

Валитова Наталья Львовна

кандидат технических наук, научный руководитель, Казанский национальный исследовательский технический университет имени

А. Н. Туполева – КАИ, Россия, г. Казань

Сонгатов Тагир Ришатович

студент, Казанский национальный исследовательский технический университет имени А. Н. Туполева – КАИ, Россия, г. Казань

songatov@gmail.com

АНАЛИЗ РЕКОМЕНДАТЕЛЬНЫХ СИСТЕМ ДЛЯ ОНЛАЙН- МАГАЗИНА КНИГ

Аннотация. В статье рассматриваются различные типы рекомендательных систем, такие, как коллаборативная фильтрация, контентная фильтрация. Автор рассуждает их преимущества и недостатки для определения наилучшего из них для внедрения рекомендательной системы в веб-приложение онлайн-магазина книг. Подчеркивается значимость использования коллаборативной фильтрации и алгоритма кластеризации k-means.

The article discusses various types of recommender systems, such as collaborative filtering, content filtering. The author discusses their advantages and disadvantages to determine the best one for implementing a recommender system in an online bookstore web application. The importance of using collaborative filtering and the k-means clustering algorithm is emphasized.

Ключевые слова: Рекомендательная система, рекомендации, k-средних, коллаборативная фильтрация.

Keywords: Recommendation system, recommendations, k-means, collaborative filter.

ВВЕДЕНИЕ

В современном мире цифровых технологий рекомендательные системы стали неотъемлемой частью практически всех онлайн-сервисов. Они

представляют собой программные алгоритмы, которые анализируют пользовательские предпочтения и поведение для предоставления персонализированных рекомендаций.

Особенно актуальны такие системы в сфере электронной коммерции, где они способствуют увеличению продаж и удержанию клиентов. В данной статье мы рассмотрим разработку рекомендательной системы на примере онлайн-магазина книг, что позволит глубже понять их значимость и сложности, связанные с их разработкой и внедрением.

Рекомендательные системы можно классифицировать на несколько основных типов: коллаборативная фильтрация, контентная фильтрация и гибридные модели. Коллаборативная фильтрация основывается на анализе предпочтений пользователей с похожими вкусами, тогда как контентная фильтрация фокусируется на характеристиках самих объектов, таких как жанр, автор и стиль. Гибридные модели объединяют оба подхода, что позволяет добиться более точных и персонализированных рекомендаций.

Современные онлайн-сервисы, такие как Amazon, Netflix и Spotify, активно используют рекомендательные системы, которые доказали свою эффективность в увеличении числа продаж и улучшении пользовательского опыта. Эти системы анализируют огромное количество данных, чтобы предлагать именно те товары или услуги, которые с наибольшей вероятностью заинтересуют конкретного пользователя. По данным различных исследований, внедрение рекомендательных систем может привести к увеличению продаж на 10–30% [1], что делает их важным инструментом в арсенале любого успешного предприятия.

Однако разработка и внедрение рекомендательных систем сопряжены с рядом сложностей. Одной из основных проблем является необходимость обработки больших объемов данных в реальном времени, что требует значительных вычислительных ресурсов. Кроме того, важно учитывать вопросы конфиденциальности и безопасности данных, поскольку пользователи должны быть уверены в том, что их личная информация не будет

использована в ущерб их интересам. Также стоит отметить проблему "холодного старта", когда система сталкивается с недостатком данных о новых пользователях или товарах, что затрудняет генерацию рекомендаций.

Учитывая данные трудности, разработка эффективной рекомендательной системы требует тщательного выбора подхода, который бы наилучшим образом решал указанные проблемы. В рамках данной темы мы рассмотрим различные методы, которые могут быть реализованы в системе рекомендаций в онлайн-магазине книг, что позволит не только повысить продажи, но и значительно улучшить пользовательский опыт.

КОНТЕНТНАЯ ФИЛЬТРАЦИЯ

В основе этого метода лежит анализ характеристик самих объектов, будь то книги, фильмы или другие товары, с целью предоставления рекомендаций, основанных на содержании [2]. В контексте онлайн-магазина книг контентная фильтрация фокусируется на таких атрибутах, как жанр, автор, стиль написания и другие метаданные, чтобы предлагать пользователю книги, которые схожи с теми, что он уже оценил.

Преимущества систем контентной фильтрации [3]:

- легко масштабируется;
- рекомендации основываются на повседневной деятельности пользователей, что повышает точность рекомендаций;
- все новые объекты с заданными характеристиками могут начинать рекомендоваться сразу после их появления.

Недостатки систем контентной фильтрации [3]:

- ограниченность рекомендаций характеристиками (отдельными категориями товаров);
- субъективность характеристик;
- проблемы с производительностью;
- холодный старт.

Контентная фильтрация, благодаря своей способности анализировать и учитывать индивидуальные характеристики объектов, является важным

инструментом в арсенале методов рекомендательных систем. Несмотря на некоторые ограничения, она предоставляет мощные возможности для персонализации рекомендаций и позволяет пользователям находить для себя схожие книги, которые соответствуют их интересам.

КОЛЛАБОРАТИВНАЯ ФИЛЬТРАЦИЯ

Коллаборативная фильтрация является одним из наиболее популярных подходов в создании рекомендательных систем. В отличие от контентной фильтрации, она фокусируется на анализе предпочтений и поведения пользователей, а не на характеристиках самих объектов. Основная идея заключается в том, что если два пользователя проявляют схожие интересы к определенной группе объектов, то, вероятно, они будут заинтересованы и в других схожих объектах [2].

Так же выделяют 3 типа коллаборативной фильтрации [5]:

- основанный на модели, анализирующий только меру сходства пользователей
- основанный на соседстве, используют самообучаемый алгоритм
- гибридный, он обходится дороже и сложнее, но не имеет недостатков двух других типов.

Преимущества коллаборативной фильтрации [3]:

- не требует знаний о предметной области и работы человека;
- возможны предложения новых объектов, ранее не знакомых конкретному пользователю;
- персонализированные рекомендации.

Недостатки коллаборативной фильтрации [3]:

- разреженность матрицы оценок;
- сложность расчета для большого количества объектов;
- холодный старт (отсутствие необходимого количества данных для корректной работы системы фильтрации);
- синонимия (сложность в умении различать схожие (синонимичные) взаимодействия).

Коллаборативная фильтрация, благодаря своей способности находить скрытые связи между пользователями через их взаимодействия с объектами, является мощным инструментом для рекомендательных систем.

СРАВНИТЕЛЬНЫЙ АНАЛИЗ РЕКОМЕНДАТЕЛЬНЫХ СИСТЕМ

Существует большое количество подходов при реализации рекомендательных систем, но каждая из типов имеет свои плюсы и минусы и используются для разных целей. Все из них имеют место для реализации под разные нужды. В данной статье мы будем опираться на рекомендательную систему, которая лучше всего подходит для онлайн-магазина книг.

В первую очередь нужно сказать, что для онлайн-магазина книг было бы идеальным создание гибридной рекомендательной системы, но такие системы очень сложны в реализации и очень дорогие, например компания Netflix потратила миллиарды долларов на разработку своей гибридной системы, поэтому мы не будем учитывать гибридную рекомендательную систему [2].

Контентная фильтрация имеет несколько минусов, она предлагает объекты, которые уже нравятся пользователю, не предлагая новые объекты, а также имеет проблемы с производительностью. Если ограниченность новизны можно устранить, введя дополнительно рекомендации новинок, то с производительностью сложнее.

В основном алгоритмы для контентной фильтрации используют сложные алгоритмы обработки текстов, в таком случае требуется качественные метрики, характеризующие объект, что для книг довольно затруднительно. Используя рекомендательную систему на основе коллаборативной фильтрации эта проблема уходит, поскольку используются метрики, характеризующие насколько пользователю понравилась книга. Однако у коллаборативной фильтрации имеется проблема “холодного старта”. Ее можно решить, дополнив рекомендации выводом дополнительно популярные книги для новых пользователей. Проблему с производительностью можно снизить, используя алгоритм кластеризации k-means [4], за счет уменьшения точности рекомендуемых объектов. Поэтому для онлайн-магазина книг

хорошим вариантом будет использование коллаборативной рекомендательной системы на основе кластеризации.

ЗАКЛЮЧЕНИЕ

В условиях современного рынка электронной коммерции персонализация предложений и улучшение пользовательского опыта являются ключевыми факторами успеха. Для онлайн-магазина книг внедрение рекомендательной системы способно значительно повысить лояльность клиентов и увеличить продажи. Коллаборативная фильтрация, благодаря своей способности анализировать предпочтения пользователей, представляет собой оптимальное решение для создания таких систем. Специфика использования алгоритма k-means в рамках коллаборативной фильтрации заключается в его способности эффективно сегментировать пользователей на основе их оценок книг. Применение k-means также способствует быстрому и точному анализу больших объемов данных, что особенно важно в условиях растущего числа пользователей и ассортимента книг. Таким образом, коллаборативная фильтрация в сочетании с алгоритмом k-means представляет собой оптимальный вариант для онлайн-магазина книг и является компромиссом между точностью рекомендаций и скоростью получения контента.

СПИСОК ЛИТЕРАТУРЫ

1. Золотарев В., Конюхов Д. Retail Rocket – платформа управления Retention [Сайт]: Рекомендательные системы: что это и как они работают // Блог. 2024. URL: <https://retailrocket.ru/blog/rekomendatelnye-sistemy/> (дата обращения: 05.06.2025).
2. Skupro.wiki [Электронный ресурс]: Рекомендательные системы в Data Science // Профессии. URL: <https://sky.pro/wiki/profession/rekomendatelnye-sistemy-v-data-science/> (дата обращения: 05.06.2025).
3. Ворошень В. А. Электронная библиотека БГТУ [Электронный ресурс]: Сравнительный анализ алгоритмов рекомендательных систем. – URL: <https://elib.belstu.by/handle/123456789/68522> (дата обращения: 05.06.2025).

4. Филиппов С. А., Захаров В. Н., Ступников С. А., Ковалев Д. Ю. Кластеризация профилей пользователей в рекомендательных системах поддержки жизнеобеспечения на основе реальных неявных данных // Труды XVIII Международной конференции DAMDID/RCDL'2016 "Аналитика и управление данными в областях с интенсивным использованием данных", Ершово, 11–14 октября 2016. URL: <https://ceur-ws.org/Vol-1752/paper16.pdf> (дата обращения: 05.06.2025).

5. Большаков Н. Что такое коллаборативная фильтрация // Платформа омниканального маркетинга: Блог. – 2023. URL: <https://www.calltouch.ru/blog/kak-rabotaet-kollaborativnaya-filtracziya/> (дата обращения: 05.06.2025).

REFERENCES

1. Zolonarev V., Konuhov D. Retail Rocket – Retention management platform [Website] : Recommender Systems: What They Are and How They Work // Blog. 2024. URL: <https://retailrocket.ru/blog/rekomendatelnye-sistemy/> (date of access: 05.06.2025).

2. Skypro.wiki [Website] : Recommender Systems in Data Science // Professions. URL: <https://sky.pro/wiki/profession/rekomendatelnye-sistemy-v-data-science/> (date of access: 05.06.2025).

3. Voroshen V. A. Electronic library of Belarusian State Technological University [Website] : Comparative analysis of recommender system algorithms. – URL: <https://elib.belstu.by/handle/123456789/68522> (date of access: 05.06.2025).

4. Filippov S. A., Zakharov V. N., Stupnikov S. A., Kovalev D. U. Clustering of User Profiles in Life Support Recommender Systems Based on Real-World Implicit Data // Proceedings of the XVIII International Conference DAMDID/RCDL'2016 "Analytics and Data Management in Data-Intensive Areas", Ershovo, 11–14 October 2016. URL: <https://ceur-ws.org/Vol-1752/paper16.pdf> (date of access: 05.06.2025).

5. Bolshakov N. What is collaborative filtering // Omnichannel Marketing Platform: Blog. – 2023. URL: <https://www.calltouch.ru/blog/kak-rabotaet-kollaborativnaya-filtracziya/> (date of access: 05.06.2025).