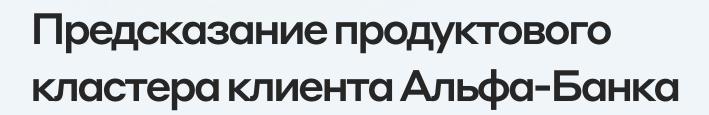
Хакатон Альфа Кампус Джуниор 2025





Введение



Mодель CLTV (Client Lifetime Value) позволяет определять ценность клиента – Юридического лица для Банка на всем сроке его жизни (периоде взаимодействия клиента с Банком).

В качестве показателя ценности клиента используется <u>операционная прибыль</u> – сумма всех доходных и расходных операций.



Продуктовый кластер

Нам известно, что основной характеристикой клиента, влияющей на его прибыльность, является набор продуктов, которыми клиент активно пользуется, то есть генерирует операционную прибыль.

Продуктовый кластер – набор (множество) продуктов клиента.



CLTV-модель

При изменении кластера клиента, например, при открытии нового продукта или закрытии существующего, у него существенным образом меняется профиль доходности.

Ядром модели CLTV будет являться модель, которая предсказывает переход клиента из одного продуктового кластера во все возможные кластеры на некотором горизонте прогнозирования.



Цель: как можно более точно предсказать продуктовый 🤲 кластер клиента на горизонте 12 месяцев.

Постановка задачи



На реальных данных построить модель, которая будет предсказывать продуктовый кластер клиента – юридического лица на горизонте в 12 месяцев.



Задание отборочного этапа

<u>Цель:</u> создать модель, которая будет выдавать вероятности перехода в каждый из 17 продуктовых кластеров.

<u>Целевая переменная:</u> продуктовый кластер, в котором клиент будет находиться через год – end_cluster.

Необходимо получить <u>вероятности перехода</u> клиента в продуктовые кластеры для последнего месяца (month 6).



Данные и материалы

Вам доступны:

- Тренировочный и тестовый датасеты табличные данные об основных характеристиках компании и транзакционной активности, а также информация о текущем кластере клиента.
- Тренировочный датасет содержит данные о 200 000 клиентах банка и их целевых переменных за три последовательных месяца (month_1, month_2, month_3).
- Тестовый датасет записи о 100 000 клиентах за 3 последовательных месяца (month_4, month 5, month 6).
- Базовое решение (jupyter notebook) и пример корректной посылки.

Постановка задачи



На реальных данных построить модель, которая будет предсказывать продуктовый кластер клиента – юридического лица на горизонте в 12 месяцев.



Обратите внимание

- В задаче присутствует важная деталь стартовый кластер клиента start_cluster (то есть тот набор продуктов, которые есть у клиента на момент прогноза) не заполнен для месяца прогноза (month_6).
- В реальных задачах часто встречаются случаи, когда данные приходят с задержками, связанными с формированием отчетности, и на момент прогноза могут присутствовать еще не все данные. Поэтому вам необходимо каким-либо образом учесть эту особенность при прогнозе.



Оценка и дополнительные файлы

- Оценка работы модели будет проводиться с помощью взвешенной по значимости кластеров метрике ROC-AUC.
- Коэффициенты значимости определяются бизнес критериями и даны в файле «cluster_weights.xlsx».
- Описание признаков приведено в файле «feature_description.xlsx».

Информация для участников



Ключевые даты Хакатона:

- **7 мая, 19:00**Старт отборочного этапа
- 18 мая 23:00 Дедлайн отборочного этапа на платформе
- 19 мая
 Подведение итогов и определение 7 финалистов
- 19 24 мая
 Доработка решений и подготовка презентаций
- **23 мая, 12:00**Дедлайн по отправке исходного кода решения
- 25 мая
 Очная защита финалистов в Конгресс-холле
 Альфа-Банка на Технопарке

Основные правила

1 этап:

Полностью онлайн, оценка решения на платформе Хакатона. Приглашение в финал получат топ-7 команд по рейтингу на платформе.

2 этап:

Заочная подготовка презентаций и очная защита перед экспертами. Результат, полученный на платформе во время первого этапа будет лишь одним из критериев!

Хакатон Альфа Кампус Джуниор 2025

