

Кейс: «Модель склонности клиента к приобретению машиноместа»



Цель кейса

**Создание модели для поиска
новых клиентов, наиболее
склонным к приобретению
машиноместа**

Команда



Пушмин Александр



Семрягина Анна



Хоружая Вероника



План работы

- Анализ данных и технического задания
- Обработка данных
- Обучение классификационной модели

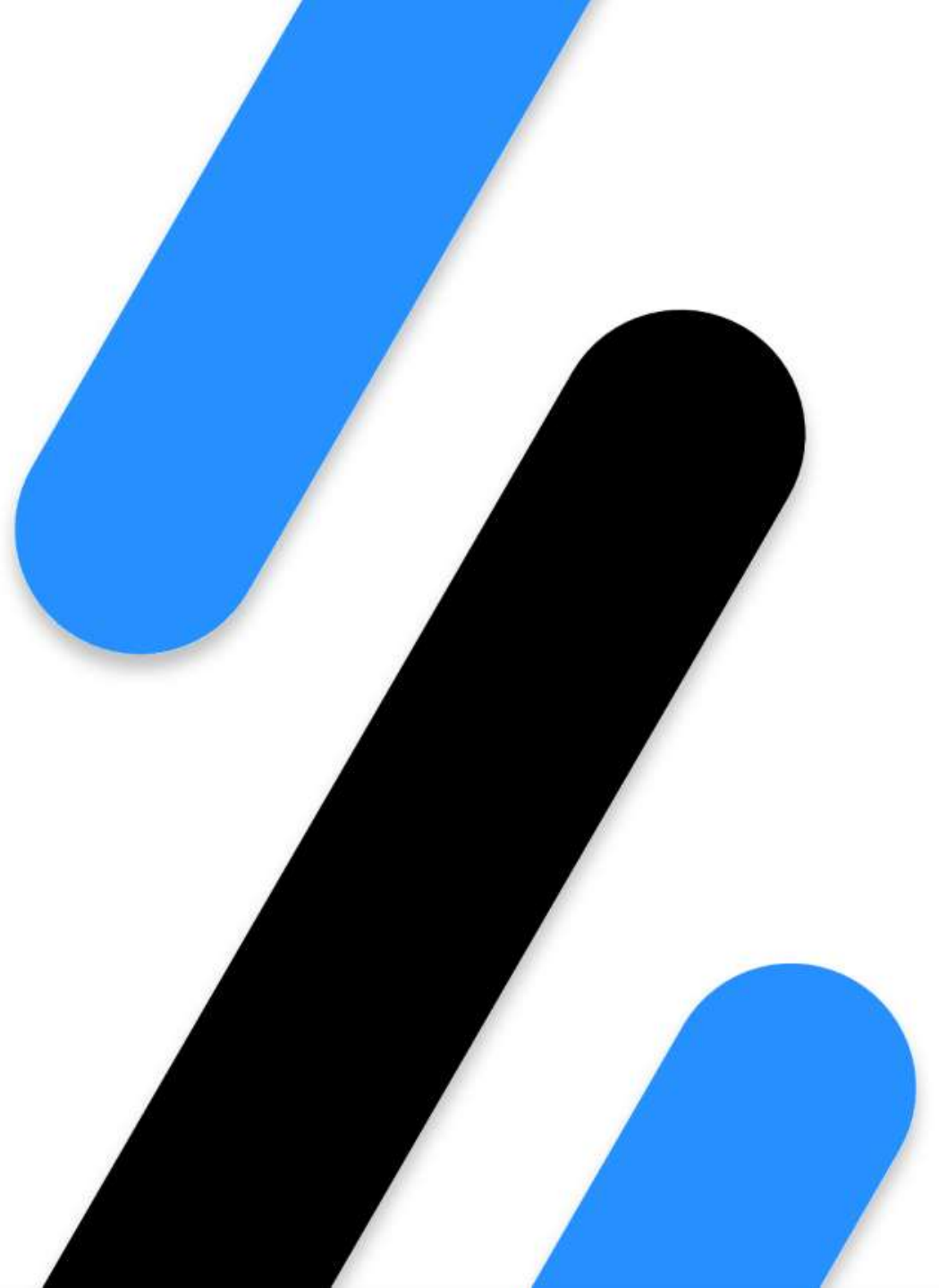
Проверка
модели
метрикой ROC-
AUC

Анализ данных

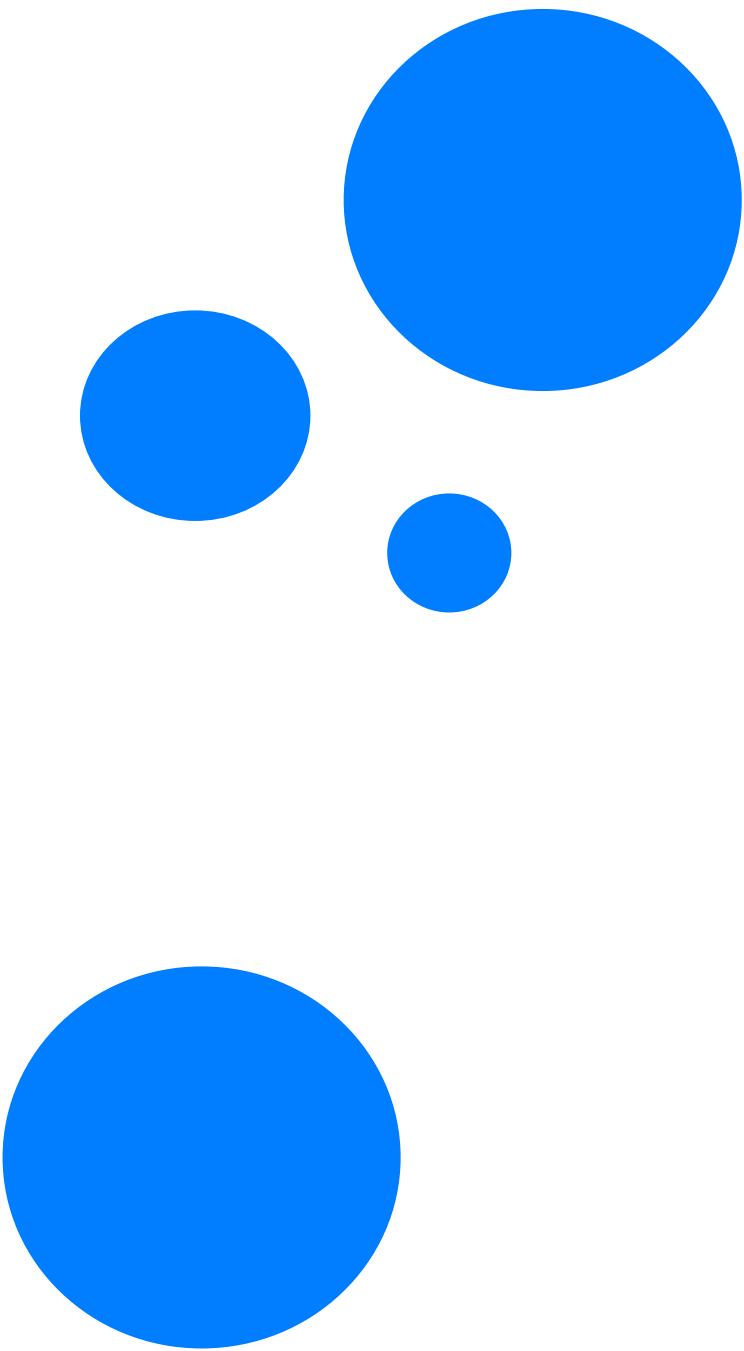
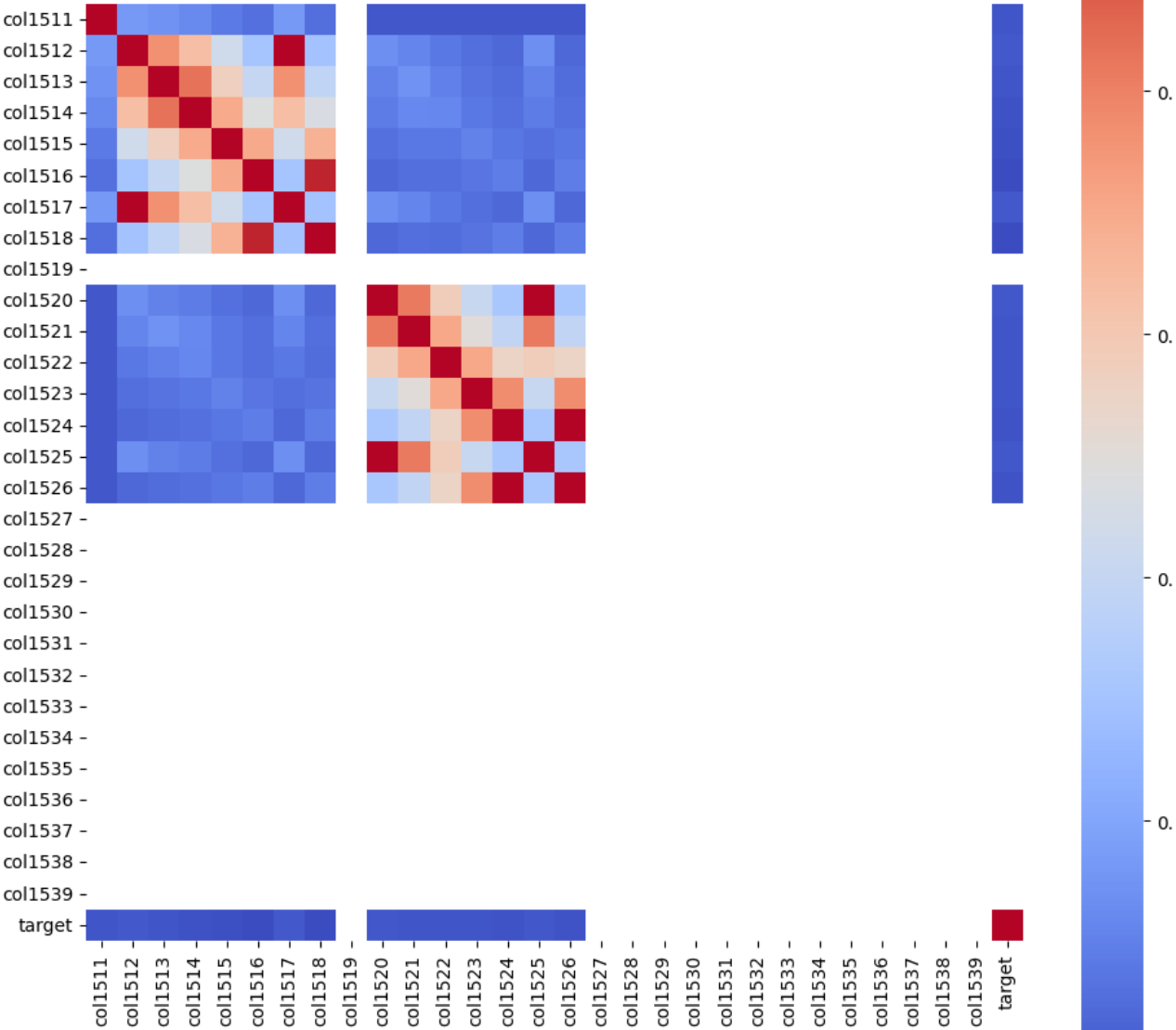
- Оценка качества данных
- Проверка на наличие пропусков, дубликатов и выбросов
- Анализ корректности данных

Обработка данных

- Заполнение пропусков
- Удаление дубликатов
- Обработка выбросов

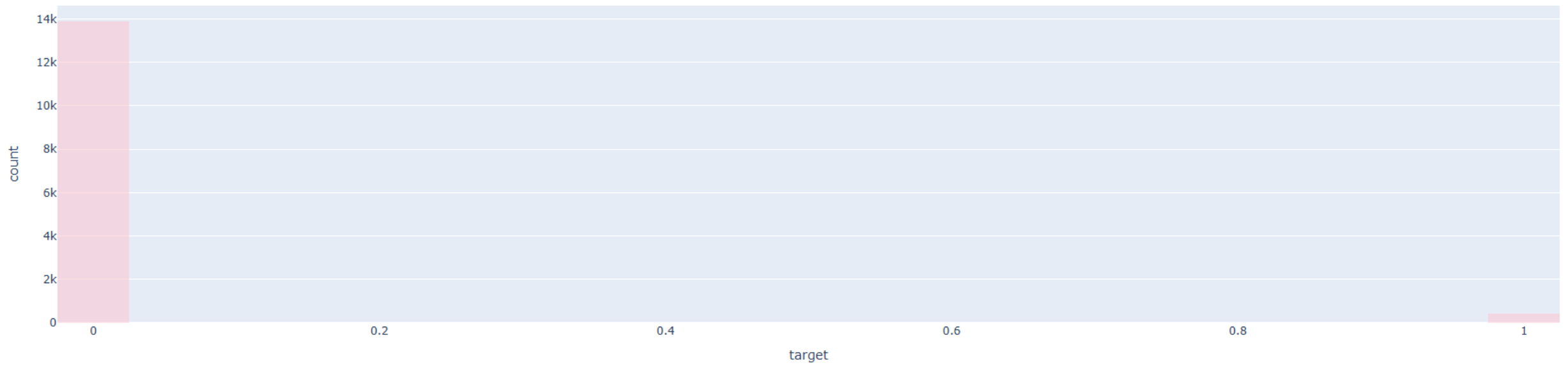


Матрица корреляций



Дисбаланс

Распределение таргета



RandomForestClassifier

Отсутствие необходимости в
масштабировании данных

Высокая точность



CatBoostClassifier

Высокая точность

Простота использования

**Отсутствие необходимости в
гиперпараметрах**

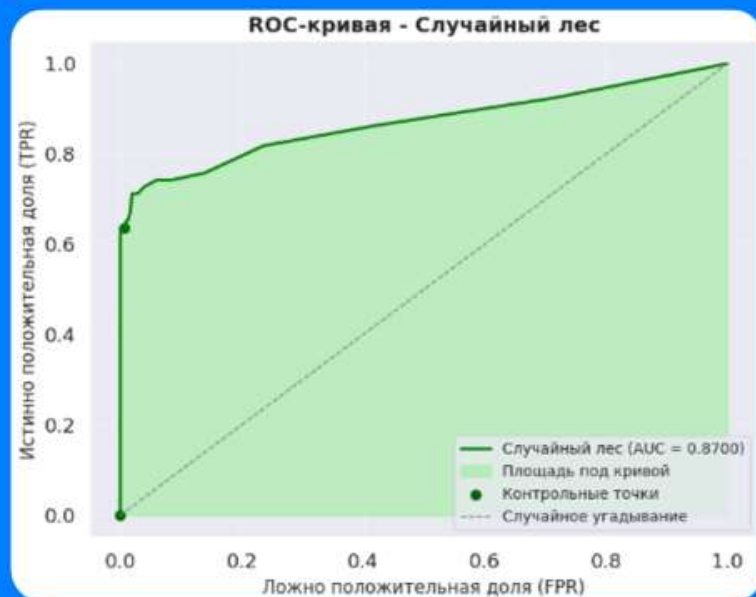
LogisticRegression

Низкая требовательность к данным

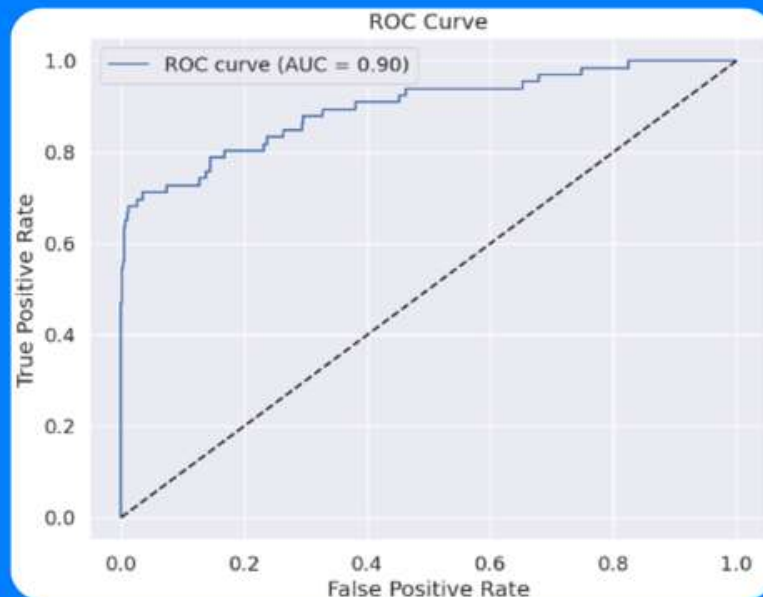
Простота использования

Интерпретируемость

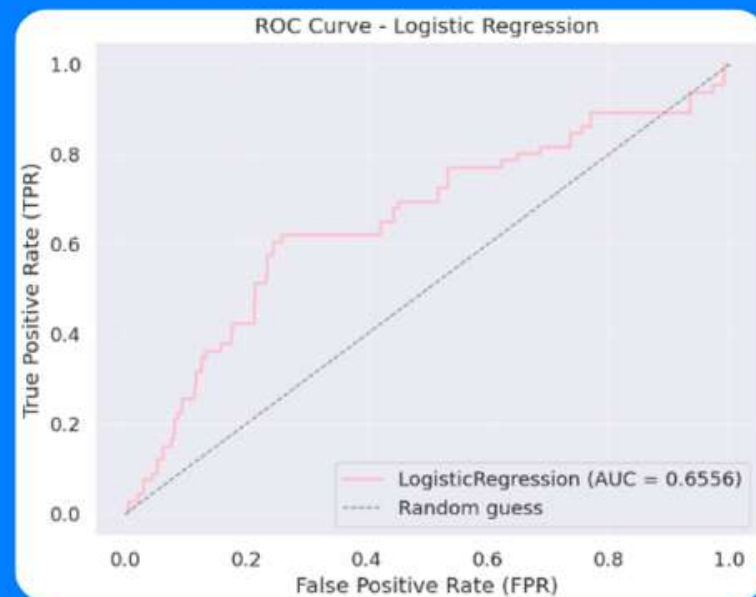
Метрики ROC-AUC



AUC: 0,87



AUC: 0,90



AUC: 0,66

Итоги работы

В процессе анализа мы изучили структуру данных, оценили их качество и проверили корректность.

На основе анализа мы разработали и обучили модель, которая решает задачу с высокой точностью.