



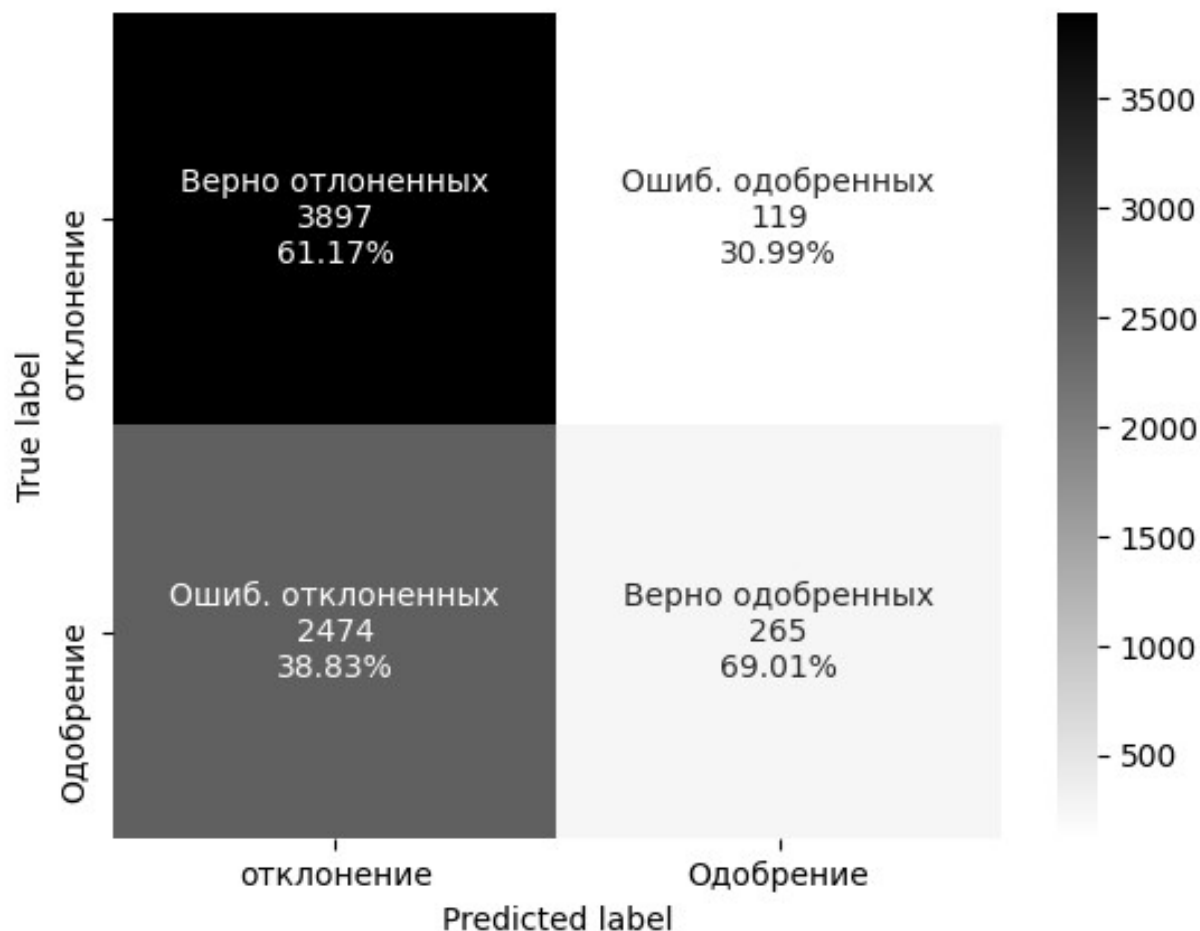
Белая зона решения андеррайтера

- 1) Если «Оценка кредитной истории ОКБ» равно «КИ отсутствует», то вес ставим -0.32, равно Хорошая КИ, то +0.175, иначе 0.023
- 2) По Формуле вычисляем вес «Эквифакс скоринг» * 0.00286 - 2.46
- 3) Если «Стаж работы» < 61 , то вес -0.02, иначе вес +0.23
- 4) Если «Семейное положение» = 'не в браке', то вес -0.26, если = 'в браке', то вес 0.03
- 5) Если $0 \leq \text{«Первоначальный взнос»} \leq 50000$, то 0.21 если $50000 \leq \text{ПВ} \leq 300000$, то -0.2, если $\text{ПВ} > 300000$, то 0.36
- 6) Если «НБКИ Скоринговый Бал» ≤ 700 , то -0.07, иначе «НБКИ Скоринговый Бал» * 0.001 — 0.84
- 7) «МБКИ», Если ОКБ КИ отсутствует, то мультипликатор = 0.9
 - МБКИ_требаналотч присутствует, то -0.18*мультип
 - МБКИ_треб_исп_пр присутствует, то -0.07*мультип
 - МБКИ_нет_огр присутствует, то 0.1*мультип

Степень уверенности в Одобрении $[0,1] = \text{вес} / 0.255$

- МАХ Одобрения: $0.1 + (950 * 0.6) / 255 - 2.2 + 0.1 + 0.02 = 0.255$

Accuracy: 0.616136
Precision: 0.690104
Одобрённых: 0.056847



Заявки 2020-01-01 по 10 мая 2022

56255 обучающих и 6917 проверяющих заявок.

Разделены по дате 2022-02-24, (та самая дата).

- верно одобрённых 265	сумма размеров кредита — 526,375,702 рублей
- ошибочно одобрённых 119	сумма размеров кредита — 253,168,559 рублей

Выделенная белая зона

- 1) Если «Оценка кредитной истории ОКБ» не равно «КИ отсутствует»
- 2) По Формуле вычисляем вес «Эквивалентный скоринг» * 0.00273 — 2.36 => > 864
- 5) ПВ<300000
- 6) «НБКИ Скоринговый Бал» * 0.0098 — 0.52 => > 53
- 7)
 - МБКИ_треб_исп_пр присутствует и ОКБ КИ отсутствует, то отклоняем
 - МБКИ_нет_огр присутствует и ОКБ КИ отсутствует, то отклоняем

Мы получили эту зону путем постановки условия в формуле, что мы одобряем заявки не имеющие отрицательные значения, а затем перебрали все возможные значения параметров.

Получили:

Одобрённых: 0.104959

Accuracy: 0.621614

Precision: 0.629055



Вариант выделения белой зоны

Отбросим (-) значения в формуле, и будем считать одобренным те заявки, которые дают больше 0.46:

Одобрённых: 0.056255

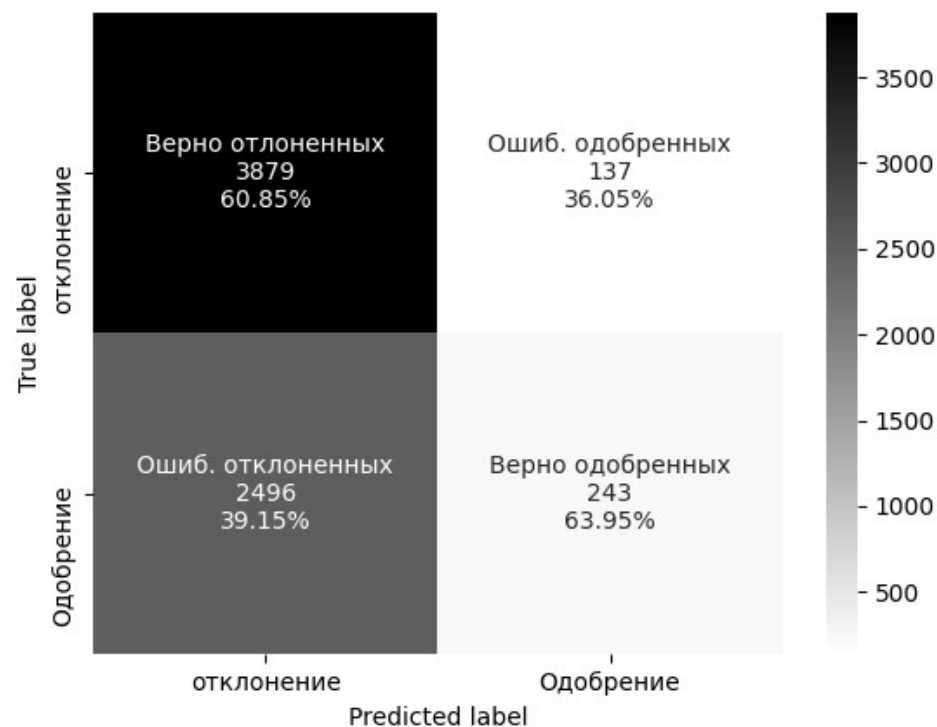
Accuracy: 0.610215 — точность

AUC: 0.609714

Precision: 0.639474 — точность одобренных

Мы модобрали 0.46, чтобы иметь такое же количество одобренных, как в полной формуле.

Как можно видеть, что мы потеряли в точности и точности одобренных.



1)

Данные



Формула
с параметрами
 $y = a \cdot x + c$



Оценки
эффективности

Подбор параметров
авто-перебором



Хотим меньше, но точнее!
Ок!

2) Улучшение самой формулы, нечасто, раз в два год.

3) Использовать степень уверенности



*** Почему в формуле много строк?**

Потому что один-два параметра не дают достаточного охвата хороших клиектов.

*** Почему это формула, а не зона?**

Формула отличается от зоны, только то что в ней есть (-) значения. Это придает гибкости. Чтобы формула стала зоной, в вашем понимании, нужно просто отбросить (-) значения в формуле.

*** Почему в белой зоне есть отрицательные значения?**

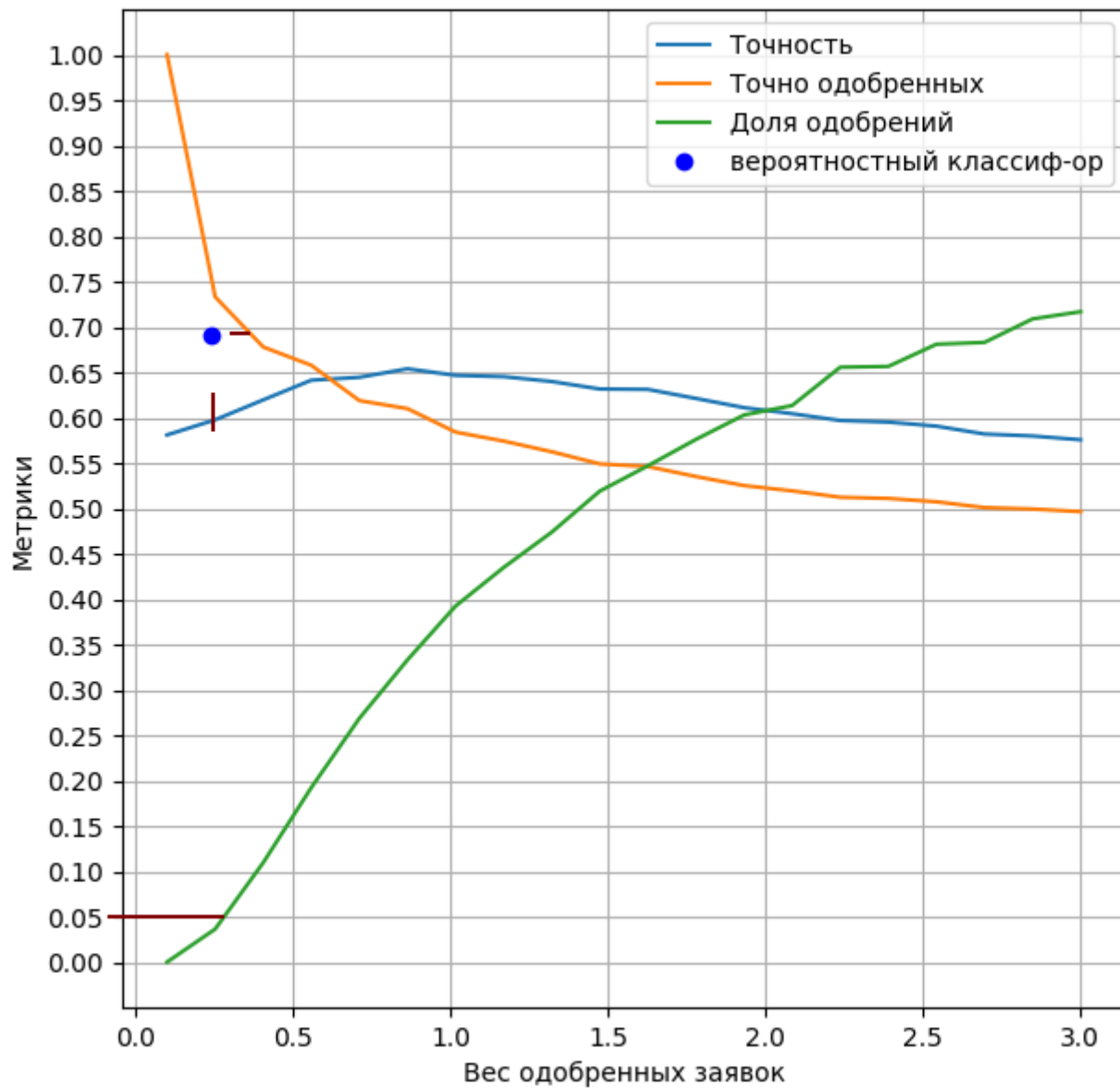
Если убрать отрицательные значения, то в белую зону (+) будет попадать меньшее количество хороших заявок, они будут отсеиваться по низкому значению одного из признаков.

*** Как эту формулу рассчитывать, это очень сложно, наверное.**

Формулу можно проверить в Экселе. Подстраивается она простым перебором всех значений параметров.

Но чем точнее настройка, тем технически сложнее.

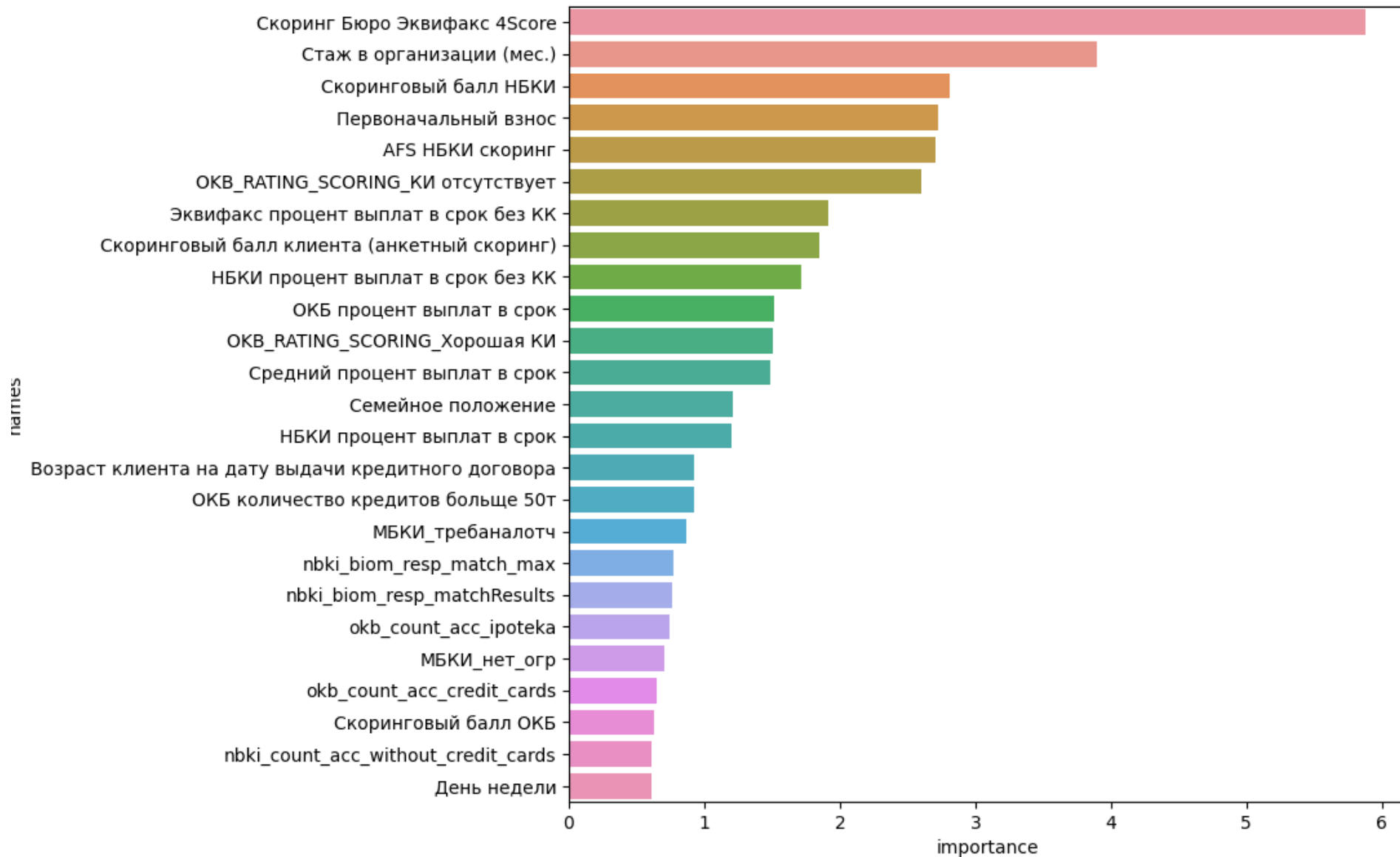




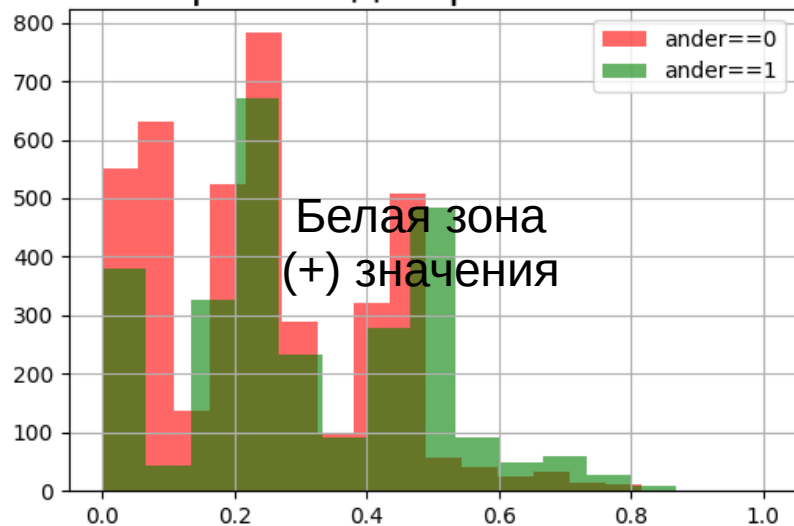
Линии — это самая лучшая модель, Которую только удастся создать.

Технически наша формула, это вероятностный Калассификатор.

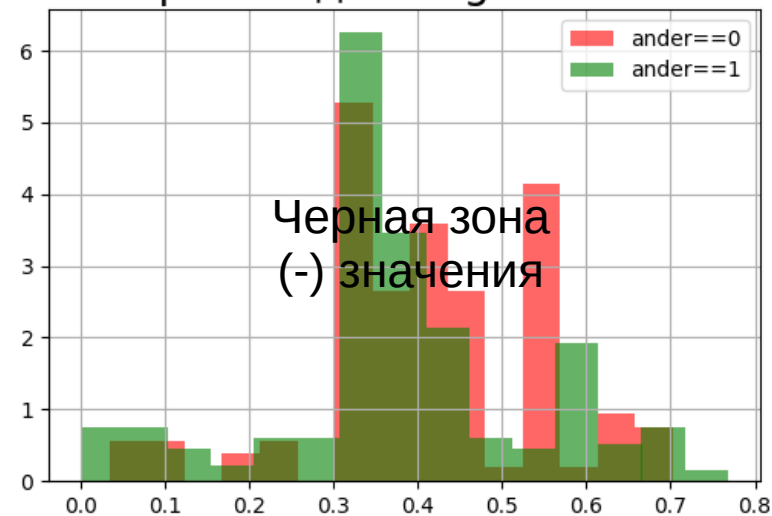
Важность признаков сумма XGBoost и RandomForest



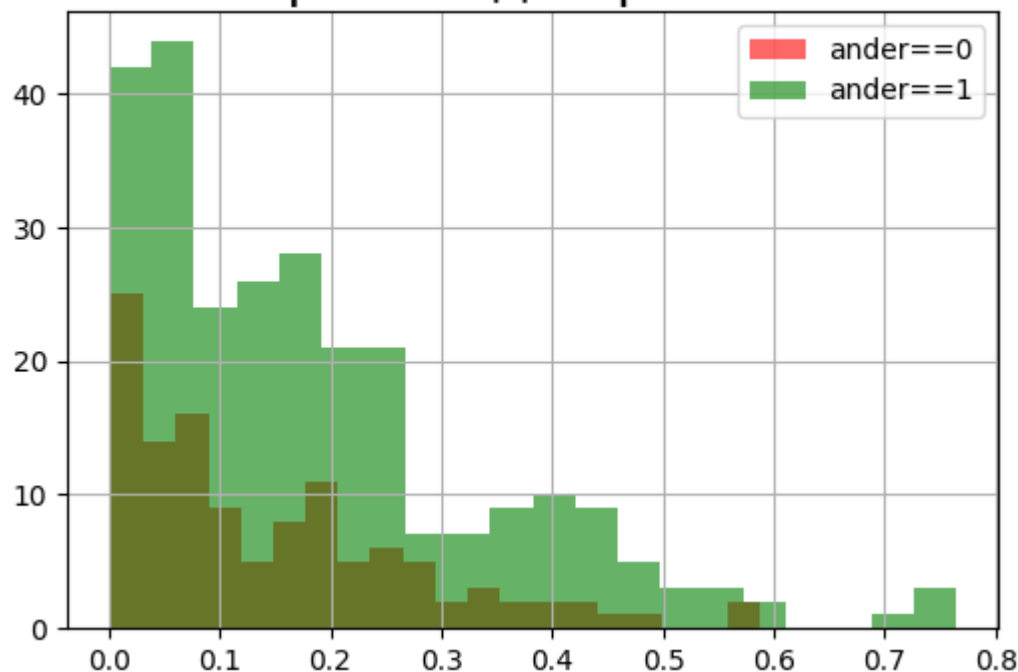
Гистограмма для positive по ander



Гистограмма для negative по ander



Гистограмма для p-n по ander



Налажение «Белой» и «Черной» зоны, где >0

