

Проект №3

Machine Learning

&

Deep Learning



Выполнил:
Озтемел Теоман ББИ-223

Структура данных

Поездки на самокатах от сервиса Т-Банк

Ключевые признаки

- **Клиент:** пол, возраст, уровень образования, семейное положение.
- **Поездка:** расстояние в км, модель самоката, стоимости аренды за минуту стоимость активации поездки.
- **Время:** начало, окончание поездки и время создания заказа.
- **Финансы:** итоговая стоимость поездки, начисленный кэшбэк в рублях, заблокированная сумма на счете клиента

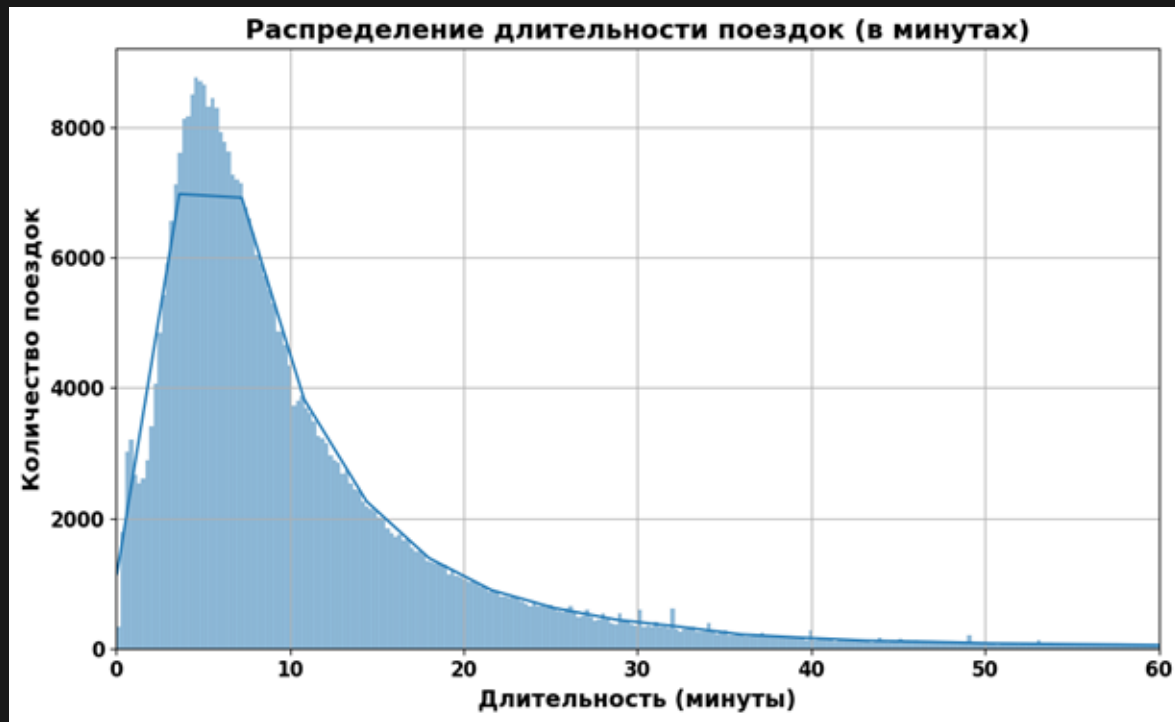
Первичный анализ данных (EDA)

1. Распределение длительности поездок (в минутах)

Большинство поездок 5–10 минут

Сервис чаще используется для коротких перемещений (до метро, магазина)

Длинные поездки – это редкость
скорее прогулки,
нестандартные случаи



Первичный анализ данных (EDA)

2. Время поездок по часам суток

Пик активности: 7–9 и 13–17 часов

Самокаты используют и утром (на работу), и днем (досуг, прогулки)

Вечером и ночью — спрос резко падает



Первичный анализ данных (EDA)

3. Сравнение будни vs выходные

В выходные поездки длиннее

В будни короче и более
утилитарные (едем по делам)

Анализируется время поездки
по типам дня, чтобы понять,
как день недели влияет на
поведение пользователей и
продолжительность поездок.

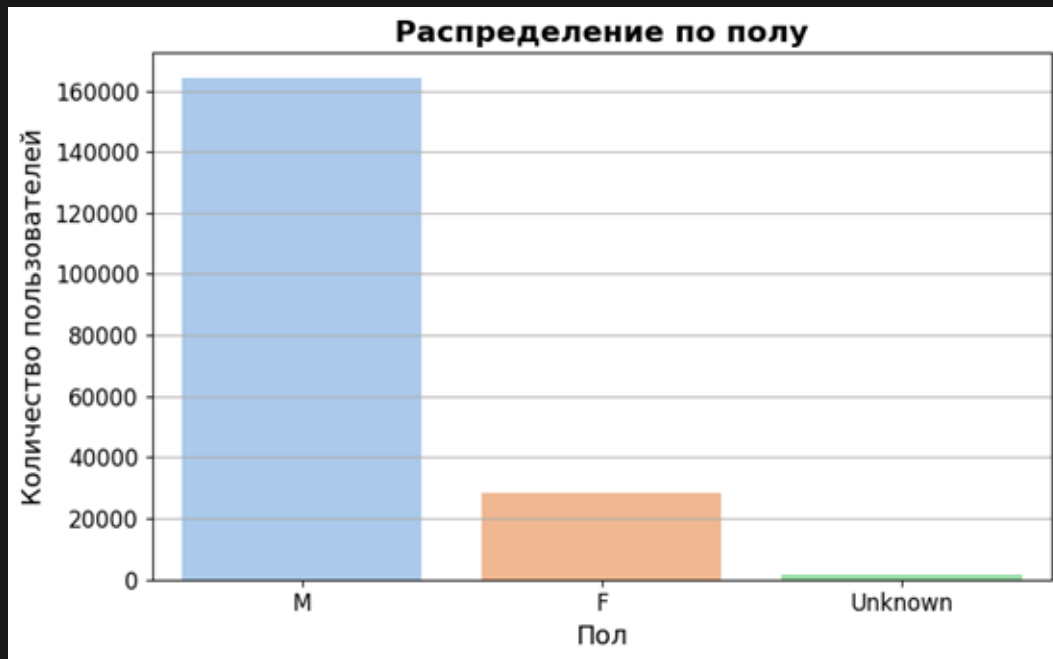


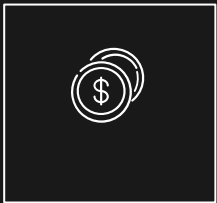
Первичный анализ данных (EDA)

4. Распределение по полу

Большинство пользователей — мужчины; женщин значительно меньше, а доля с неуказанным полом минимальна.

Этот дисбаланс важно учитывать при моделировании и анализе пользовательского поведения.





Бизнес-задача №1 – Сегментация клиентов

Выделить группы пользователей с разным стилем использования самокатов, чтобы проводить персонализированный маркетинг, повышать удержание и выявлять ценные сегменты.

Почему это важно для бизнеса:

- Не все клиенты одинаково ценны: одни ездят редко, другие часто и далеко
- Позволяет **таргетировать рекламу и бонусы**, а также понимать, **кого теряем**
- Упрощает разработку новых тарифов и предложений



Первый сегмент

"Элита / VIP" (около 500 человек)

Самые активные и самые прибыльные клиенты. Пользуются сервисом постоянно, тратят очень много (~13 341 руб).

Стратегия: удержание и эксклюзивность. Этих клиентов нельзя "купить" простой скидкой. Им нужен статус и особое отношение.

Ключевой показатель (KPI): удержание (Churn rate) этого сегмента должен стремиться к нулю.





Второй сегмент

"Лояльные" (около 4200 человек)

Очень активные и прибыльные клиенты (~**3938 руб**), составляют ядро нашей выручки. Их ценность для бизнеса: высокая. Это "золотой фонд" компании.

Стратегия: рост и геймификация. Наша цель — удержать их и мотивировать стать "Элитой".





Третий сегмент

"Стабильные середнячки" (около 26 000 человек)

Самый массовый активный сегмент. Пользуются сервисом регулярно, но без фанатизма. Тратят немного (~600 руб).

Их ценность для бизнеса: основа. Это наша основная аудитория, которую нужно "раскачивать".

Стратегия: вовлечение и Upsell. Наша цель — увеличить их средний чек и частоту поездок.





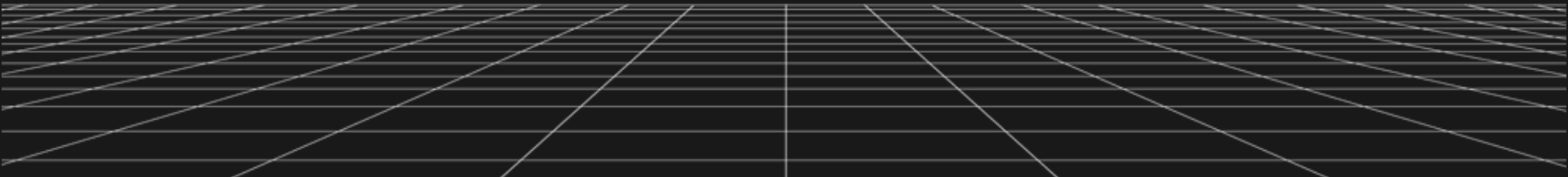
Четвертый сегмент

"Пользователи выходного дня" (около 14 000 человек)

Очень четкий паттерн – катаются почти исключительно в выходные.

Их ценность для бизнеса: стабильный источник дохода в досуговом сегменте.

Стратегия: усиление ниши. Не пытаться заставить их ездить на работу, а стать для них лучшим сервисом для отдыха.





Пятый сегмент

"Пользователи маршрутов до работы" (около 19 000 человек)

Почти неактивны (`days_since_last_trip` ~83 дня). Катались только по будням. Это люди, которые попробовали сервис для поездок на работу и, скорее всего, нашли альтернативу (метро, личный транспорт).

Их ценность для бизнеса: низкая, но это огромная база для реактивации.

Стратегия: реактивация (осторожная). Попытаться вернуть их одним мощным предложением, не тратя много денег.





Бизнес-задача №2 – Прогноз оттока клиентов

Определить, какие клиенты с высокой вероятностью **перестанут пользоваться сервисом**, чтобы заранее среагировать через бонусы, напоминания, персональные предложения и др.

Почему это важно для бизнеса:

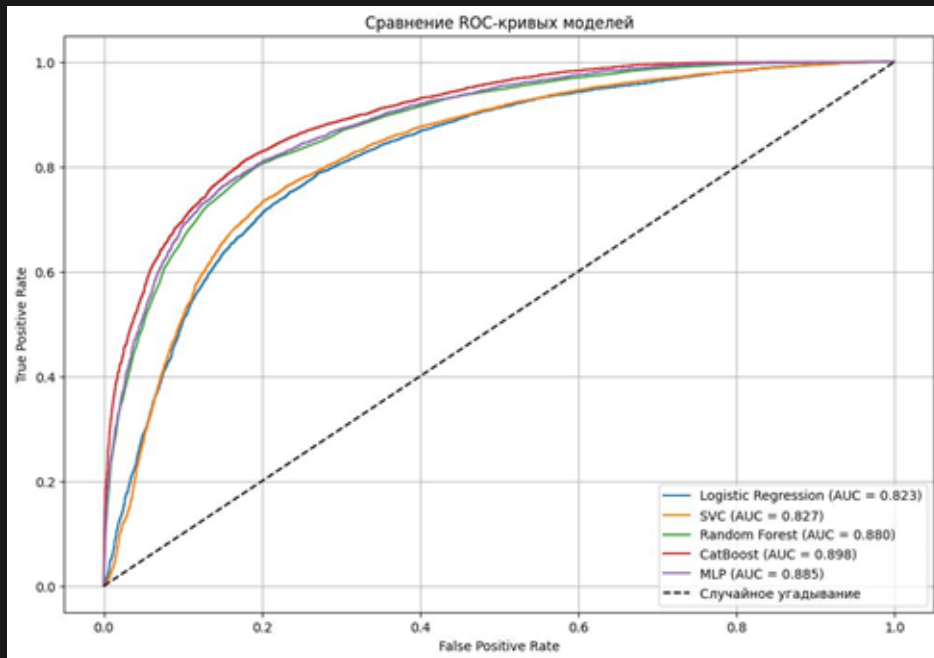
- **До 40–60% новых клиентов совершают только 1 поездку**, а значит влечет за собой большой потенциальный убыток
- Удержать старого клиента дешевле, чем привлечь нового
- Раннее выявление позволяет запускать **кампании по удержанию**

Прогноз оттока клиентов

MODEL	CatBoost	Neural Network (MLP)	Random Forest	SVC	Logistic Regression
F1-score (на обычной)	0.764	0.751	0.724	0.702	0.699
F1-score (с подобранными гиперпараметрами)	0.768	0.751	0.743	0.707	0.697

Прогноз оттока клиентов

ROC-AUC



Выводы для бизнеса (по сегментации клиентов)

Было выделено **5 основных сегментов** клиентов, по которым можно сделать ключевые выводы:

- **Профиль элиты:** активные и прибыльные пользователи – **нужно удерживать** и предоставлять **эксклюзивные предложения**.
- **Лояльные клиенты** — основа для роста, нужно **повышать их активность** через геймификацию и апселлы.
- **Средний сегмент** — необходимо **увеличить их расходы** через активные предложения и вовлеченность.
- **Пользователи выходных** — стабильный источник дохода в досуговом сегменте, для которых важно **не пытаться их заставить ездить на работу**.
- **Пользователи маршрутов до работы:** высокая вероятность **churn**, требуется **реактивация** с помощью целевых предложений.



Рекомендации для бизнеса (по прогнозу оттока клиентов)

Прогнозирование **churn** позволяет **предупредить уход клиентов** и **предложить персонализированные предложения** (скидки, бонусы).

- Для **элитных и лояльных клиентов**: работать на **удержание** и предложить **персонализированные привилегии**.
- Для **"выходных пользователей"**: усилить **программы лояльности**, чтобы сделать их **постоянными пользователями**.
- Для **клиентов с риском оттока**: запустить **акции и напоминания**, направленные на **реактивацию**.

Бонусные рекомендации:

- Использовать **кластеризацию клиентов** для более точного таргетинга.
- Запустить **адаптивную программу удержания**, где клиентам будет предложено то, что соответствует их паттернам поведения.



**Спасибо за
внимание!**

