

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

28 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.

Гибридные подходы

- ▶ Взвешенный;
- ▶ Комбинация признаков;
- ▶ Проблемы:
 - ▶ Вычислительные сложности;
 - ▶ Отзывчивость системы на действия пользователей;
 - ▶ Векторизация рекомендуемого предмета;
 - ▶ Разреженность данных.

Описание подхода

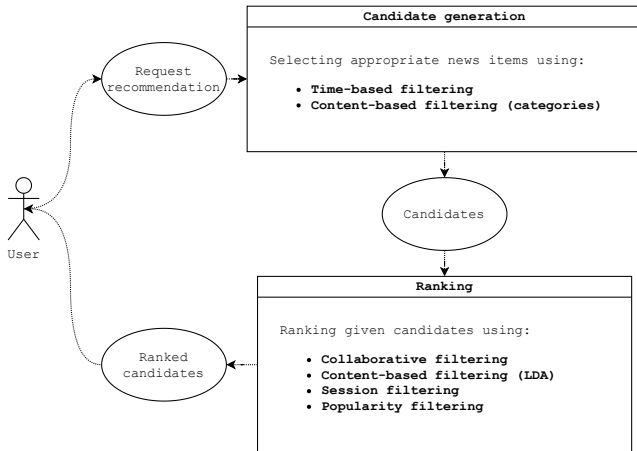


Рис. 1: Обзор подхода

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;
- ▶ Вклад каждого алгоритма находится оптимизацией $MAP@20$;
- ▶ Веса меняются с получением информации о действиях пользователя.

Описание подхода (продолжение)

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

Составление таблицы предпочтений

- ▶ По оси x : пользователи;
- ▶ По оси y : новости;
- ▶ Содержимое: как пользователь оценил новость:
 - ▶ Факт показа новости;
 - ▶ Факт просмотра новости;
 - ▶ Какой % новости прочитан;
 - ▶ Оставленная эмоция;
 - ▶ Оставленный комментарий.

Фильтрация на основе содержимого

- ▶ Векторизация текста:
 - ▶ LDA;
 - ▶ TF-IDF;
- ▶ Составления вектора пользовательских предпочтений:
 - ▶ Взвешенная сумма оценённых векторизованных новостей;
- ▶ Сравнение вектора пользователя и вектора новости.

Коллаборативная фильтрация

- ▶ Матричное разложение Funk MF;
- ▶ Пытаемся получить матрицы меньшей размерности, которые содержат скрытые признаки.

Оценка качества (offline)

method	recall@20
Multi-Gradient Descent	0.418
H+Vamp Gated	0.413
RaCT	0.403
Hybrid Recommender	0.398
Mult-VAE PR	0.395

Таблица 1: Benchmark on MovieLens 20M

- ▶ Метод не сильно хуже лучших моделей на benchmark датасете:
 - ▶ Подход обобщается на другие доменные области;
- ▶ Преимущество в решение продуктовых проблем.

Оценка качества (online)

feature	popularity	content	collaborative	hybrid
session duration	121	110	156	302
emoji number	1.1	0.8	1.4	2.5
comments number	0.2	0.1	0.4	0.8

Таблица 2: Online evaluation

- ▶ 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

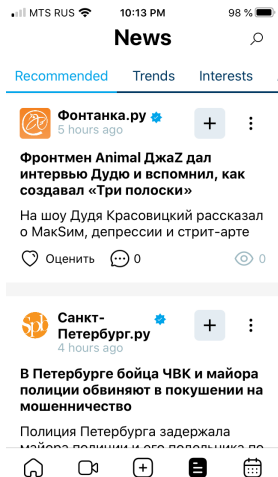


Рис. 2: Персонализированная рекомендательная лента

Результаты

- ▶ Проведён обзор существующих решений;
- ▶ Собрано уникальное решение;
- ▶ Оценено качество предложенного решения;
- ▶ Решение реализовано и внедрено в экосистему приложения ЯRUS;
- ▶ Разработана платформа для А/В тестирования;
- ▶ Увеличено время нахождения пользователей в приложении и повышена вовлечённость.

Акт о внедрении

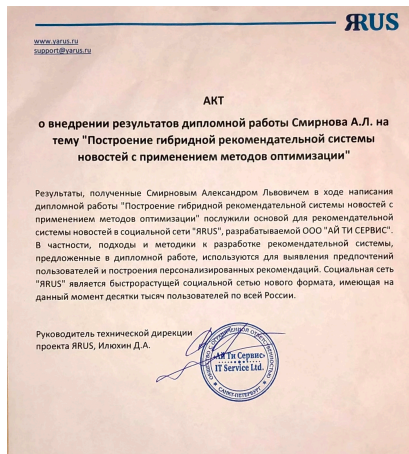


Рис. 3: Акт о внедрении

Заключение

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