

Аналитика оттока клиентов компании Gestaro

Галстян Карина

Цель исследования

- Определить ключевые признаки, оказывающие влияние на отток
- Составить портрет «отточного» клиента
- Разобраться в причинах оттока клиентов
- Предложить решения по минимизации оттока

Данные

- Для анализа были предоставлены 4 таблицы, содержащие сведения об активности клиентов, о платежах, о компаниях и о сотрудниках компаний.
- Исследование показало, что в датасетах не совпадают периоды и количество компаний.
- Датасеты содержат плохо структурированную информацию, затрудняющую анализ. В данных есть пропуски и дубликаты. Небольшой размер датасета, содержащего информацию о зарегистрированных компаниях, не позволил провести удаление выбросов, так как это сокращало размер данных более чем на 20%.

Датасет	Период	Число компаний
activities	с января 2019 по март 2024 года	716
billings	с 23 июля 2019 по 15 марта 2024 года	1841
payed_billings	с 23 июля 2019 по 15 марта 2024 года	878
companies	с 23 января 2019 по 7 февраля 2024 года	716 (7 компаний были удалены)
employees	с 23 января 2019 по 13 декабря 2023 года	712

Выводы

- Проведенное исследование позволяет предположить, что в отток чаще уходят маленькие компании с нерегулярным доходом, для которых цена сервиса не компенсируется приносимой пользой. Условно говоря, если в какие-то месяцы у фирмы нет подрядов на ремонты или отделку, платить по 21 тыс. рублей в месяц становится ощутимой финансовой нагрузкой.
- Кроме того, мониторинг отзывов позволяет предположить наличие у сервиса проблем с функциональностью, что также является триггером оттока.

Рекомендации

- Мониторинг негативных отзывов в сети позволяет предположить наличие функциональных проблем, постепенное устранение которых увеличит количество довольных и активных клиентов.
- Кажется целесообразным провести тесты наподобие NPS (Net Promoter Score), чтобы выяснить, насколько клиенты удовлетворены сервисом.
- Собрать предложения по улучшению функционала сервиса при помощи email-рассылок или форм с тем, чтобы постепенно улучшать сервис, избавляясь от багов и недоработок.
- Некоторые некритичные доработки позволят облегчить анализ и построение моделей в будущем. Например, можно добавить в личном кабинете выбор рода деятельности через список со множественным выбором, это позволит избежать ошибок и разных вариантов написания. Кроме того, имеет смысл оптимизировать саму базу данных. В текущем виде она лишь затрудняет анализ — разные периоды, разное количество уникальных компаний в каждой таблице, наличие дубликатов, непонятно наличие огромного количества компаний с выставленными, но неоплаченными платежками, непонятна причина пропусков в датах платежей, если флаг оплаты выставлен в True и прочее.
- Модели эластичности позволят найти клиентов, чей спрос эластичен цене. Таких клиентов можно будет удерживать в сервисе за счет более низких цен (остальным клиентам можно назначать цену выше).
- Для снижения расходов на коммуникации можно использовать аплифт-модели, они позволят выбрать только тех клиентов, напоминания которым принесут желаемый результат (например, продление подписки).

Обоснование выводов

Матрица корреляции позволили выявить признаки, влияющие на отток (по убыванию влияния)

- last_prolong - длительность последней по времени подписки,
- last_tariff_total - общая длительность последнего тарифа (с учетом предыдущих подписок),
- total_prolong - общая длительность всех подписок компании,
- avg_income_coef - средний коэффициент дохода,
- income_coef - коэффициент дохода,
- month_activities_cnt - количество месяцев, когда клиент был активен,
- last_cat_tariff - группа последнего тарифа (1-4),
- last_month_activities - количество событий за последний месяц,
- avg_cat_tariff - средняя группа тарифов для компании,
- total_activities - общая активность компании за все время.

Признаки, имеющие нулевое влияние на отток

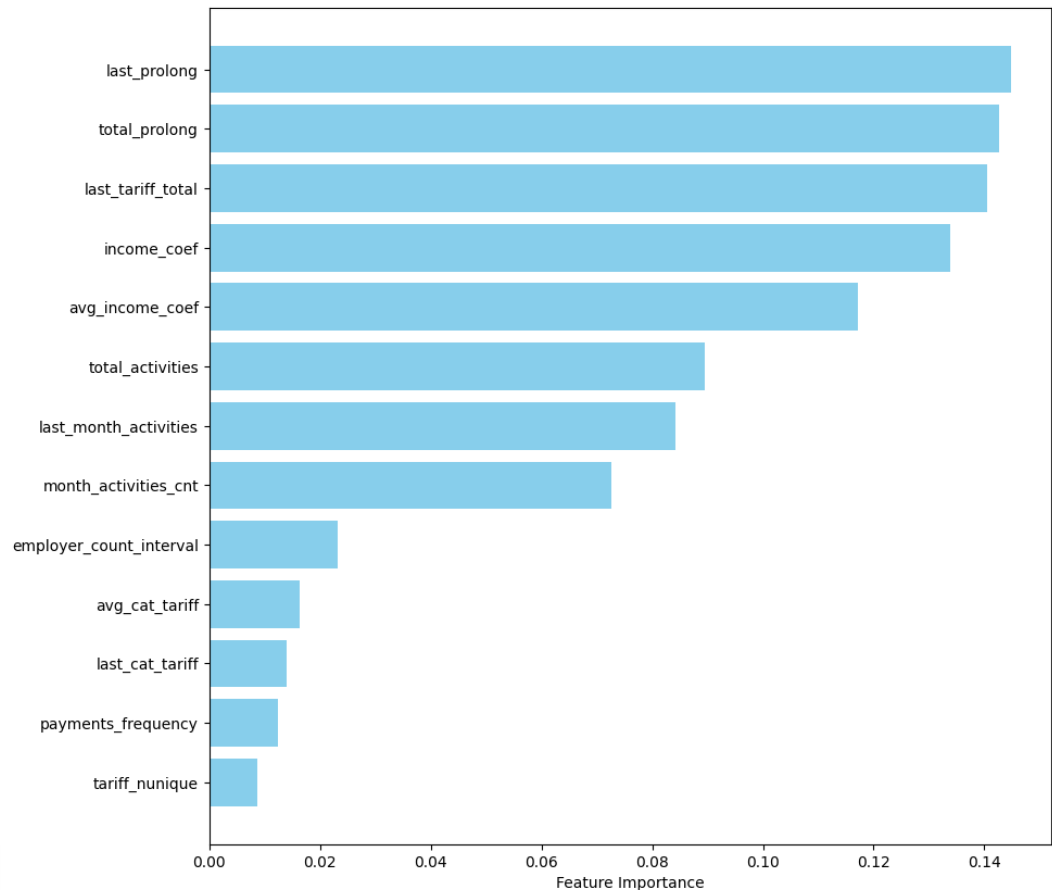
- payments_frequency - частота платежей,
- tariff_nunique - количество уникальных тарифов,
- employer_count_interval - интервал количества сотрудников.

last_prolong	0.692982
last_tariff_total	0.690185
total_prolong	0.682222
avg_income_coef	0.649822
income_coef	0.636879
month_activities_cnt	0.287445
last_cat_tariff	0.268954
last_month_activities	0.223178
avg_cat_tariff	0.171031
total_activities	0.158615
working_direction	0.050056
payments_frequency	0.000000
tariff_nunique	0.000000
employer_count_interval	0.000000

Обоснование выводов

Таблица важности признаков, полученная при помощи модели классификации:

	feature	importance
4	last_prolong	0.144781
6	total_prolong	0.142686
5	last_tariff_total	0.140454
7	income_coef	0.133824
8	avg_income_coef	0.117252
12	total_activities	0.089552
11	last_month_activities	0.084275
10	month_activities_cnt	0.072706
9	employer_count_interval	0.023118
3	avg_cat_tariff	0.016349
2	last_cat_tariff	0.013892
0	payments_frequency	0.012478
1	tariff_nunique	0.008633



Обоснование выводов

Сфера деятельности

	departed	loyal
tag_construction	23.86	28.19
tag_repair	6.22	5.29
tag_finishing	2.90	1.76
tag_engineering	2.90	2.64
tag_electrical	1.24	0.44
tag_installation	1.24	0.88
tag_project	1.24	1.32
tag_facade	0.83	0.44
tag_production	0.83	0.88
tag_design	0.41	0.44
tag_trade	0.21	0.88

Количество сотрудников компаний

	eci	departed	loyal
0	1	0.699170	0.656388
1	2	0.203320	0.211454
2	3	0.058091	0.074890
3	4	0.018672	0.022026
4	5	0.010373	0.017621
6	6	0.004149	NaN
7	7	0.002075	0.008811
5	8	0.004149	0.008811

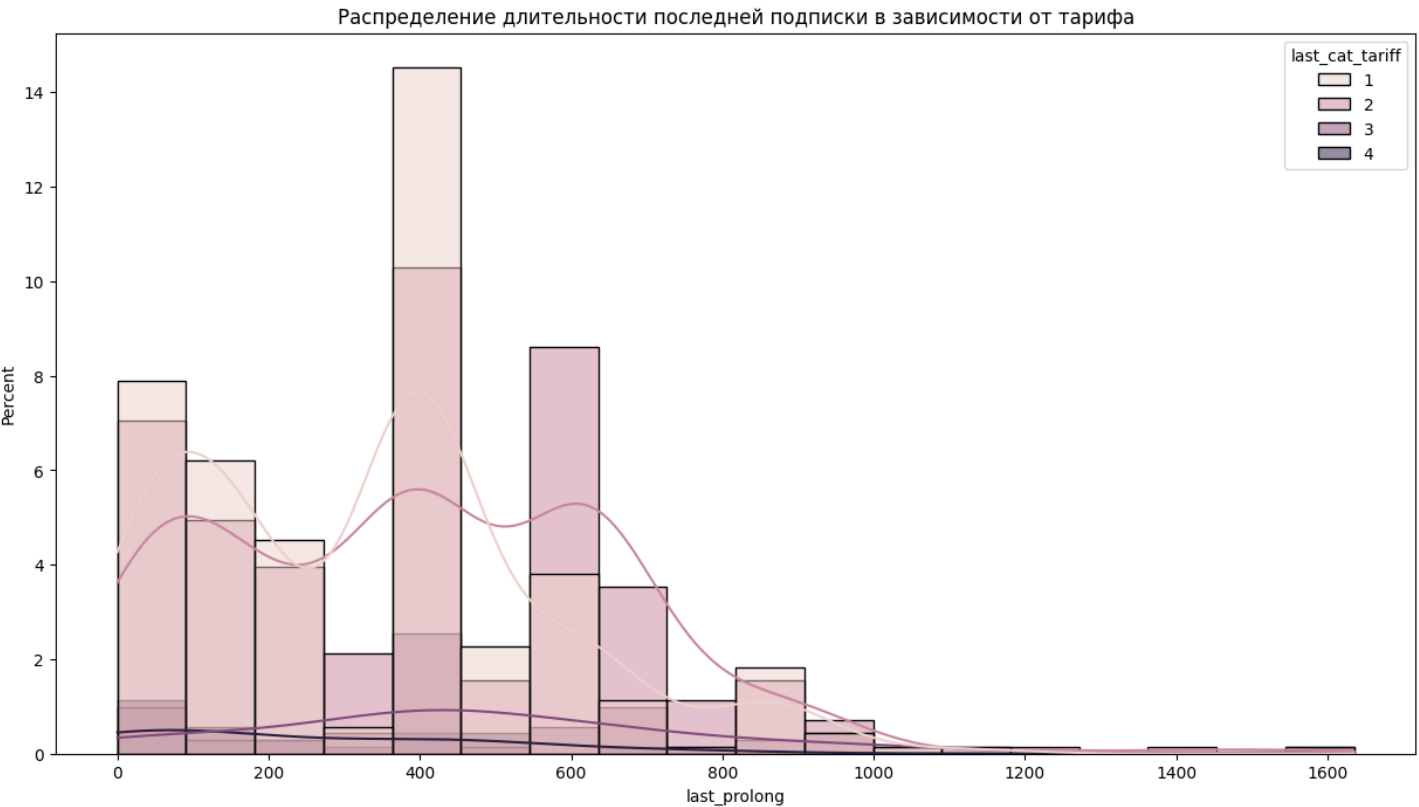
Обоснование выводов

Отток по тарифам

Старт	76%
Базовый	64%
Бизнес	43%
Профи	74%

Распределение по тарифам

Отток	Старт	Базовый	Профи	Бизнес
нет	10.58	16.50	4.09	0.85
да	33.00	29.48	3.10	2.40



Обоснование выводов

Портрет отточного клиента:

- около 80% отточных клиентов совершают всего один платеж,
- около 90% использовали только 1 тариф,
- около 50% выбрали тариф группы "Старт", еще приблизительно 45% - "Базовый",
- медианное последнее продление подписки 195 дней (среднее - 275 дней),
- медианная жизнь в сервисе 364 дня (средняя - 339 дней),
- медианный усредненный коэффициент доходности 363 (средний - 432),
- около 70% компаний имеют 1-4 сотрудников, 20% - 5-10,
- у 75% компаний количество активных месяцев меньше 11,
- почти у 30% активность в последний месяц равна 1 событию, у 15% - 2 событиям,
- у 29% активность за весь период не превышает 50 событий,
- 24% указали в роде деятельности строительство, 6% - ремонт, по 3% - отделку и инженерные системы.