

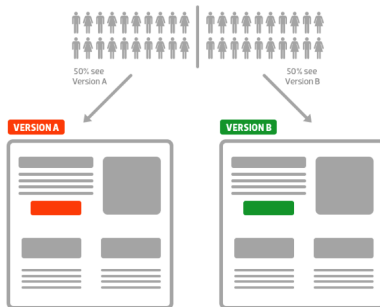
Моделирование поведения пользователей сервиса Яндекс.Толока

Высшая школа экономики

Студент: Гусев Андрей

Научный руководитель: Куралёнок Игорь Евгеньевич

2019



- А/Б тестирование используется при принятии решений о выпуске новой версии сервиса
- Существуют схемы тестирования с различными характеристиками
- Наилучшей схемы для всех сервисов не существует
- Запустить сразу несколько схем и сравнить результаты дорого

Сессия - период непрерывной активности пользователя

$$SpU = \frac{1}{N} \sum \frac{1}{\Delta t_i}$$

Взаимодействие пользователя с различными версиями системы моделируется взаимодействием со множеством объектов

- Яндекс.Толока - краудсорсинговая платформа
 - Однотипные задания сгруппированы в проекты
 - Сопоставляем проекты версиям системы
- Last.fm - сервис музыкальных рекомендаций
 - Сопоставляем версиям исполнителей

- Цель - разработать модель аудитории, подходящую для предсказания сессионных метрик А/Б тестирования на данных Яндекс.Толоки и last.fm
- Задачи:
 - Проанализировать распределение интервалов между сессиями в данных
 - Проанализировать существующие модели для предсказания времён действий пользователей
 - Оценить точность предсказания аудиторных метрик с использованием существующих моделей
 - При необходимости реализовать новую модель, приспособленную для данных и задачи
 - Оценить точность предсказания метрик с помощью новой модели

Существующие решения

- Для оценки используют метрику Mean absolute error

$$MAE = \frac{1}{N} \sum |t_{real} - t_{pred}|$$

- Самый простой ориентир - всегда предсказывать медиану интервалов обучающей выборки
- Модели обучения с учителем[1] - требуют факторы, которые не универсальны и ограничены
- Анализ выживаемости
 - Модель Кокса[2]
 - Точечные процессы[3]

¹Pudipeddi, J. et al. "User Churn in Focused Question Answering Sites: Characterizations and Prediction".

²Kapoor, K. et al. "A Hazard Based Approach to User ReturnTime Prediction"

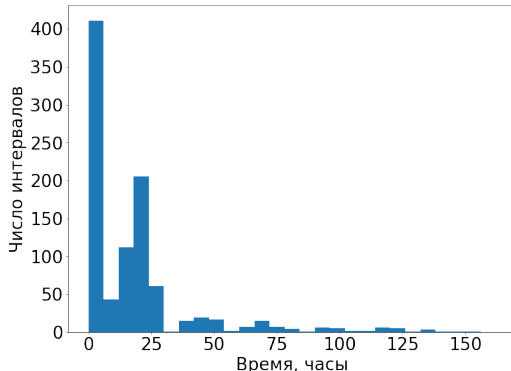
³Wang, Y. et al. "Coevolutionary Latent Feature Processes for Continuous-Time User-Item Interactions"

- Интенсивность процесса Хоукса: $\lambda(t) = \lambda_0 + \alpha \sum_{t_i \in \mathcal{T}} e^{(t-t_i)/\sigma}$.
- λ_0 и α - параметры, требующие оптимизации
- Эмбединги для переноса информации на новые пары из пользователя и объекта

⁴Du, N. et al. "Time-Sensitive Recommendation From Recurrent User Activities"

Несоответствие модели данным

- В LowrankHawkes и аналогах время до действия пользователя моделируется экспоненциальным распределением
- Интервалы между сессиями в рассматриваемых данных имеют другое распределение



- Перейдём к моделированию экспоненциальным распределением числа дней до следующей сессии
- Предсказываем возвращение пользователя на сервис, а не к конкретному объекту
- Обучаем максимизацией функции правдоподобия

Параметр распределения вычисляется по истории пользователя:

$$\lambda_{u,t} = \lambda_{u0} + \sum_{\substack{i \in H_u \\ t_i < t}} e^{-\beta \Delta_{i,t}} (v_u \cdot v_{p_i})$$

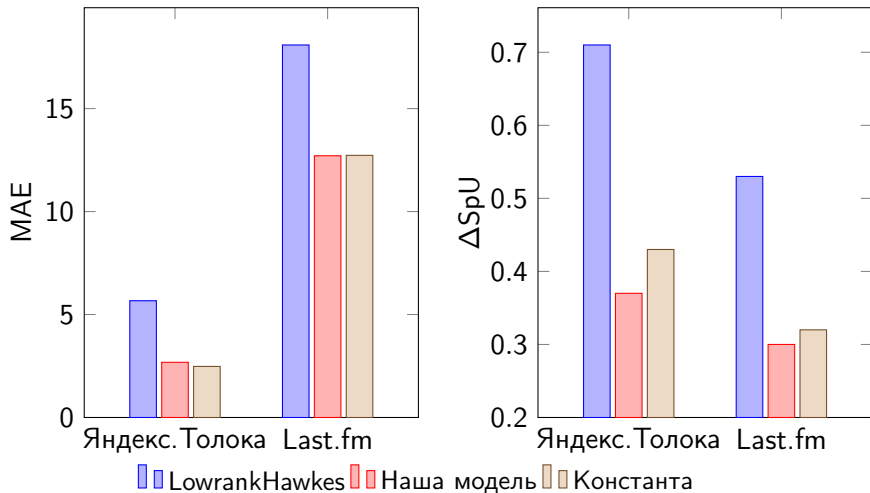
- H_u - множество взаимодействий пользователя
- $\Delta_{i,t}$ - число взаимодействий после i -го на момент запроса t
- v_u и v_{p_i} - эмбединги пользователя и объекта
- β - гиперпараметр, определяющий скорость убывания веса
- λ_{u0} - начальное значение, вычисляемое как оптимальный постоянный параметр на обучающей выборке

Для перехода к точному предсказанию в часах берётся средний интервал для данного числа пропущенных дней

Результаты

ΔSpU - разница SpU с тестовой выборкой.

Константа - медиана интервалов обучающей выборки



- Распределение времени между сессиями пользователя в рассматриваемых данных не является экспоненциальным
- Применение моделей, предполагающих экспоненциальное распределение, затруднено
- Количество дней без активности на сервисе поддаётся моделированию экспоненциальным распределением
- Реализована модель, позволяющая предсказывать SpU (ошибка 0.37 против 0.43 и 0.71)
- Метрики MAE и SpU слабо связаны