



КАФЕДРА ИУ7 «Программное обеспечение ЭВМ и информационные технологии»

2025 г.

РЕФЕРАТ

Расчетно-пояснительная записка 14 с., 0 рис., 1 таблица, 9 источников, 1 приложение.

РЕКОМЕНДАТЕЛЬНЫЕ СИСТЕМЫ, КОЛЛАБОРАТИВНАЯ ФИЛЬТРАЦИЯ, ФИЛЬТРАЦИЯ НА ОСНОВЕ СОДЕРЖАНИЯ, ФИЛЬТРАЦИЯ НА ОСНОВЕ ЗНАНИЯ, ГИБРИДНЫЕ МЕТОДЫ, МАРКЕТПЛЕЙСЫ

Цель работы — сравнение методов рекомендации товаров в маркетплейсах по сформулированным критериям.

В данной работе были выделены методы решения задачи рекомендательной системы и проведено сравнение методов по различным сформулированным критериям. По результатам сравнения был сделан вывод, что каждый метод решения задачи рекомендательной системы имеет свои особенности и предназначен для решения определённых задач при определённых условиях.

СОДЕРЖАНИЕ

РЕФЕРАТ	3
ВВЕДЕНИЕ	5
1 Анализ предметной области	6
1.1 Основные определения	6
1.2 Формализация задачи	6
2 Методы решения	8
2.1 Основные методы и их классификация	8
2.1.1 Методы коллаборативной фильтрации	8
2.1.2 Методы на основе содержания	8
2.1.3 Методы на основе знаний	9
2.1.4 Гибридные методы	10
3 Сравнение методов	11
ЗАКЛЮЧЕНИЕ	12
СПИСОК ИСПОЛЬЗОВАННЫХ ИСТОЧНИКОВ	13
ПРИЛОЖЕНИЕ А	14

ВВЕДЕНИЕ

В настоящее время маркетплейсы активно используют методы рекомендательных систем для улучшения качества обслуживания и количества продаж товаров. Такие системы на основе имеющейся информации рекомендуют пользователям товары, которые соответствуют их предпочтениям.

Целью данной научно-исследовательской работы является исследование методов рекомендации товаров в маркетплейсах.

В рамках работы были поставлены следующие задачи:

- 1) провести анализ предметной области;
- 2) формализовать задачу рекомендательной системы;
- 3) перечислить существующие методы решения задачи;
- 4) сформулировать критерии сравнения методов решения;
- 5) провести сравнительный анализ методов решения на основе сформулированных критериев.

1 Анализ предметной области

1.1 Основные определения

Рекомендательные системы — это программные инструменты и методы, предоставляющие пользователю предложения относительно объектов, которые могут быть ему полезны [2]. В случае маркетплейсов такими предлагаемыми объектами будут являться товары.

Обратная связь пользователя — реакция пользователя собираемая системой для использования в качестве обучения для рекомендательных систем [2].

Обратную связь пользователя разделяют на два вида:

- 1) Явная обратная связь — пользователь явно оценил объекты;
- 2) Неявная обратная связь — обратная связь извлекается на основе наблюдения и анализа действий пользователя.

Проблема нового пользователя или товара (проблема «холодного старта») — ситуация, при которой невозможно точно вычислить рекомендации для новых пользователей или товаров, так как недостаточно информации о пользователе или товаре [5].

Пользователи «серые овцы» (grey-sheep users) — пользователи с уникальными или нестандартными предпочтениями, которые ведут к низким корреляциям с большинством пользователей и, как следствие, приводят к нетипичным рекомендациям для них [3].

1.2 Формализация задачи

Пусть U — множество пользователей, I — множество объектов. Для каждого пользователя $u \in U$ существует множество объектов $I_u \subset I$, с которыми он взаимодействовал, и которым он присвоил рейтинги согласно выражению 1.1.

$$R = (r_{ui})_{i \in I_u}, \quad (1.1)$$

где R — матрица отношения размером $n_{users} \times n_{items}$, r_{ui} — характеристика обратной связи пользователя u с объектом i .

Задача рекомендательных систем формулируется следующим образом: для каждого пользователя $u \in U$ необходимо оценить значения r_{ui} для объектов

$i \in I \setminus I_u$, выбрать заданное количество k объектов с наибольшими предсказанными рейтингами \hat{r}_{ui} , что записывается в виде выражения 1.2.

$$Recommend_k(u, \dots) = (i_1, \dots, i_k) \quad (1.2)$$

где $Recommend_k$ — функция, принимающая пользователя u , для которого выполняется рекомендация, и необходимые для конкретного метода параметры, и возвращающая упорядоченный набор (i_1, \dots, i_k) из k объектов, с наибольшими предсказанными рейтингами \hat{r}_{ui} для пользователя u .

2 Методы решения

2.1 Основные методы и их классификация

Методы рекомендаций обычно классифицируют следующим образом [1, 2, 6, 8, 9]:

- 1) коллаборативная фильтрация;
- 2) на основе содержания;
- 3) на основе знания;
- 4) гибридные.

2.1.1 Методы коллаборативной фильтрации

Основная идея коллаборативных рекомендательных подходов заключается в использовании информации о прошлом поведении или мнениях существующего сообщества пользователей для прогнозирования того, какие объекты, вероятнее всего, будут наиболее релевантны текущему пользователю [1].

Чистые подходы коллаборативной фильтрации не используют и не требуют информации о самих товарах и пользователях. Преимуществом этого метода является то, что эти данные не нужно вводить в систему или поддерживать.

Но данный метод обладает и недостатками, выделяют следующий ряд проблем характерных для коллаборативной фильтрации:

- 1) проблема нового пользователя;
- 2) проблема нового товара;
- 3) проблема информационного пузыря;
- 4) проблема «серой овцы»;
- 5) невозможность явно задать предпочтения;

2.1.2 Методы на основе содержания

Рекомендации на основе содержимого строятся на наличии описаний объектов и профиля, который присваивает этим характеристикам определённое значение. Для маркетплейса возможными характеристиками товара могут быть категория, бренд, описание [1]. Подобно описаниям объектов, профили пользователей также могут автоматически формироваться и обучаться либо на основе анализа поведения и обратной связи пользователей, либо путём прямого опроса

об их предпочтениях, например, через заполнение анкеты при регистрации на маркетплейсе.

По сравнению с подходами, не использующими информацию о содержании, рекомендации на его основе имеют два преимущества. Во-первых, им не требуются большие группы пользователей для достижения достаточной точности рекомендаций. Во-вторых, новые объекты можно сразу рекомендовать, как только становятся доступны их характеристики [1]. Таким образом, у данных методов отсутствуют проблемы нового пользователя, нового товара, «серой овцы», при этом поддерживается высокий уровень персонализации. Основными недостатками этого метода являются низкая точность при малом количестве данных о характеристиках товаров или пользователях и проблема возникновения информационного пузыря, так как пользователь начнёт взаимодействовать с рекомендованными товарами, что повлечёт ещё большую рекомендацию подобных товаров [7].

2.1.3 Методы на основе знаний

В рекомендательных системах, основанных на знаниях, рекомендации формируются на основе существующих знаний или правил о потребностях пользователей и функциях объектов. В отличие от методов, основанных на содержании или коллаборативной фильтрации, системы, основанные на знаниях, хранят базу знаний, созданную на основе информации, извлечённой из предыдущих записей пользователя. Эта база знаний содержит сведения о прошлых задачах, ограничениях и соответствующих решениях. Знания из базы используются системой при возникновении новой задачи рекомендации. Метод рассуждений на основе прецедентов применяет предыдущие случаи для решения текущей задачи и является широко используемой техникой в системах на основе знаний. В отличие от систем рекомендаций, основанных на содержании, поиск сходства между продуктами требует более структурированного представления. При этом осуществляется сравнение предыдущего случая с текущим и адаптация решения [4].

Применение методов рекомендаций на основе знаний особенно ценно в областях, которые характеризуются высокой специфичностью знаний предметной области, и каждый случай представляет собой уникальную ситуацию. Одним из

преимуществ этого подхода является отсутствие проблемы нового товара и проблемы нового пользователя, так как предварительные знания уже приобретены и сохранены в базе знаний. Ещё одним преимуществом является возможность пользователей накладывать ограничения на результаты рекомендаций. Однако у данного метода есть и недостатки — дополнительные затраты на настройку и управление системой при создании и поддержке базы знаний [4].

2.1.4 Гибридные методы

Гибридные рекомендательные системы представляют собой технические подходы, которые комбинируют несколько реализаций алгоритмов или компонентов рекомендаций [1].

На результат гибридной рекомендательной системы влияет архитектура системы и используемые методы рекомендательных систем. Таким образом, они приобретают необходимые свойства описанных ранее методов, что позволяет получить рекомендательные системы, которые минимально подвержены недостаткам, свойственным другим методам, при этом обладая их преимуществами.

3 Сравнение методов

Для сравнения методов были сформулированы следующие критерии [1,4]:

- 1) отсутствие проблемы нового пользователя;
- 2) отсутствие проблемы нового товара;
- 3) учёт параметров товара;
- 4) учёт параметров пользователя;
- 5) не требует параметров товара;
- 6) не требует параметров пользователя;
- 7) отсутствие проблемы информационного пузыря;
- 8) высокая персонализация;
- 9) отсутствие проблемы «серой овцы»;
- 10) возможность явно задать предпочтения.

В таблице 3.1 представлено сравнение перечисленных методов по сформулированным критериям.

Таблица 3.1 — Сравнение классов методов

Метод Номер критерия	Коллаборативная фильтрация	На основе содержания	На основе знаний	Гибридный
1	-	+	+	+
2	-	+	+	+
3	-	+	+	±
4	-	±	-	±
5	+	-	-	±
6	+	+	+	±
7	-	-	+	±
8	+	±	-	+
9	-	+	-	±
10	-	±	+	±

Гибридный метод удовлетворяет наибольшему количеству сформулированных для сравнения критериев, так как объединяет в себе другие методы для достижения необходимых свойств. Так, например, недостатки метода коллаборативной фильтрации компенсируются методом на основе содержания.

ЗАКЛЮЧЕНИЕ

Цель работы достигнута: были исследованы методы рекомендательных систем. При выборе методов решения задачи рекомендательных систем необходимо определить наличие и возможность использования данных о товарах, пользователях, предпочтении пользователей, а также уровень персонализации и необходимость решения основных проблем рекомендательных систем.

В ходе выполнения работы все поставленные задачи были выполнены:

- 1) проведён анализ предметной области;
- 2) формализована задача рекомендательной системы;
- 3) перечислены существующие методы решения задачи;
- 4) сформулированы критерии сравнения методов решения;
- 5) проведён сравнительный анализ методов решения на основе сформулированных критериев.

СПИСОК ИСПОЛЬЗОВАННЫХ ИСТОЧНИКОВ

1. Jannach D., Zanker M., Felfernig A., Friedrich G. Recommender Systems: An Introduction // Cambridge University Press, 2010. — 352 с.
2. Ricci F., Rokach L., Shapira B. Recommender Systems Handbook. 3-е изд. // New York: Springer, 2022. — 1060 с.
3. Zheng Y. Using Outlier Detection to Identify Grey-Sheep Users in Recommender Systems: A Comparative Study // Computers, Materials & Continua. — 2025. — Vol. 83, № 3. — P. 4315-4328.
4. Qian Zhang, Jie Lu, Yaochu Jin. Artificial intelligence in recommender systems // Complex and Intelligent Systems. — 2021. — Vol. 7. — P. 439-457.
5. Zangerle E., Bauer C. Evaluating Recommender Systems: Survey and Framework // ACM Computing Surveys. — 2022. — Vol. 55, № 8. — P. 1–38.
6. Черняков А. Н., Дибиров М. Ш. О некоторых способах построения рекомендательных систем онлайн-маркетинга на основе алгоритмов машинного обучения // Инновации и инвестиции — 2023 — № 6.
7. Волин М. В., Вовк Л. П. Рекомендательная система с контентной фильтрацией в электронной коммерции // Вестник науки и образования. — 2025. №6 (161)-1.
8. Михайлов А. Н. Разработка рекомендательных систем: методы и алгоритмы для электронной коммерции // Вестник науки. — 2024. — № 12 (81).
9. Штовбонько А. Искусственный интеллект в электронной коммерции: алгоритмы рекомендаций и их влияние на поведение пользователей // Вестник науки. — 2025. — № 4 (85).

ПРИЛОЖЕНИЕ А

Презентация к научно-исследовательской работе состоит из 3-х слайдов.