

Хакатон Baltic **Alfa** Hack

Оптимизация работы
с корпоративными клиентами
банка на основе больших данных

6 → 8 октября
Офлайн



Участникам хакатона



Внимательно ознакомьтесь с расписанием хакатона.
Если возникнут вопросы – пишите в чат хакатона в телеграме.

Расписание мероприятий:

6 октября

18:00 – Открытие хакатона

19:00 – Публикация заданий и вводная QA-сессия

7 октября

10:00 – Открытие очной площадки для решений

12:00 – Брифинг и QA-сессия

17:00 – Окончание приема посылок решений

18:00 – Объявление команд-финалистов

8 октября

10:30 – Открытие очной площадки

13:30 – Окончание приема презентаций

14:00 – Защиты решений

17:00 – Интерактив и награждение

17:30 – Сессия обратной связи

Вся информация «в прямом эфире»
публикуется здесь:

Информационный канал
Хакатона



Расписание мероприятий



6 октября

18:00 – Открытие хакатона
Акт. зал

**19:00 – Публикация заданий
и вводная QA-сессия**



7 октября

**10:00 – Открытие очной
площадки для решений**

**12:00 – Брифинг и QA-
сессия**

**17:00 – Окончание приема
посылок решений**

**18:00 – Объявление
команд-финалистов**

8 октября

**10:30 – Открытие очной
площадки**

**13:30 – Окончание приема
презентаций**

14:00 – Защиты решений

**17:00 – Интерактив и
награждение**

**17:30 – Сессия обратной
связи**

Постановка задачи: 1 этап

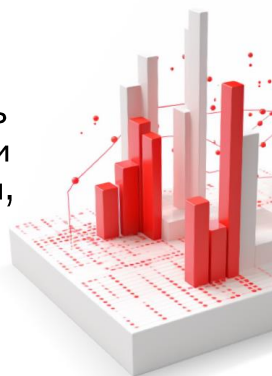


На клиентских данных для юридических лиц необходимо спрогнозировать отток клиентов из банка. Под оттоком подразумевается прекращение финансовой активности клиента или закрытие РКО*.

1 этап: 19:00 6.10 – 17:00 7.10

Изучить данные, обработать и визуализировать.

- Предобработать данные: обработать выбросы и категориальные признаки
- Отобрать наиболее важные факторы, проанализировать на наличие зависимостей между ними и визуализировать



Выполнить задачу классификации клиентов:

- Построить модель, прогнозирующую отток клиента из банка
- Оценить точность предсказания по метрике ROC-AUC на открытой выборке

Задача и датасет

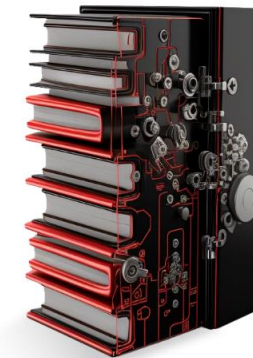
Датасет представляет собой таблицу со 100 факторами и содержит информацию о клиентах и их активностях внутри банка

300 000 записей – обучающая выборка
100 000 записей – тестовая выборка (public+private)

Целевая переменная:

- Прекращение финансовой активности клиента (Y1) или закрытие РКО(Y2)
- Итоговый таргет $Y = \max(Y1, Y2)$

Оценка решения на первом этапе производится только по метрике ROC-AUC



*Расчетно-кассовое обслуживание

Постановка задачи: 2 этап



На клиентских данных для юридических лиц необходимо спрогнозировать отток клиентов из банка. Под оттоком подразумевается прекращение финансовой активности клиента или закрытие РКО.

2 этап: 18:00 7.10 – 13:30 8.10

- Доработать модель, создать репозиторий с документацией и воспроизводимым решением
- Сформулировать бизнес-инсайты по результатам визуализации и моделирования. Определить наиболее интерпретируемые для бизнеса метрики, с помощью которых оценить модель.
- Презентовать свое решение и результаты перед членами жюри: подготовить выступление на 5-7 минут

Презентация решения

- Описать подход к решению задачи и полученные результаты
- Подготовить визуализацию данных
- На основе построенной модели сформулировать полезные и понятные для бизнеса выводы

Критерии оценки решений на финале

Техническая сторона решения:

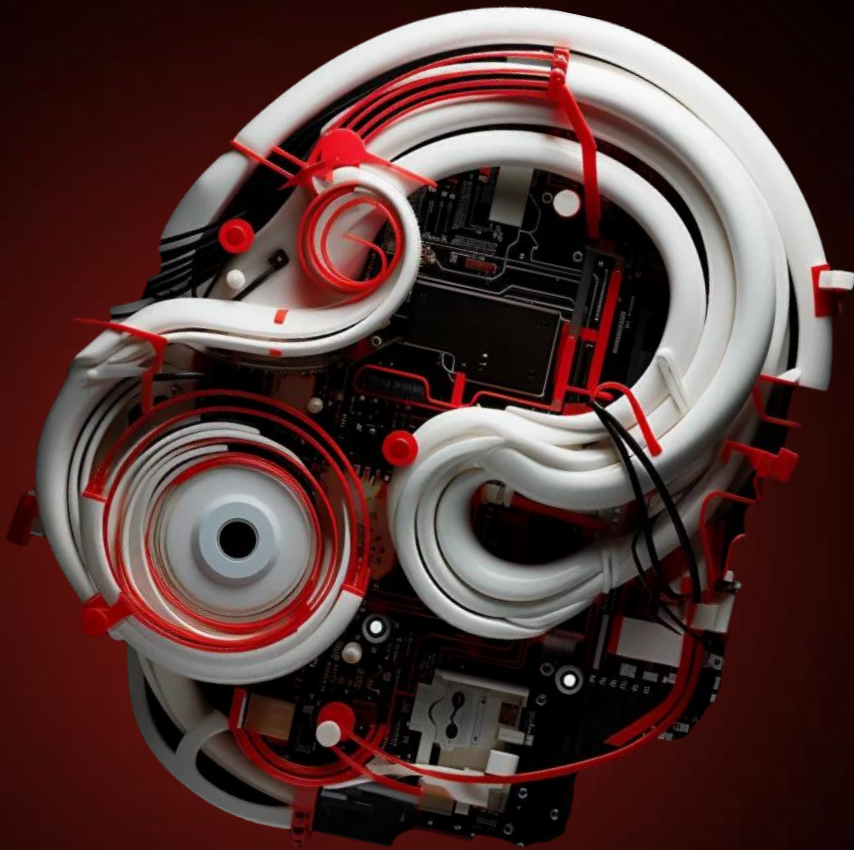
- Значение метрики ROC-AUC
- Воспроизводимость и читабельность кода – возможна проверка при защите решений
- Широта подхода: использование различных ML-алгоритмов, метрик качества модели

Бизнес-сторона:

- Качество и обоснованность инсайтов (выводов), полученных по результатам создания модели
- Оригинальность подхода и защиты решения
- Качество визуального оформления и логика презентации

Ваши решения будут оценены по нескольким критериям командой экспертов Альфа-Банка и БФУ.

Платформа сбора и оценки решений А



Ссылка на
соревнование

Я Контест

Тестовая выборка состоит из 2х частей: public и private

Ответ на задание – *.csv файл с предсказанием оттока на тестовой выборке

Во время первого этапа участники ранжируются по лучшему результату на public части тестовой выборки

Итоги первого этапа подводятся на private тестовой выборке по последней попытке команды.

Важно отправить нужное вам решение последней попыткой перед завершением 1 этапа!

Полезные материалы



Машинное обучение

Гайды:

- [Визуализация данных](#)
- [Генерация признаков](#)
- [Метрики качества](#)
- [Отбор признаков](#)
- [Машинное обучение для начинающих](#)

Python библиотеки:

- Обработка данных – [NumPy, pandas](#)
- Визуализация – [matplotlib, seaborn](#)
- Построение модели – [scikit-learn, LightGBM, CatBoost, XGBoost, PyTorch](#)

Специфика данных

- [Канал Центра продвинутой аналитики Альфа-Банка](#) с информацией о специфике применения моделей в банковской сфере
- [Видео от Альфа-Банка](#) о решении задач склонности и оттока с помощью нейронных сетей

- [Организация рабочего процесса](#)
- [Интересный подход к решению задачи оттока](#) с проработанной бизнес-постановкой задачи
- [Соревнование по предсказанию оттока пользователей](#) для тренировки и вдохновения



Дисклеймер

Команда оператора Хакатона Talent Case подготовила данный кейс по заказу АО «Альфа-Банк».

Все решения, полученные в процессе проведения Хакатона Baltic Alfa Hack, являются собственностью АО «Альфа-Банк». Персональные данные, приведенные в кейсе, могли быть изменены или выдуманы с целью сохранения конфиденциальности данных.

Информационный канал
Хакатона

Платформа Хакатона —
Яндекс Контекст