Баданина Наталья ПМ3-1

Разработка и построение рекомендательной системы

План:

1. Постановка задачи
   1. Формулировка
   2. Обозначения
2. Признаки
   1. Статистические
   2. Коллаборативная фильтрация
      1. Сравнение пользователей и товаров
      2. Тривиальные рекомендации
      3. Сходства пользователей
      4. Сходства товаров
3. Контентные модели
4. Ошибки
   1. Качество предсказаний
   2. Покрытие
   3. Новизна

Источники:

Статья в двух частях от разработчика ГК Ланит

<https://habr.com/ru/company/lanit/blog/420499/>

<https://habr.com/ru/company/lanit/blog/421401/>

Датасет: <https://www.kaggle.com/c/santander-product-recommendation/data>

Или <https://www.kaggle.com/retailrocket/ecommerce-dataset>

Источники «не хабр»

<http://www.ijircce.com/upload/2017/may/334_34_Intelligent.pdf>

<http://files.grouplens.org/papers/FnT%20CF%20Recsys%20Survey.pdf>

Федеральное государственное образовательное бюджетное  
 учреждение высшего профессионального образования   
**«Финансовый университет при Правительстве   
Российской Федерации»**

**Департамент анализа данных, принятия решений и финансовых**

**технологий**

**Курсовая работа**

*По дисциплине* «**Машинное обучение**» *на тему*:

**«Разработка и построение рекомендательной системы»**

Вид исследуемых данных:

Данные клиентов банка Santander Bank

Выполнила:

Студентка группы ПМ3-1

Баданина Наталья Дмитриевна

Научный руководитель:

к.э.н., доцент

Гринева Наталья Владимировна

Москва 2020

Раньше маркетинг был традиционным. До появления современных технологий продавец в локальном магазине знал о своих покупателей гораздо больше, чем требовалось для продажи. Можно сказать, что предложения продавца покупателю были персонализированными или таргетинговыми.

С развитием современных технологий и экстенсивным ростом индустрии, большие бизнесы, получающие доход с продаж, уже не знаю своих клиентов лично. В лучшем случае о клиенте известные общая информация, например пол или регион проживания. На этом этапе появляется не персонализированная реклама: баннеры, вывески и т. д. Телевидение, несмотря на свой охват, тоже предлагает не таргетинговую рекламу. Допустим, некоторый канал в определенное время смотрят работающие мужчины, больше информации извлечь невозможно. На основе таких данных рекомендации будут неэффективные: слишком общие или неактуальные. Проблема сбора данных решается с появление сети Интернет. Сейчас о пользователе можно узнать геопозицию, каким телефоном и браузером он пользуется, личные данные и увлечения из профилей социальных сетей. Особенностью сбора данных через интернет является то, что пользователь сам указывает начальную информацию о себе: ФИО, город, номер телефона и т. д.. Остальные данные получают исходя из поведения клиента.

Однако и при этом подходе рекомендации могут быть слабо персонализированными в силу специфичности данных. Сложно сделать вывод что нужно некому Васе, который зашел в мобильное приложение (МП) в 13:00. На этом этапе появляется машинное обучение, которое обработав массив данных пользователей может сделать неочевидный вывод, что некому Васе можно продать обед. Алгоритмы машинного обучения, которые предсказывают какой фильм/статью/товар/услугу купит пользователей называют рекомендательными системами. В основном такие системы используются для рекомендаций фильмов (Netflix, YouTube), товаров (Яндекс.Маркет, Ozon) и статей (Яндекс.Дзен). Но у персональных рекомендаций большое будущее в банковской сфере. У современных банков из продуктов уже не только кредиты. Сейчас на рынке предлагают пенсионные проекты, депозиты, брокерские услуги, услуги для SME бизнесов: инкассация, бухгалтерия, интернет эквайринг, конструкторы сайтов. Перечень продуктов и услуг исчисляется десятками и не о всех клиент знает, но и не все он готов купить. Банк не может позволить себе часто предлагать открыть свой продукт, так как может столкнуться с негативом и уходом клиента. Более того, кроссейл стоит денег. Звонок с предложением оформить кредит обойдется в 1-10 рублей. При случайном обзвоне конверсия в покупку может составить <1%. Это значит, что банк не заработает, а потеряет деньги. Внедряя рекомендательные движки, персональные менеджеры (ПМ) получают список потенциально заинтересованных клиентов, обзвон которых увеличивает конверсию, принося компании платящих клиентов.