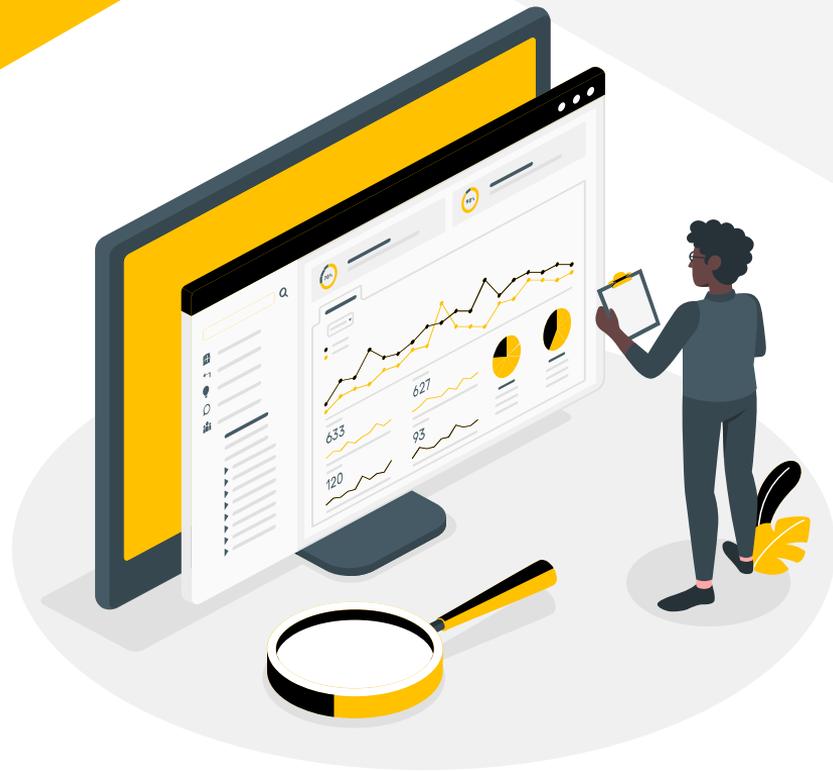


Анализ данных

РАНО



Команда «ленашики»

Исследовательский вопрос

Какие характеристики клиента определяют, какой вид кредитования для него наиболее актуален?

данные

гипотеза

обработка

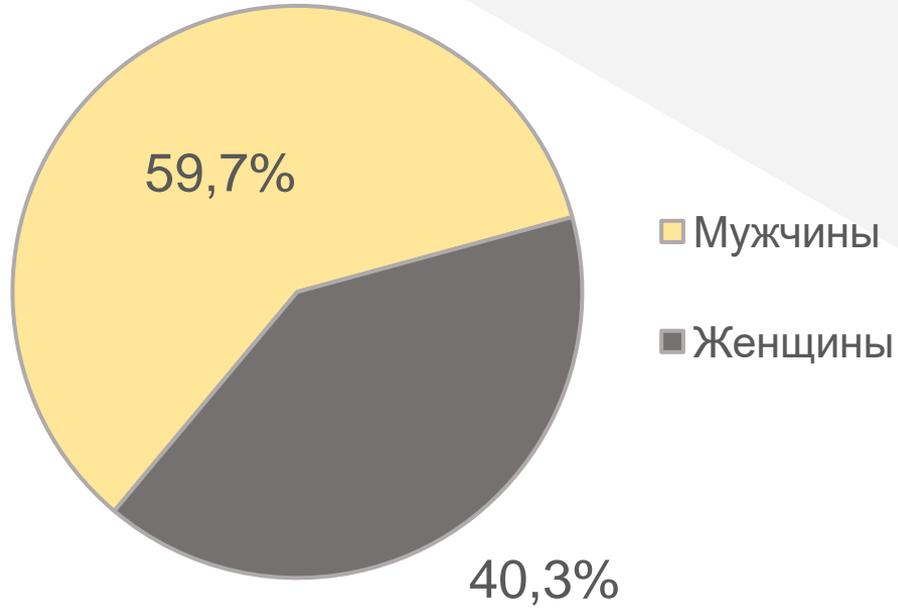
проверка

выводы

Структура базы данных

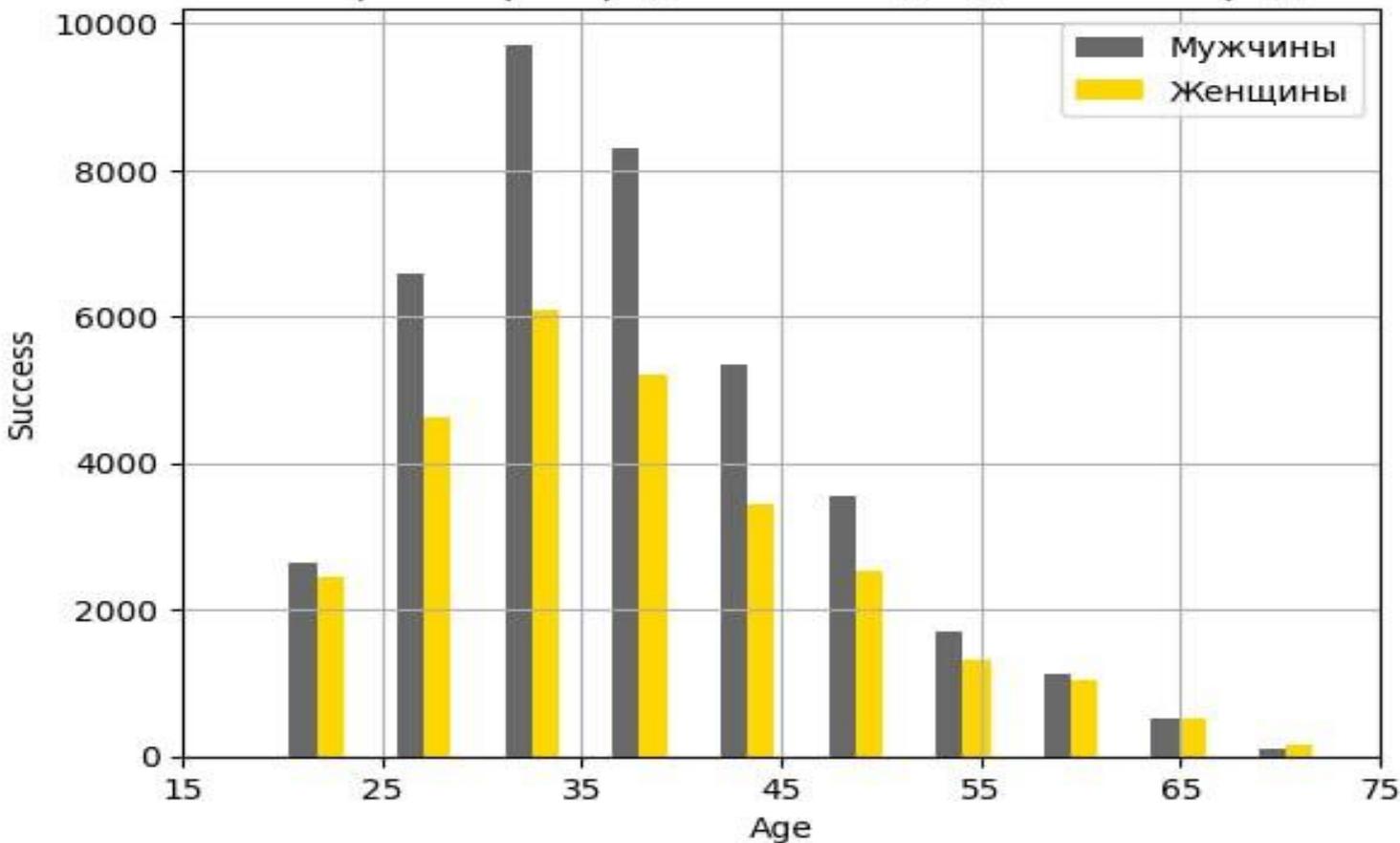
| Количественные | | Качественные | |
|--------------------|------------------------------|------------------------|--|
| req_amt | Запраш. сумма кредитования | req_product_code | код запрошенного продукта кредитования |
| req_term | Запраш. срок кредитования | gender_code | пол клиента |
| monthly_income_amt | Доход клиента в месяц | car_flg | наличие авто у клиента |
| age | Возраст | realty_flg | наличие недвижимости у клиента |
| car_price | Цена авто клиента | cc_flg | наличие кредитной карты у клиента |
| realty_price | Цена недвиж. клиента | agreement_product_code | код согласованного клиентом продукта |
| account_amt | Согл. сумма открытого счета | success_activate_flg | успешная активация кредита по заявке |
| account_amt | Согл. срок открытого счета | | |
| account_rate | Согл. ставка открытого счета | | |

Предварительный анализ



Предварительный анализ

Гистограмма распределения людей, кто взял кредит



Предварительный анализ

Корреляционная матрица

| | Unnamed: 0 | req_amt | req_term | monthly_income_amt | age | gender_code | car_flg | car_price | realty_flg | realty_price | cc_flg | agreement_product_code |
|------------------------|------------|-----------|-----------|--------------------|-----------|-------------|----------|-----------|------------|--------------|-----------|------------------------|
| Unnamed: 0 | 1.000000 | -0.003928 | -0.004204 | 0.001503 | -0.001578 | -0.009536 | 0.006751 | -0.006832 | 0.001964 | 0.002081 | -0.001711 | 0.001398 |
| req_amt | -0.003928 | 1.000000 | 0.321617 | 0.171232 | 0.116564 | 0.036208 | 0.059420 | 0.228981 | 0.099588 | 0.084471 | -0.038518 | 0.072697 |
| req_term | -0.004204 | 0.321617 | 1.000000 | 0.026286 | -0.018701 | 0.012964 | 0.039734 | 0.008615 | 0.069218 | -0.012162 | 0.023405 | 0.098592 |
| monthly_income_amt | 0.001503 | 0.171232 | 0.026286 | 1.000000 | -0.005269 | 0.071799 | 0.094824 | 0.269322 | 0.098906 | 0.111089 | 0.002425 | 0.068253 |
| age | -0.001578 | 0.116564 | -0.018701 | -0.005269 | 1.000000 | -0.035848 | 0.088020 | 0.061769 | 0.040468 | 0.060864 | -0.046703 | 0.081967 |
| gender_code | -0.009536 | 0.036208 | 0.012964 | 0.071799 | -0.035848 | 1.000000 | 0.218317 | 0.007040 | -0.032577 | 0.010004 | 0.035289 | 0.102437 |
| car_flg | 0.006751 | 0.059420 | 0.039734 | 0.094824 | 0.088020 | 0.218317 | 1.000000 | nan | 0.052613 | 0.023085 | 0.003873 | 0.284252 |
| car_price | -0.006832 | 0.228981 | 0.008615 | 0.269322 | 0.061769 | 0.007040 | nan | 1.000000 | 0.130786 | 0.145721 | -0.049036 | 0.056690 |
| realty_flg | 0.001964 | 0.099588 | 0.069218 | 0.098906 | 0.040468 | -0.032577 | 0.052613 | 0.130786 | 1.000000 | nan | -0.029305 | 0.217665 |
| realty_price | 0.002081 | 0.084471 | -0.012162 | 0.111089 | 0.060864 | 0.010004 | 0.023085 | 0.145721 | nan | 1.000000 | -0.000465 | -0.015045 |
| cc_flg | -0.001711 | -0.038518 | 0.023405 | 0.002425 | -0.046703 | 0.035289 | 0.003873 | -0.049036 | -0.029305 | -0.000465 | 1.000000 | -0.088012 |
| agreement_product_code | 0.001398 | 0.072697 | 0.098592 | 0.068253 | 0.081967 | 0.102437 | 0.284252 | 0.056690 | 0.217665 | -0.015045 | -0.088012 | 1.000000 |

Предварительный анализ



Предварительный анализ



Предобработка данных

Изменение данных

- Есть машина или недвижимость, но нет стоимости – медианные значения;
- Нет месячного дохода, запрашиваемой суммы, возраста, периода выплаты – медианные значения;
- Нет пола – мужской;
- В столбце приоритета продаж продуктов на согласовании условий отредактировать данные.

- Строки, где код согласованного с клиентом продукта начинается с Т0...;
- Строки, где кредит не активирован (кроме кредитной карты)

Выбросы данных

данные

гипотеза

обработка

проверка

выводы

Механизм

Клиент запрашивает
кредит

Банк анализирует данные
пользователя

На основе данных банк
предлагает выгодные для
себя вид кредитования

Согласование условий

отказ

Кредит выдан



Какие характеристики клиента определяют, какой вид кредитования для него наиболее актуален?

ГИПОТЕЗА

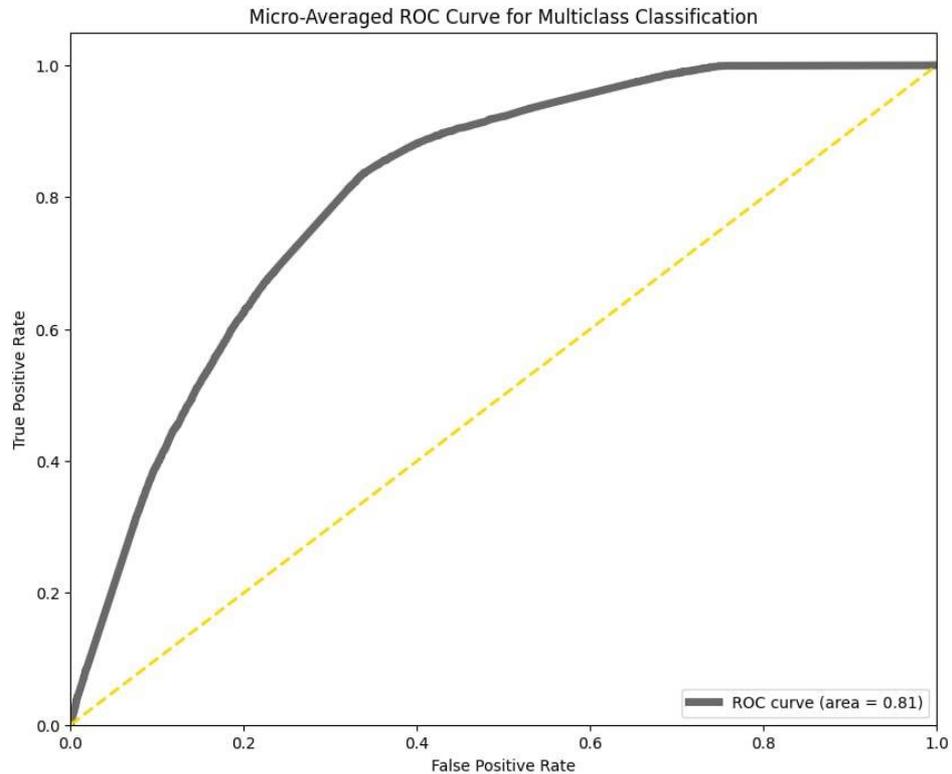
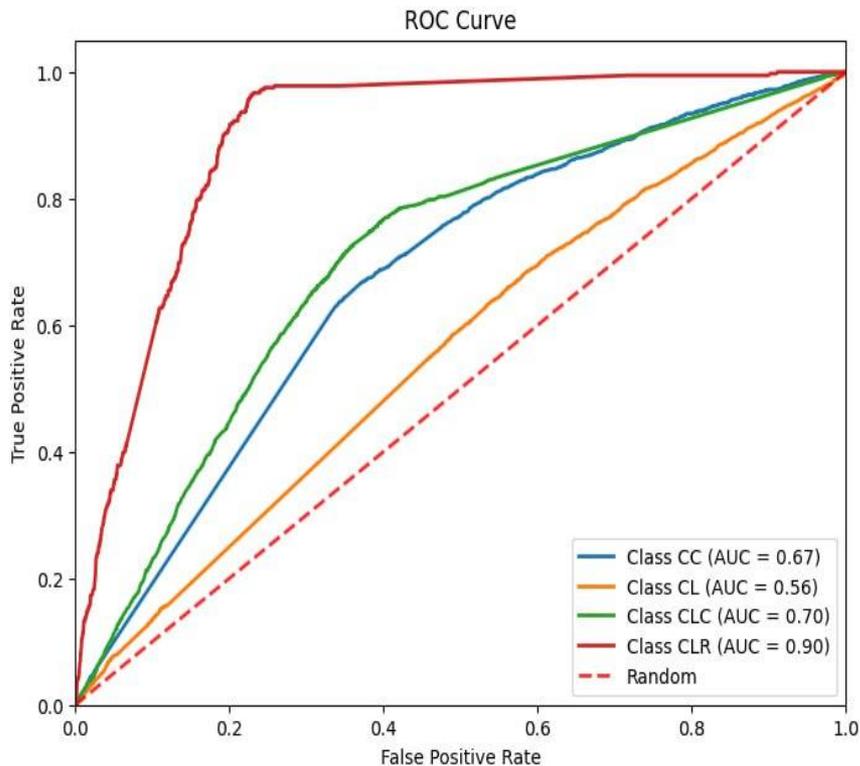
Наиболее актуальный для клиента вид кредитования определяется по таким характеристикам, как стоимость недвижимости и стоимость автомобиля

МОТИВАЦИЯ

Благодаря сокращению количества данных, которые нужно анализировать, можно увеличить скорость принятия банком решения о том, какой вид кредитования подходит клиенту.

Анализ

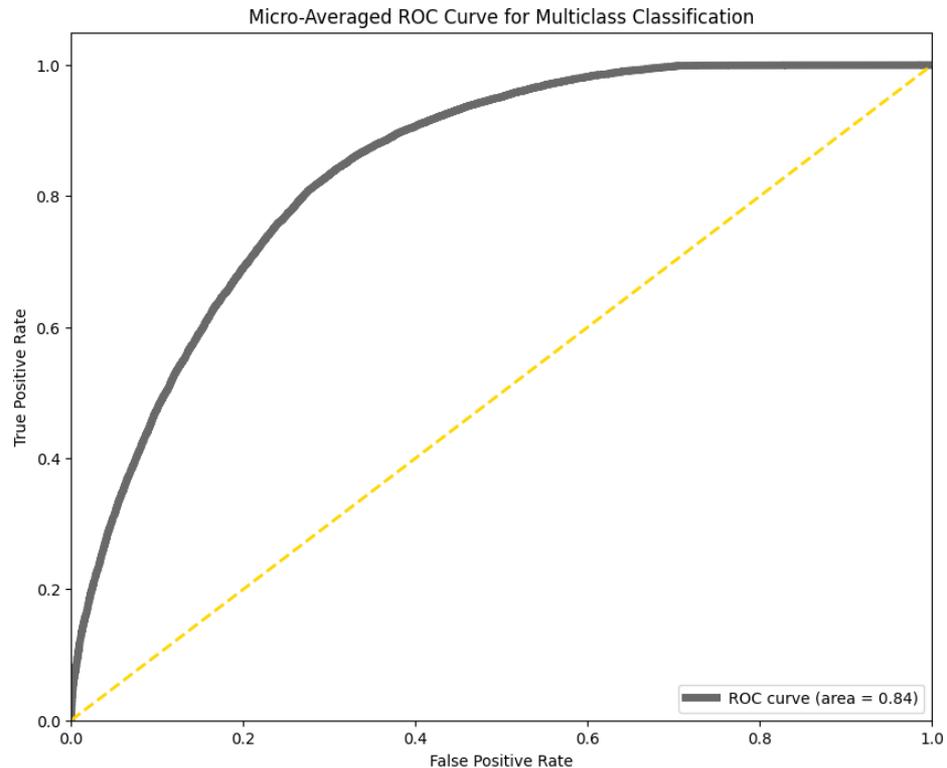
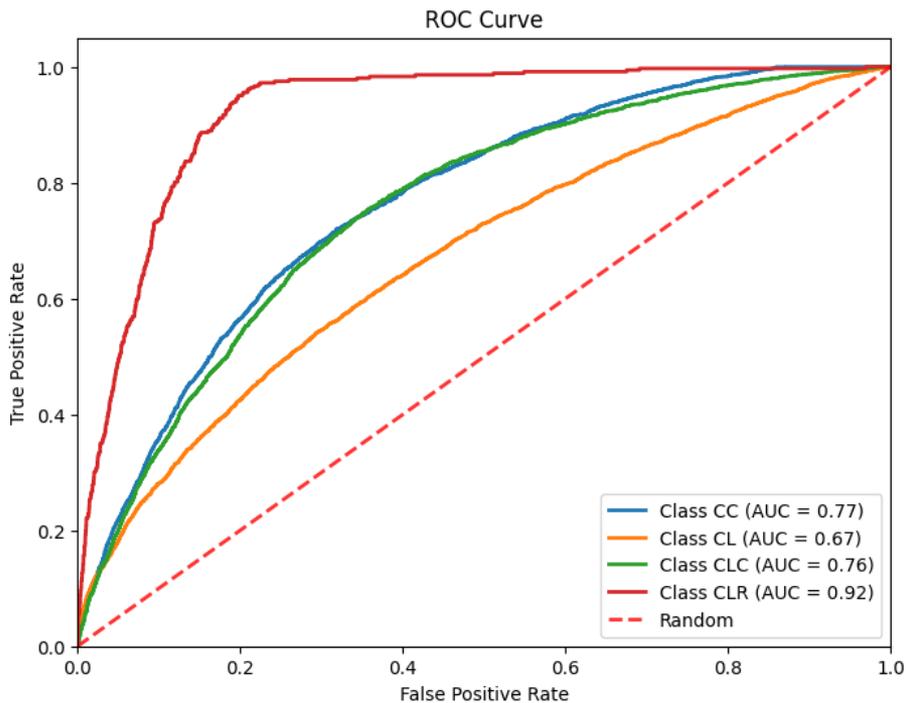
Вывод из гипотезы



Анализ. Уточнение данных

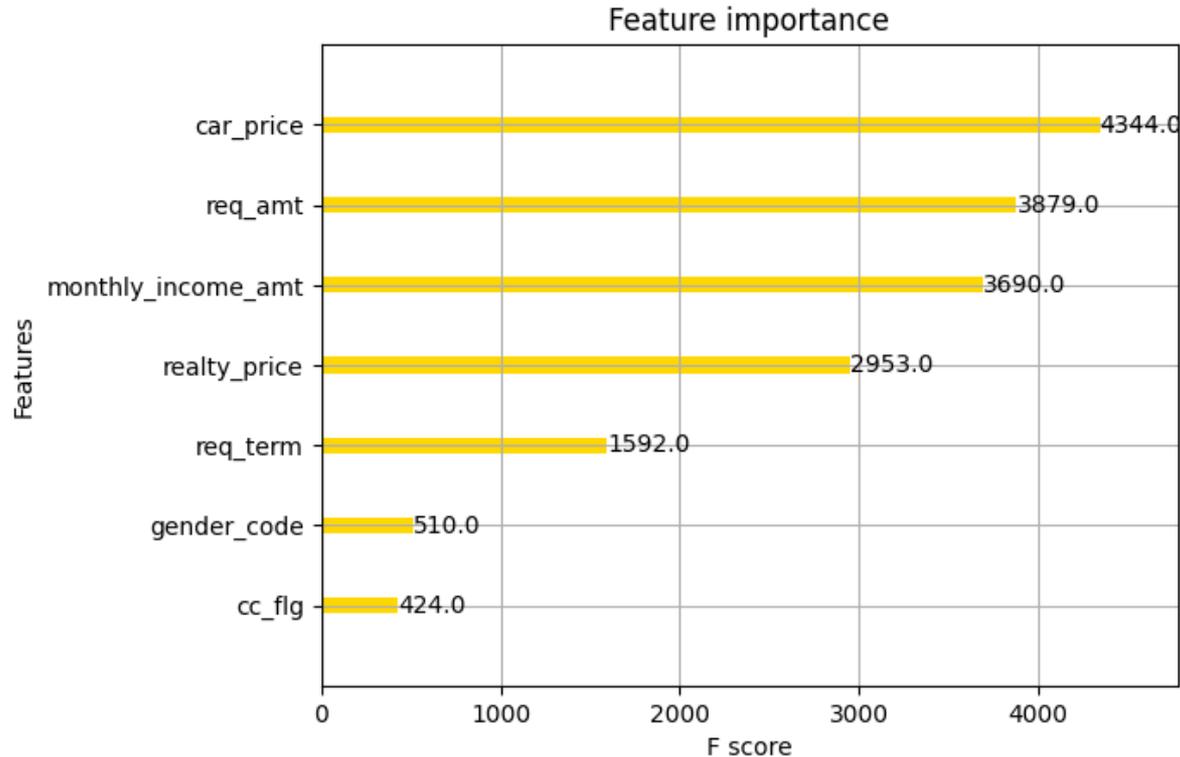
Добавили зависимость от: пола, месячного дохода, наличия кредитной карты, запрашиваемой суммы, срока кредитования.

ИТОГ: +3%



ИТОГОВЫЙ ВЫВОД

Наибольшая зависимость от стоимости машины, стоимости недвижимости и
месячного дохода



Рекомендации

При поступлении новых клиентов собирать сведения о стоимости машины и недвижимости

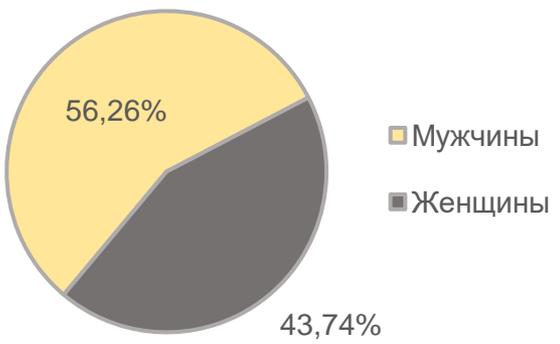
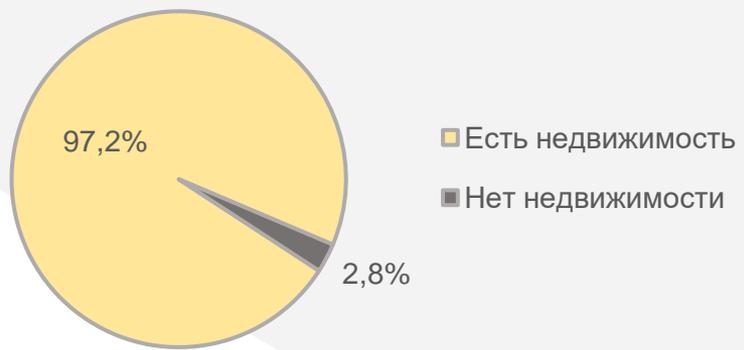
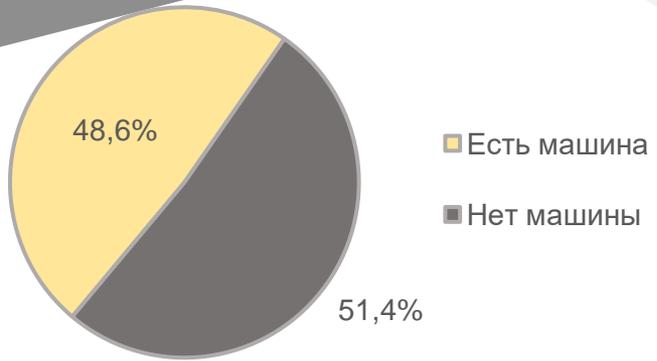
Ограничения

- Выбор модели
- Субъективный выбор параметров

Перспективы

- Подбор идеальных параметров
- Увеличение выборки обучающих данных
- Внедрение алгоритма банка

Анализ ЦА кредита с залогом на недвижимость



Наша команда



Капитан команды, разработчик
мат. модели
Большакова Анжелика



Визуализатор
Горохова Ольга



Программист
Молеваник Алексей



Презентер
Фатхуллина Софа

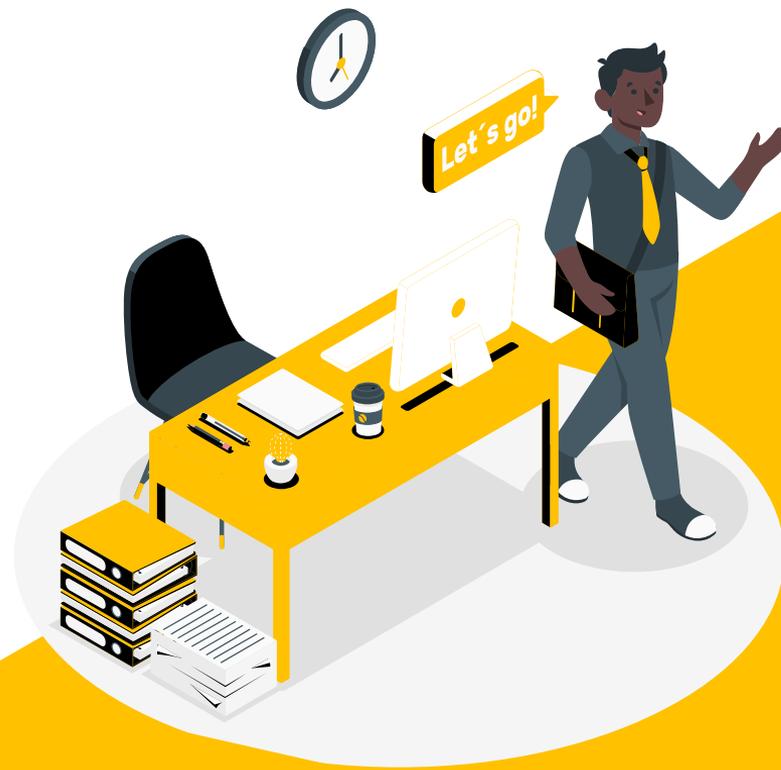


Аналитик
Малышев Виктор



Презентер
Павлухина Анна

СПАСИБО



Команда «ленашики»