

Построение гибридной рекомендательной системы новостей с применением методов оптимизации

Смирнов Александр 17.Б07-мм

Научный руководитель: к.ф.-м.н., доц. Михайлова Елена Георгиевна

Рецензент: руководитель отдела инженерии ООО “АЙ ТИ Сервис”, Осипов Евгений Валерьевич

27 мая 2021 г.

Введение

- ▶ Приложение ЯRUS:
 - ▶ Агрегатор новостей;
 - ▶ Социальная сеть;
- ▶ Огромный объём информации:
 - ▶ Необходима персонализация.

Постановка задачи

- ▶ Цель:
 - ▶ Реализация рекомендательной системы новостей в приложении ЯRUS;
- ▶ Задачи:
 - ▶ Исследование предметной области;
 - ▶ Анализ проблем существующих подходов;
 - ▶ Реализация подходов;
 - ▶ Совмещение подходов в единую систему;
 - ▶ Анализ качества работы рекомендательной системы;
 - ▶ Оценка влияния решения на ключевые показатели эффективности.

Подходы

- ▶ Коллаборативная фильтрация;
- ▶ Фильтрация на основе содержимого;
- ▶ Фильтрация на основе популярности.

Коллаборативная фильтрация

- ▶ Рекомендации в зависимости от истории похожих пользователей;
- ▶ Проблемы:
 - ▶ Холодный старт;
 - ▶ Вычислительные сложности;
 - ▶ Отзывчивость системы на действия пользователей;
 - ▶ Разреженность данных.

Фильтрация на основе содержимого

- ▶ Рекомендации в зависимости от истории взаимодействия пользователя;
- ▶ Проблемы:
 - ▶ Холодный старт;
 - ▶ Векторизация рекомендуемого предмета;
 - ▶ Однообразность содержимого.

Фильтрация на основе популярности

- ▶ Рекомендуемые самые “трендовые” новости;
- ▶ Проблемы:
 - ▶ Отсутствие персонализации.

Описание подхода

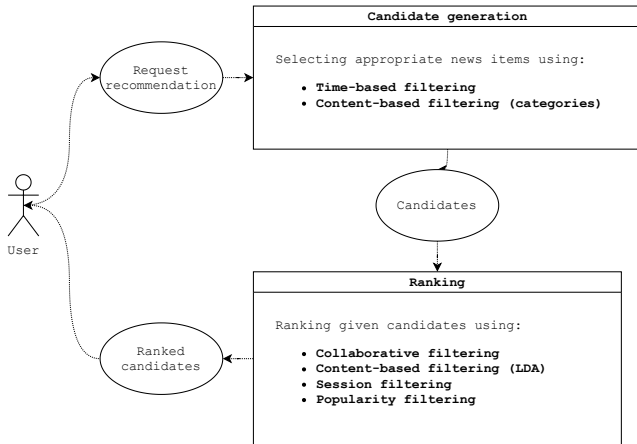


Рис. 1: Обзор подхода

Описание подхода (продолжение)

- ▶ Отбор кандидатов:
 - ▶ Штраф за устаревание новости;
 - ▶ Грубая фильтрация по категориям;
- ▶ Ранжирование:
 - ▶ Коллаборативная фильтрация;
 - ▶ Фильтрация на основе содержимого;
 - ▶ Фильтрация на основе текущей сессии;
 - ▶ Фильтрация на основе популярности.

Описание подхода (продолжение)

- ▶ Решённые проблемы:
 - ▶ Холодный старт;
 - ▶ Вычислительные сложности;
 - ▶ Отзывчивость системы на действия пользователей;
 - ▶ Векторизация рекомендуемого предмета;
 - ▶ Однообразность содержимого;
 - ▶ Отсутствие персонализации.

Фильтрация на основе содержимого



Коллаборативная фильтрация



Оценка качества (offline)

- ▶ MAP@20
- ▶ NDCG
- ▶

Оценка качества (online)

- ▶ A/B тестирование:
 - ▶ Время нахождения на вкладке “новости” за одну сессию;
 - ▶ Вовлечённость:
 - ▶ Количество эмоций;
 - ▶ Количество комментариев.



Внедрение

- ▶ Микросервисная архитектура:
 - ▶ Алгоритмы рекомендаций;
 - ▶ Алгоритмы подготовки и обработки данных;
- ▶ python, flask: разработка;
- ▶ docker: упаковка решений;
- ▶ k8s: оркестрация;
- ▶ gitlab: версионирование;
- ▶ gitlab CI: непрерывная интеграция.

Апробация

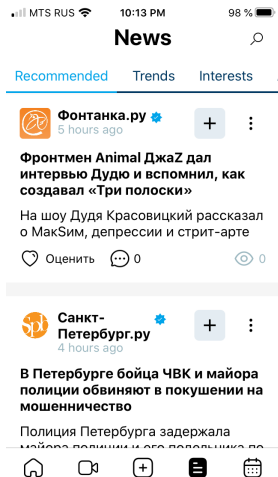


Рис. 2: Персонализированная рекомендательная лента

Результаты

- ▶ Проведён обзор существующих решений;
- ▶ Собрано уникальное решение;
- ▶ Оценено качество предложенного решения;
- ▶ Решение реализовано и внедрено в экосистему приложения YRUS;
- ▶ Увеличено время нахождения пользователей в приложении и повышена вовлеченность.

Акт о внедрении

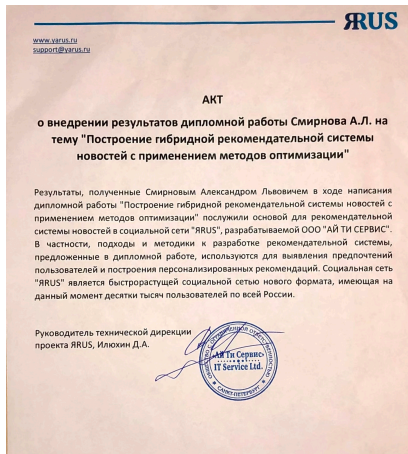


Рис. 3: Акт о внедрении

Заключение

- ▶ Спасибо за внимание!
- ▶ Задавайте вопросы.
- ▶ Ссылки:
 - ▶ yarus.ru
 - ▶ [@furioustebag](https://t.me/furioustebag)