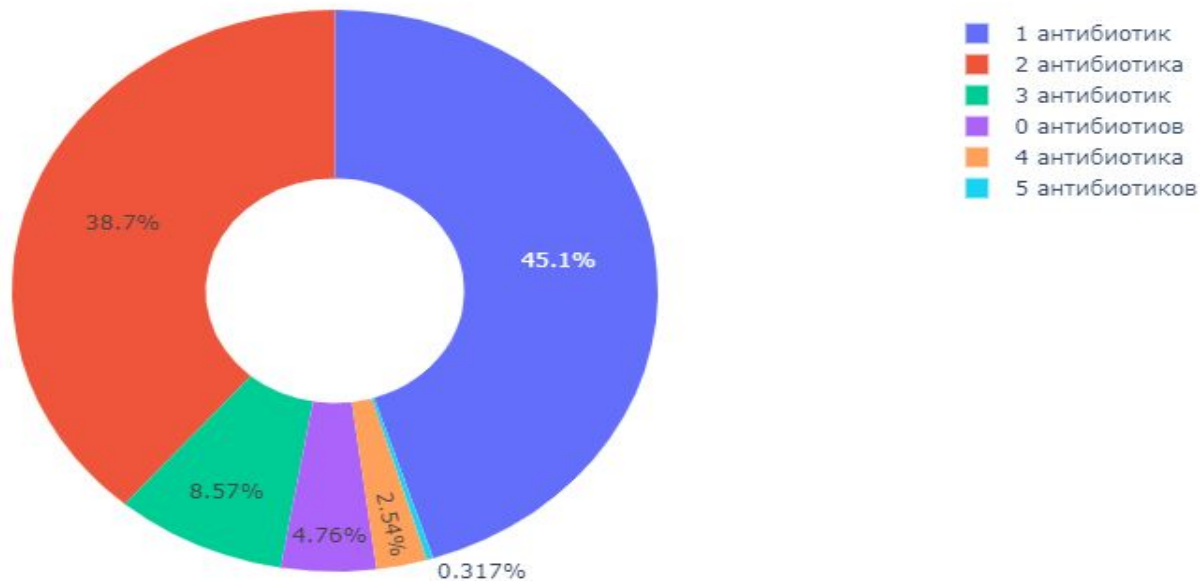
Типичная пневмония (воспаление лёгких) в 95% и более случаях обусловлена бактериальной инфекцией, поэтому антибактериальная терапия в случае пневмонии назначается в 100% случаев. Однако, с приходом коронавирусной инфекции баланс этиологических факторов поменялся.

При вирусных инфекциях - антибиотики не показаны, т.к. нет точки приложения данной группы препаратов, однако, с другой стороны, при присоединении бактериальной флоры - антимикробные препараты показаны, т.к. доказано, что они улучшают прогноз у такой категории пациентов.

Таким образом, имеет место обстоятельство, которое диктует врачу назначать антибактериальные средства в случае, если невозможно исключить наличие бактериальной флоры у пациентов с вирусными инфекциями

В ГОКБ МР 95 % пациентов с COVID-19 получают

Антибактериальные схемы



Понимание цели

В какой-то мере, решением проблемы необоснованной антибиотикотерапии - исследование уровня прокальцитонина

Повышение уровня прокальцитонина в крови выше 0.5 нг/мл указывает на развитие бактериальной инфекции (чувствительность — 80–95%, специфичность — 88–93%)

Однако, доступность данного анализа является значимым ограничением данного решения, поэтому возникает потребность в поиске метода, который позволял бы определять показания к антибиотикотерапии помимо уровня прокальцитонина

Формулировка цели


Построить модель машинного обучения, которая смогла бы минимизировать количество назначенных антибиотиков при нулевом уровне ложноотрицательных результатов

→ Пример эпикриза 2. Доступные данные

- ВЫПИСНОЙ ЭПИКРИЗ № 170 Находился(ась) в пульмонологическом отделении ГОКБ МР Ф.И.О.: Дата рождения: 02.08.1966 Адрес регистрации: Адрес проживания: Место работы: Должность: Поступил: 13.01.2021 Выписан: 26.01.2021
- Диагноз клинический: Тяжёлая коронавирусная (ТОРС-КОВ-2 ПЦР+ от 12.01.2021) инфекция. Острое вирусное повреждение лёгких, КТ 3 ДН 1 ст.
- Диагноз заключительный клинический: В34.2 / (основной) / Тяжёлая коронавирусная (ТОРС-КОВ-2 ПЦР+ от 12.01.2021) инфекция. Острое вирусное повреждение лёгких, КТ 3 в стадию обратного развития, ДН 0 ст. ИБС: Атеросклеротический кардиосклероз. Атеросклероз аорты, коронарных артерий. АГ 2ст., риск 4. Н 1ст. Сахарный диабет, тип 2, клинико-метаболическая субкомпенсация. МКБ (оперативное лечение 2019) ХБП С2 (рСКФ -72 мл/мин/1,73м2). Дискогенная радикулопатия S1 слева, стойкий выраженный болевой синдром. Диабетическая полинейропатия, дистальный тип, с умеренными сенсо-моторными нарушениями. Ожирение 3 ст. (ИМТ - 42 кг/м2)
- Консультации врачей-узких специалистов, консилиумы врачей:
- Консультация невролога 14.01.2021, 20.01.2021 Консультация нейрохирурга 20.01.2021 показаний к экстренному оперативному вмешательству нет
- Результаты инструментальных и аппаратных методов исследования:
- ЭКГ от 13.01.2021 Ритм синусовый, правильный, ЧСС - 93 уд в мин, норм.полож.ЭОС, неспецифические изменения в миокарде нижней стенки ЛЖ МРТ от 19.01.2021 Заключение: МР-картина дегенеративно-дистрофических изменений пояснично-крестцового отдела позвоночника в виде остеохондроза с экструзией диска в Л5-С1 сегменте Результаты лабораторных исследований: Гликемический профиль 14.01.2021: 8 ч. 00 мин. 25.8 ммоль/л ; 12 ч. 00 мин. 24.3 ммоль/л ; 16 ч. 00 мин. 25.5 ммоль/л ; Гликемический профиль 20.01.2021: 8 ч. 00 мин. 9.8 ммоль/л ; 12 ч. 00 мин. 10.3 ммоль/л ; 16 ч. 00 мин. 9.5 ммоль/л ; Гемостазиограмма 14.01.2021: Активированное частичное тромбопластиновое время 22.1 с ; Ratio АЧТВ 0.744 ; Протромбированное время 11.2 с ; Протромбиновый индекс 1.01 ; Международное нормализованное отношение 0.981 ; Фибриноген 8.1 г/л ; Биохимическое исследование крови 14.01.2021: Общий белок 71 г/л ; Мочевина 10.7 ммоль/л ; Креатинин 98.1 мкмоль/л ; С-реактивный белок 41.5 мг/л ; Билирубин общий 10.5 мкмоль/л ; Аспартатаминотрансфераза 27.8 Ед/л ; Аланинаминотрансфераза 54.0 Ед/л ; Натрий 138.7 ммоль/л ; Калий 5.89 ммоль/л ; э\ Прокальцитонин от 14.01.2021 0.14 нг/мл Биохимическое исследование крови 15.01.2021 Глюкоза - 22,1 ммоль/л Общий анализ крови 14.01.2021: WBC (Лейкоциты) 9.1 109/ ; RBC (Эритроциты) 4.09 1012/л ; HGB (Гемоглобин) 129 g/l ; PLT (тромбоциты) 495 10*9/л ; СОЭ 35 мм/ч ; Нейтрофилы палочкоядерные 3 % ; Нейтрофилы сегментоядерные 68 % ; Лимфоциты 26 % ; Моноциты 3 % ; \C у м а к л е т о к : 100 ;
- Состояние при выписке: относительно удовлетворительное. При выписке сатурация крови кислородом при дыхании атмосферным воздухом - 97%, ЧД 18 в мин. уровень лейкоцитов более 3 тыс. на мл., не лихорадит более 3 дней без жаропонижающих. Уровень СРБ менее 2х норм
- Проведенное лечение: гепарин, пантопрозол, ацецезон, гликлазид, метформин, лизиноприл, метопролол, индапамид, эуфиллин, новокаин, димедрол, вит группы В, ксефокам, толперизон, карбомазепин, кетаролак, зопиклон, пентоксифиллин, трамадол, мелоксикам, ФТЛ
- Рекомендации: Наблюдение участкового терапевта, эндокринолога, невролога по месту жительства (ДЗ) Избегать переохлаждений, витаминно-минеральный комплекс Снижение употребления поваренной соли менее 5грамм в сутки, животных жиров, рафинированных углеводов. Повышение употребления овощей , омега-3 ненасыщенные ЖК, умеренные аэробные физические нагрузки Контроль гликированного гемоглобина Целевой уровень ЛПНП - ниже 1,8 ммоль/л Ривароксабан 10 мг 1 раз в день (45 дней), при невозможности - аспикард 75 мг после обеда (не менее 3 месяцев) Ацетилцистеин 1,2 в день - 45 дней Гликлазид 60мг 1 табл. утром Метформин 1000мг 2 раза в день Лизиноприл/гидрохлортиазид 10/12,5мг утром Метопролол 25 мг 2 раза в день Невролог: 1) Дексаметазон 0,5 мг № 5 2) Нимесулид 100 мг 2 раза в день (утро, вечер) при необходимости 3) Мидокалм 150 мг по 1 табл 3 раза в день - 14 дней 4) Боровит 2,0 в/м № 10

—

Написаны методы, позволяющие получать все интересующие данные



Использованные
библиотеки:

-Glob

-Os

-Regex

-String

-NLTK

Имеющиеся данные

Анализируемые независимые признаки:

Численные данные:

- 1) Тромбоциты
- 2) СОЭ
- 3) Фибриноген
- 4) Лактатдегидрогеназа
- 5) Лейкоциты,
- 6) С-реактивный белок
- 7) Аланин-трансаминаза
- 8) Аспартат-трансаминаза
- 9) Длительность пребывания
- 10) Возраст
- 11) Креатинин
- 12) Скорость клубочковой фильтрации
- 13) Лейкоцитарно-тромбоцитарное соотношение

Категориальные данные: Пол, Тяжесть COVID-19, Ишемическая болезнь сердца, Хроническая обструктивная болезнь лёгких, Бронхиальная астма, Ожирение, Диабет, Дыхательная недостаточность (0, 1, 2, 3)

Зависимый признак: показание к антибиотикотерапии: уровень прокальцитонина более 0.4 нг/мл выбран как порог показания АБ терапии

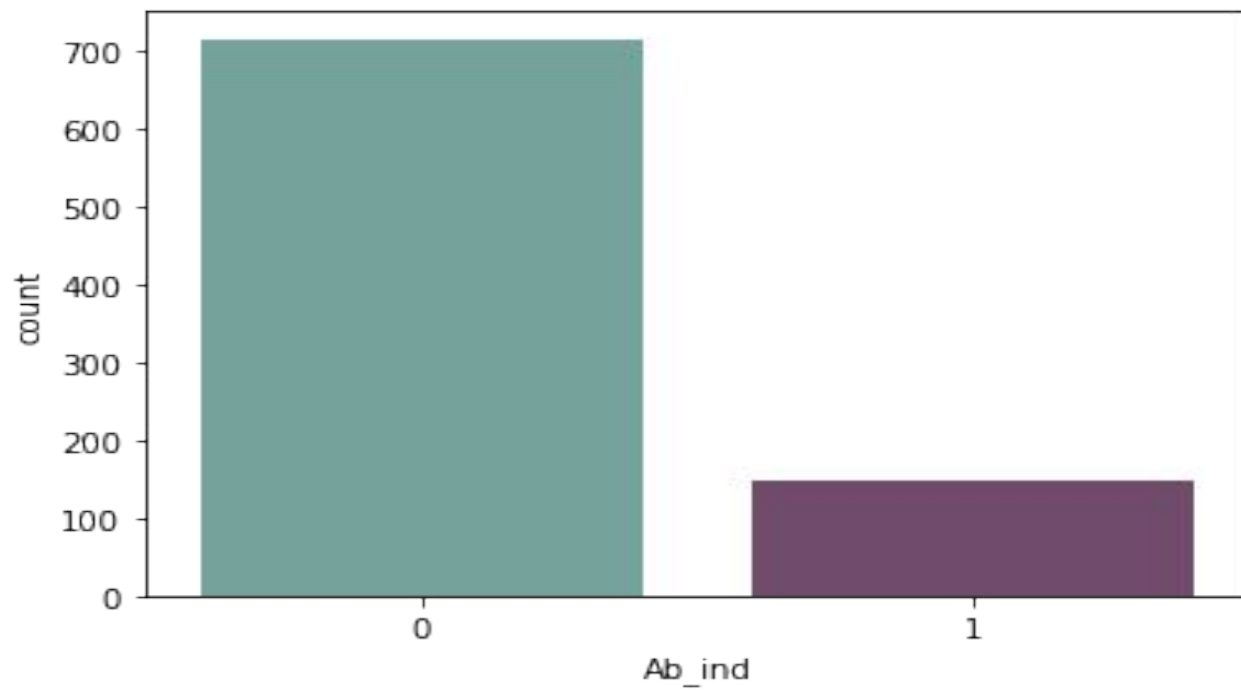
- 1 – АБ показаны
- 0 – АБ не показаны

Итого мы имеем матрицу размерности 863 на 25

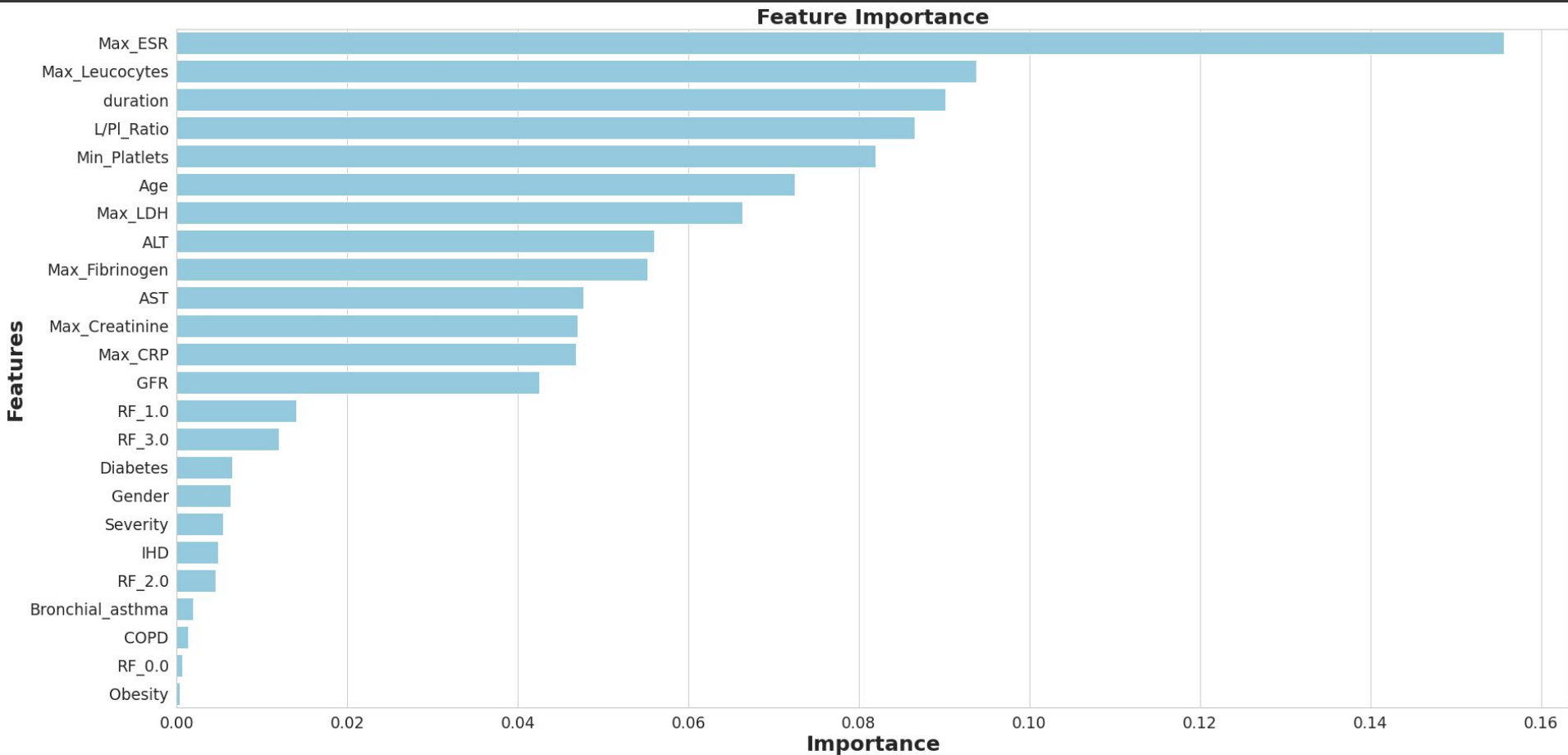
```
In [ ]: covid_data.head()
```

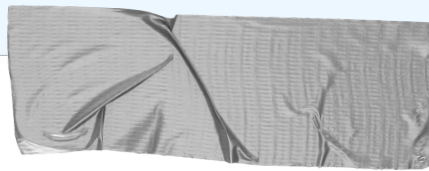
Out []:	Min_Platlets	Max_ESR	Max_Fibrinogen	Max_LDH	Max_Leucocytes	Gender	Max_CRP	Severity	IHD	COPD	Bronchial_asthma	Obesity	Diabetes	ALT
0	142.8	54.0	4.8	513.0	10.0	1	88.7	0	0	1	0	0	0	57.2%
1	261.5	36.0	10.0	287.6	8.3	0	47.4	1	0	0	0	0	0	101.6%
2	225.4	36.0	9.3	851.4	12.5	0	57.4	0	0	0	0	0	0	69.5%
3	323.3	57.0	3.7	690.3	10.9	0	116.0	0	0	0	0	0	0	46.8%
4	168.8	52.0	7.3	567.1	7.7	0	189.6	0	0	0	0	0	0	43.2%

Баланс классов



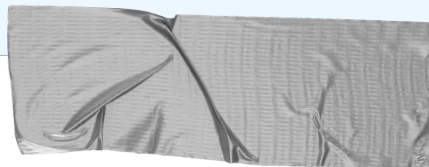
Определяем наиболее важные признаки, влияющие на исход





2. Первичные выводы, исходя из классических статистических тестов

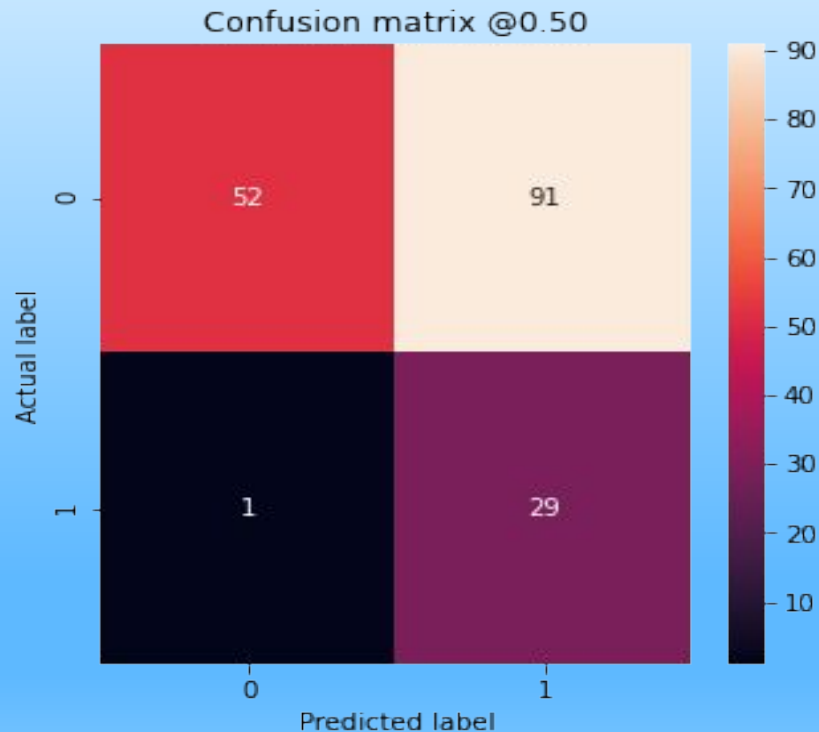
1. Данные в группах (1, 0) – в основном распределены ненормально
2. В наиболее важных группах имеет место статистически значимая разница по группам (1, 0) (Тесты: Крускал-Уолисс, Манна-Уитни $p < .01$)
3. Таким образом, есть основания полагать, что в группах имеет место объективные различия, которые попробуем найти



3. Машинное обучения с учителем. Использованы следующие алгоритмы:

1. **LogisticRegression**
 2. **SVM**
 3. **CatBoost**
 4. **LGBM**
 5. **XGB**
 6. **RandomForest**
- **Подбор параметров: библиотека FLAML и использовался поиск по сетке с кросс-валидацией (где FLAML не работала)**
 - **Метрикой качества была выбрана AUC (площадь под ROC-кривой)**

Результат голосования моделей на тестовой выборке



—

***Модели машинного обучения
смогли продемонстрировать
98 % чувствительность (recall),
24 % специфичность (precision)***

– Глубокое обучение в TensorFlow с бустрапированием

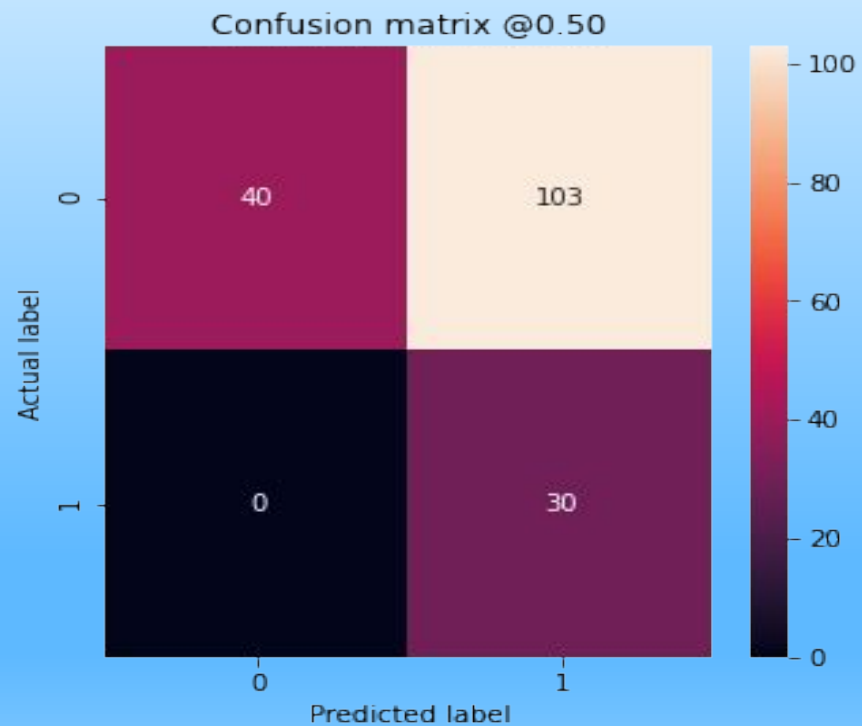
Использовался подход «перенабора» признаков с лейблом «1», тем самым позволим лучше нейронной сети понять закономерности и взаимосвязи таргетного значения с признаками

Глубокое обучение в TensorFlow с бустрапированием



```
# чтобы попытаться улучшить модель, попробуем использовать бустрап-подход, а именно:  
# уже существующие признаки случайным образом перенаберём, тем самым, позволим  
# "лучше" нейронной сети понять закономерности взаимосвязи таргетного значения с признаками  
bool_train_labels = train_labels != 0  
pos_features = train_features[bool_train_labels]  
neg_features = train_features[~bool_train_labels]  
pos_labels = train_labels[bool_train_labels]  
neg_labels = train_labels[~bool_train_labels]  
ids = np.arange(len(pos_features))  
choices = np.random.choice(ids, len(neg_features))  
res_pos_features = pos_features[choices]  
res_pos_labels = pos_labels[choices]  
resampled_features = np.concatenate([res_pos_features, neg_features], axis=0)  
resampled_labels = np.concatenate([res_pos_labels, neg_labels], axis=0)  
order = np.arange(len(resampled_labels))  
np.random.shuffle(order)  
resampled_features = resampled_features[order]  
resampled_labels = resampled_labels[order]  
resampled_features.shape
```


$class_weight = \{0:1, 1:7.2\}$



Оценка качества

- *Нейронные сети продемонстрировали способность предсказывать показания к антибактериальной терапии со 100 % чувствительностью и 22,5 % специфичностью*

Выводы

- 1) Целевой показатель ложнонегативное предсказание равен 0, что являлось принципиальным условием
- 2) Лучшая модель способна снизить назначение антибиотиков с 95 % (изначально) до 76 %
- 3) Поставленная цель достигнута при помощи глубокого обучения, чтобы улучшить модель необходимо больше данных
- 4) Стекинг моделей может восполнить недостатки каждой из моделей машинного обучения и помочь в достижении целей, но в моем случае – «стэкинг» не помог достичь лучшего результата

Внедрение модели

tk

Предикция показаний антибактериальной терапии у пациентов с COVID-19

СКФ

90

Возраст

34

Уровень тромбоцитов

142

Уровень лейкоцитов

10

Уровень ЛДГ

700

Уровень АЛТ

32

Уровень АСТ

34

Уровень креатинина

80

Уровень СОЭ

29

Уровень фибриногена

5

Длительность

10

Наличие ИБС

Нет

Наличие ХОБЛ

Нет

Наличие бронхиальной астмы

Да

Наличие диабета

нет диабета

Степень дыхательной недостаточности

1

Пол

женский

Тяжесть

Нетяжелая

Получить результат

Сбросить

Выйти

Предсказание

Антибиотик не показан

OK

Благодарю за внимание