**Для решения задачи применялся поэтапный подход:**

* Был произведен обзор и анализ предоставленных данных (Data analysis).
* Определены и реализованы алгоритмы масштабирования и формирования признаков (Preprocessing & Feature engineering).
* Произведена минимизация влияния дисбаланса классов при помощи оверсэмплинга – метода синтетического создания образцов меньших классов (Over-sampling).
* В целях выбора модели были рассмотрены градиентные бустинги: XGBoost CatBoost и LightGBM (Machine learning)
* В результате обучена на всех данных модель машинного обучения на основе градиентного бустинга (XGBoost), с помощью которой сформирован результат по тестовой выборке.

**Технические особенности:**

при реализации решения использовался следующий стек Open Source технологий: **python, numpy, pandas, matplotlib, seaborn, sklearn, xgboost**

**Уникальность:**

Решение реализовано с помощью оригинальных алгоритмов масштабирования и формирования признаков для обучающей выборки (Preprocessing & Feature engineering). Обученная на «обогощенных», в результате оверсэмплинга, данных модель демонстрирует увеличение показателя по метрике Recall более чем на 10 процентных пунктов по сравнению с таким же показателем модели, обученной на данных без добавления синтетических элементов меньших классов.