

ПОИСК LOOK-ALIKE-АУДИТОРИЙ НА ОСНОВЕ АНАЛИЗА ДАННЫХ О КЛИЕНТАХ МАГАЗИНОВ «ПЕРЕКРЁСТОК»



Drill Team

Ильназ Зиязиев

ilnazziaziev@gmail.com

Иван Богуш

shigurui2@gmail.com

Станислав Безвытный

stasbezv@gmail.com

Федор Бардин

fedya_bardin@mail.ru

Гипотезы

Гиперпараметры

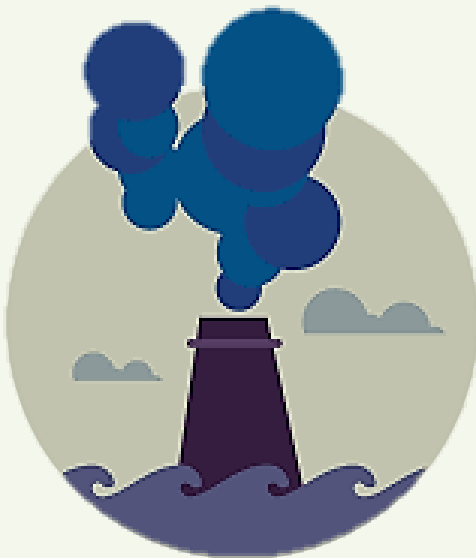
Рекомендации

EDA

Метрики

Первая гипотеза

- Подаваемые данные сильно засорены и содержат выбросы



Вторая гипотеза

- Гиперпараметры оптимизированы недостаточно хорошо



Гипотезы

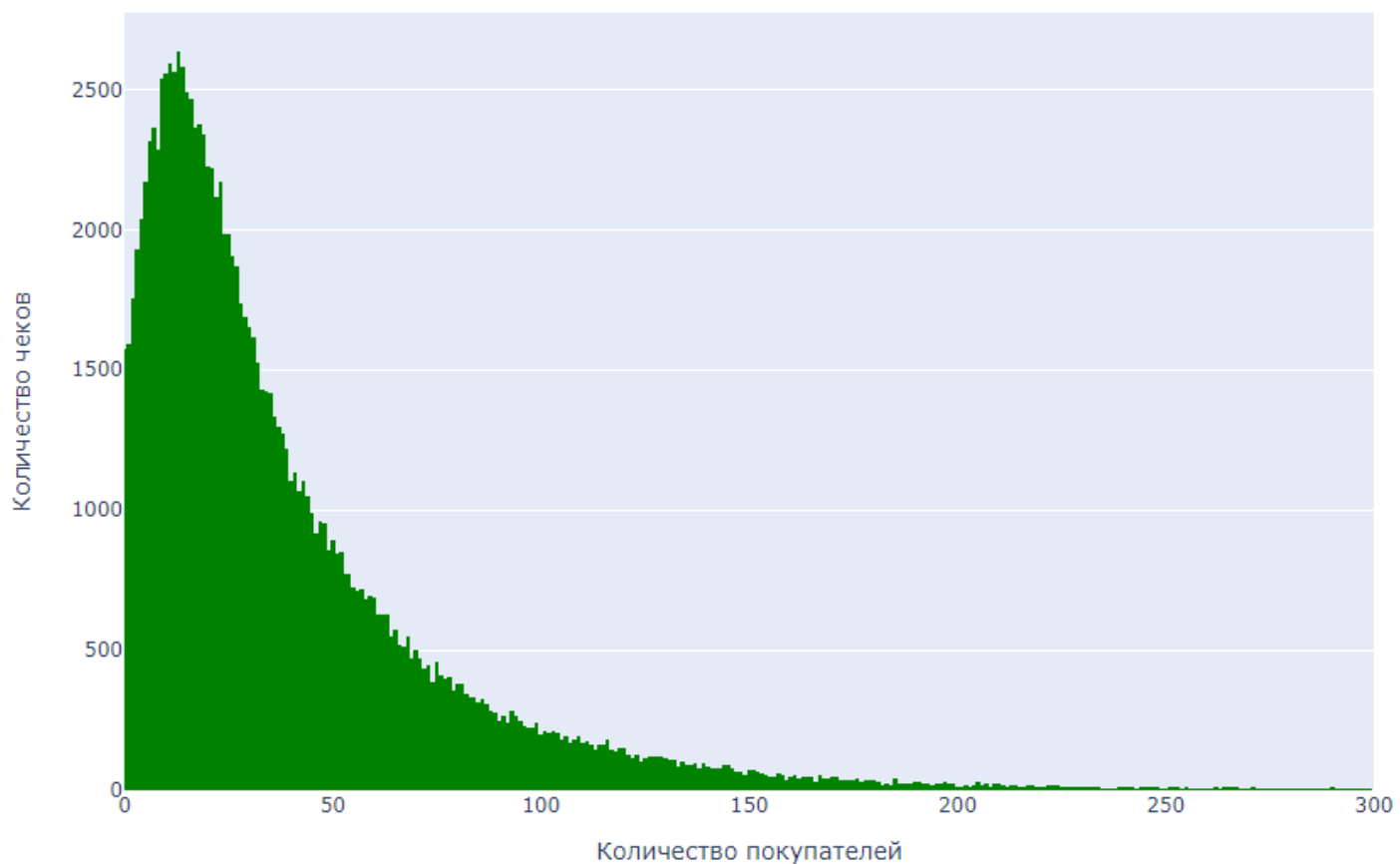
Гиперпараметры

Рекомендации

EDA

Метрики

Распределение суммарного количества чеков



min	39.6
std	40.4
min	0
5%	3
25%	13
50%	52
75%	52
95%	118
max	1114

Гипотезы

Гиперпараметры

Рекомендации

EDA

Метрики

Обучали на тренировочной части датасета, который получили путём деления всего датасета в соотношении 70:30 с перемешиванием и стратификацией.

Подбор занял 500 итераций или 3 часа реального времени.

Обучение модели на датасете из 124 тысяч строк и 234 признаков занимает 22,6 секунды.



Для подбора гиперпараметров был использован open-source framework - Optuna

Гипотезы

Гиперпараметры

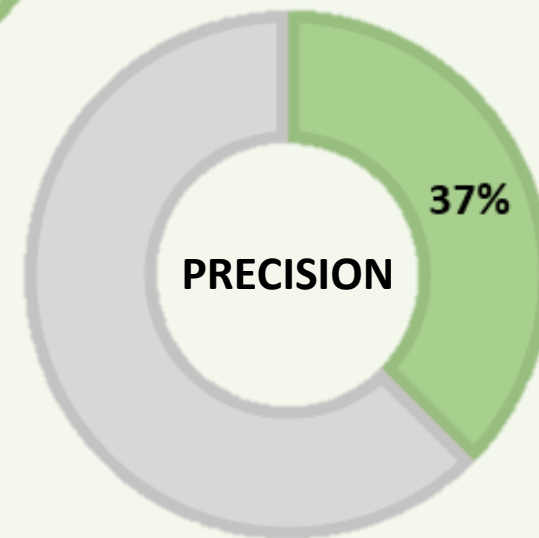
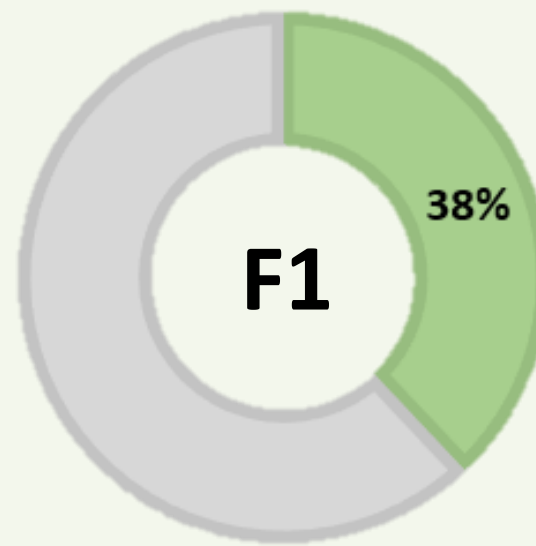
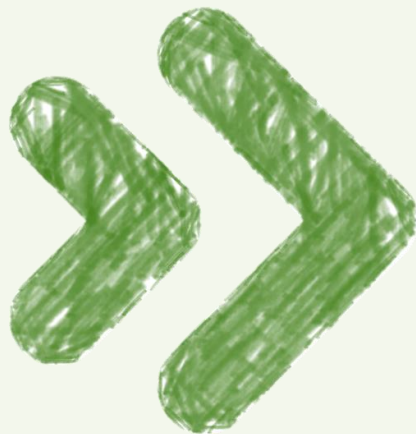
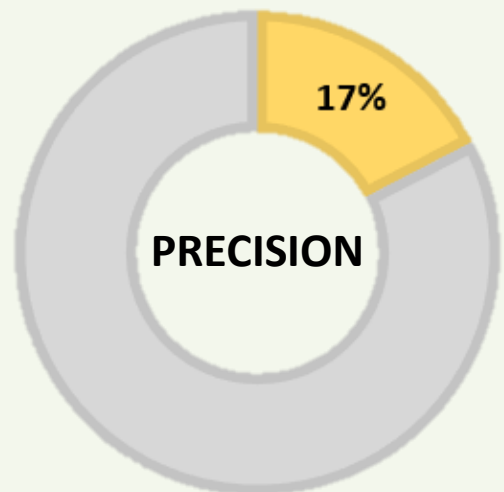
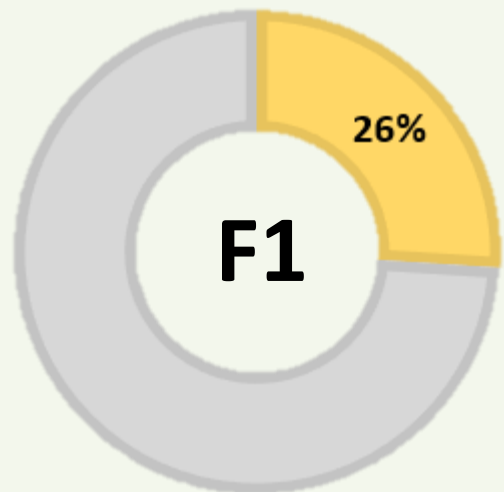
Рекомендации

EDA

Метрики

Метрики модели первого тура

Метрики финальной модели



Итоговая модель была улучшена
на 12% по F1 и на 20% по Precision

Гипотезы

Гиперпараметры

Рекомендации

EDA

Метрики

Отброшенные значения

4917 покупателей с количеством чеков < 3 за 7 месяцев

Решение:

Завлечь таких покупателей выгодными предложениями, баллами и скидками.

6099 покупателей с количеством чеков > 118 за 7 месяцев

Решение:

Покупателям с огромным количеством чеков незамедлительно предлагаем вступить в клуб.

Информируем на кассах самообслуживания, смс сообщениями или рассылкой на почту.

Варианты взаимодействия с целевой аудиторией (предложение вступить в клуб):

- Через смс и письмо на почту, указанные при регистрации;
- После оформление товаров через кассу самообслуживания при использовании карты лояльности;
- Информирование о возможности вступить в клуб на чеках.

После анализа, предобработки и обучения модели мы смогли предсказать, кто из покупателей является потенциально возможным членом клуба здоровых привычек.

Решение состоит из:

- исполняемого кода в виде файла Jupyter Notebook
- файла .csv с предсказанием по пользователям
- презентации в формате .pdf
- readme файл



Спасибо за внимание!



Сведения о команде

Drill Team

Бардин Федор
Москва
РГУ нефти и газа им
И.М. Губкина
Разработка нефтяных
и газовых
месторождений, 2021

Контакты:
fedya_bardin@mail.ru,
+7-906-794-79-33

Безвытный Станислав
Санкт-Петербург
СПб ГУАП
Институт
радиотехники,
электроники и связи,
2019

Контакты:
stasbezv@gmail.com,
+7-921-772-32-35

Зиязьев Ильназ
Москва
РГУ нефти и газа им И.М.
Губкина
Разработка нефтяных и
газовых месторождений,
2021

Контакты:
ilnazziaziev@gmail.com,
+7-937-576-45-87

Богуш Иван
Москва
РГУ нефти и газа им И.М.
Губкина
Разработка нефтяных и
газовых месторождений,
2021

Контакты:
Ivan.s.bogush@gmail.com,
+7-916-975-16-61