

Прогнозирование вероятности оттока пользователей для сети фитнес-центров

Статус проекта: *завершен*

Использованные

библиотеки: pandas, matplotlib, seaborn, numpy, datetime, sys, sklearn, scipy, itertools

Описание проекта:

Сеть фитнес-центров «Культурист-датасаентист» разрабатывает стратегию взаимодействия с клиентами на основе аналитических данных.

Распространённая проблема фитнес-клубов и других сервисов — отток клиентов. Для фитнес-центра можно считать, что клиент попал в отток, если за последний месяц ни разу не посетил спортзал. Конечно, не исключены частные случаи, после устранения которых клиент обязательно продолжит ходить на фитнес. Однако чаще бывает наоборот. Если клиент немного походил в спортзал, а потом пропал — скорее всего, он не вернётся.

Чтобы бороться с оттоком, отдел по работе с клиентами «Культуриста-датасаентиста» перевёл в электронный вид множество клиентских анкет. «Культурист-датасаентист» предоставил сведения в csv-файле, который содержат данные на месяц до оттока и факт оттока на определённый месяц.

Наша задача — провести анализ и подготовить план действий по удержанию клиентов.

Ход исследования:

1. Обзор данных и предобработка данных.
2. Прогноз вероятности оттока (на уровне следующего месяца) для каждого клиента.
3. Формирование типичных портретов клиентов:
 - выделить несколько наиболее ярких групп;
 - охарактеризовать их основные свойства.
4. Анализ основных признаков, наиболее сильно влияющих на отток.
5. Формулировка выводов и разработка рекомендации по повышению качества работы с клиентами:
 - выделить целевые группы клиентов;
 - предложить меры по снижению оттока;
 - определить другие особенности взаимодействия с клиентами.

Выводы по проекту:

Группы клиентов по доле оттока каждой группы от общего числа оттекших клиентов отличаются довольно сильно:

- самый сильный отток происходит в третьей группе - больше 51%;
- чуть больше 44% у второй группы;
- группа №1 на третьем месте по оттоку - почти 27%;
- в четвертой и пятой отток не столь сильный - ~ 7% и ~3% соответственно (можно считать клиентов в этих группах более надежными).

В первой группе клиенты, для которых клуб расположен достаточно близко, они часто являются сотрудниками партнеров клуба или пришли по промо-акции, абонемент у них, в среднем, больше чем на пол года, они чаще остальных посещают групповые занятия и довольно много тратят на сопутствующие услуги клуба. С момента первого обращения прошло почти 5 месяцев, а до окончания абонементов еще далеко, в среднем около 9 месяцев. Посещают клуб они около двух раз в неделю.

Во второй группе клиенты, для которых клуб расположен также достаточно близко, у большинства из них абонемент сроком до 6 месяцев, по потреблению услуг они где-то посередине между всеми группами, до окончания абонемента, в среднем, почти 5 месяцев, а вот с момента обращения прошло почти 4 месяца, менее половины из них посещают групповые, являются сотрудниками партнеров клуба или пришли по промо-акции, посещают зал около 2х раз в неделю.

В третьей группе клиенты, которым далеко до клуба добираться, менее половины из них являются сотрудниками партнеров клуба, чаще встречаются месячные абонементы, чем на более длительный срок, они реже остальных ходят на групповые занятия, но сопутствующие услуги потребляют, ходят скорее один раз в неделю, чем больше.

В четвертой группе клиенты, наиболее близко работающие/проживающие от клуба, среди них не так много акционных или партнерских, абонементы предпочитают краткосрочные и услугами пользуются меньше всех, группы посещают, но не часто да и в клуб ходят преимущественно раз в неделю.

В пятой группе клиенты, также базирующиеся недалеко от клуба, не особо часто это партнерские или акционные клиенты, зато среди них чаще попадаются любители абонементов подлиннее или посещающие групповые, у них также выше показатель времени с момента первого обращения и показатель потребления сопутствующих услуг клуба, к тому же они самые частые клиенты клуба по посещениям - почти 3 раза в неделю.