МИНИСТЕРСТВО НАУКИ И ВЫСШЕГО ОБРАЗОВАНИЯ

РОССИЙСКОЙ ФЕДЕРАЦИИ

ФЕДЕРАЛЬНОЕ ГОСУДАРСТВЕННОЕ БЮДЖЕТНОЕ

ОБРАЗОВАТЕЛЬНОЕ УЧРЕЖДЕНИЕ ВЫСШЕГО ОБРАЗОВАНИЯ

«ОРЛОВСКИЙ ГОСУДАРСТВЕННЫЙ УНИВЕРСИТЕТ

имени И.С. ТУРГЕНЕВА»

ВЫПУСКНАЯ КВАЛИФИКАЦИОННАЯ РАБОТА

по направлению подготовки 02.03.03 Математическое обеспечение

и администрирование информационных систем

направленность (профиль): Математическое обеспечение

интеллектуальных систем

Студента Багрова Владимира Андреевичашифр 210845

Факультет (институт): физико-математический

Кафедра алгебры и математических методов в экономике

Тема выпускной квалификационной работы

**Разработка рекомендательного сервиса для интернет-магазина**

|  |  |  |
| --- | --- | --- |
| Студент |  | В.А. Багров |
| Научный  Руководитель \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_ |  | Т.Н. Русских |
| Руководитель  образовательной  программы \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_ |  | С.П. Строев |

Орёл 2025

МИНИСТЕРСТВО НАУКИ И ВЫСШЕГО ОБРАЗОВАНИЯ

РОССИЙСКОЙ ФЕДЕРАЦИИ

ФЕДЕРАЛЬНОЕ ГОСУДАРСТВЕННОЕ БЮДЖЕТНОЕ

ОБРАЗОВАТЕЛЬНОЕ УЧРЕЖДЕНИЕ ВЫСШЕГО ОБРАЗОВАНИЯ

«ОРЛОВСКИЙ ГОСУДАРСТВЕННЫЙ УНИВЕРСИТЕТ

имени И.С. ТУРГЕНЕВА»

Физико-математический факультет

Кафедра алгебры и математических методов в экономике

Направление подготовки (специальность) 02.03.03 Математическое обеспечение

и администрирование информационных систем

Направленность (профиль) Математическое обеспечение

интеллектуальных систем

|  |  |
| --- | --- |
| УТВЕРЖДАЮ:  Руководитель  образовательной программы | |
|  | \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_С.П. Строев |
| «28» ноября 2024 г. | |

**ЗАДАНИЕ**

на выполнение выпускной квалификационной работы

студента Багрова Владимира Андреевича шифр 210845

1. Тема ВКР: Разработка рекомендательного сервиса для интернет-магазина.

Утверждена приказом по университету «27» ноября 2024 года №2-3655.

2. Срок сдачи студентом работы: «3» июня 2025 года.

3. Исходные данные к работе: материалы учебных пособий, научных статей, статистические данные, справочные данные сети Internet.

4. Содержание ВКР (перечень подлежащих разработке вопросов)

1). Изучение предметной области.

2). Обзор и сравнительный анализ программных решений в предметной области.

3). Обзор методов и авторских подходов к решению к решению задачи разработки рекомендательного сервиса.

4). Разработка рекомендательного сервиса с использованием библиотек языка программирования Python.

5). Разработка интерфейса рекомендательного сервиса для интернет-магазина.

6). Интеграция ChatGPT.

5. Перечень графического материала не предусмотрен.

6. Консультанты по ВКР не предусмотрены.

Дата выдачи задания «28» ноября 2024 года.

|  |  |  |
| --- | --- | --- |
| Научный руководитель ВКР |  | Т.Н. Русских |
| Задание принял к исполнению |  | В.А. Багров |

КАЛЕНДАРНЫЙ ПЛАН

|  |  |  |
| --- | --- | --- |
| Наименование этапов ВКР | Срок выполнения этапов работы | Примечание |
| Анализ литературы по теме исследования, обоснование актуальности темы  исследования. | 29.11.2024-6.12.2024 |  |
| Описание предметной области. | 7.12.2024-18.12.2024 |  |
| Обзор и сравнительный анализ программных решений в предметной области. | 19.12.2024-30.12.2024 |  |
| Обзор методов и авторских подходов к решению задачи разработки рекомендательного сервиса. | 9.01.2025 -31.01.2025 |  |
| Разработка рекомендательного сервиса с использованием библиотек языка программирования Python. | 1.02.2025-20.02.2025 |  |
| Разработка интерфейса рекомендательного сервиса для интернет-магазина. | 21.02.2025-  14.03.2025 |  |
| Интеграция ChatGPT. | 15.03.2025-  10.04.2025 |  |
| Структурирование материала ВКР,  оформление текста работы. | 16.05.2025-31.05.2025 |  |

|  |  |  |
| --- | --- | --- |
| Студент |  | В.А. Багров |
| Научный руководитель ВКР |  | Т.Н. Русских |

**АННОТАЦИЯ**

Выпускная квалификационная работа на тему «Разработка рекомендательного сервиса интернет-магазина» содержит 85 страниц, 23 рисунка, 5 таблиц. Список литературы включает **40** источников.

**Ключевые слова:** рекомендательная система, интернет-магазин, машинное обучение, коллаборативная фильтрация, контентная фильтрация, глубокое обучение, нейронные сети, гибридная модель, персонализация, UML-проектирование, разработка, веб-разработка.

Данная выпускная квалификационная работа посвящена проектированию и разработке рекомендательного сервиса для интернет-магазина, обеспечивающего персонализированный подбор товаров для пользователей. Актуальность темы определяется необходимостью повышения эффективности взаимодействия с клиентами в условиях высокой конкуренции в электронной коммерции, где успешная персонализация напрямую влияет на рост конверсии и удержание аудитории.

Целью исследования является разработка программного решения, способного автоматически анализировать поведение пользователей, учитывать их предпочтения и формировать индивидуальные товарные рекомендации на основе алгоритмов машинного обучения.

При решении поставленных задач использовались методы анализа и обобщения, методы предобработки данных, методы машинного обучения, методология проектирования и разработки информационной системы.

В работе представлен обзор существующих решений в области рекомендательных систем и электронной коммерции, определены требования к функциональности и архитектуре сервиса, выполнено его проектирование с использованием методологии UML. Реализация прототипа выполнена на языке программирования Python, а также с применением HTML, CSS, PHP, системы управления базами данных MySQL и специализированных библиотек Python для обработки данных и построения моделей рекомендаций. При разработке сервиса использованы методы коллаборативной и контентной фильтрации, а также технологии нейросетевого моделирования.

Разработанный сервис предусматривает интеграцию с данными интернет-магазина, генерацию персонализированных рекомендаций, использование ChatGPT, визуализацию метрик и пользовательский интерфейс, удобный для встраивания в существующую веб-платформу.

Практическая значимость работы заключается в возможности использования предложенного решения для повышения уровня персонализации, увеличения среднего чека, вовлечённости клиентов и автоматизации части маркетинговых функций интернет-магазинов малого и среднего бизнеса.

**СОДЕРЖАНИЕ**

**ВВЕДЕНИЕ**

**Актуальность**. В условиях стремительного роста электронной коммерции интернет-магазины сталкиваются с необходимостью повышения эффективности взаимодействия с пользователями. Одним из ключевых инструментов для решения этой задачи являются рекомендательные системы, которые позволяют персонализировать контент и предлагать товары, соответствующие предпочтениям покупателей. Это не только улучшает пользовательский опыт, но и способствует увеличению таких ключевых бизнес-показателей, как конверсия, средний чек и лояльность клиентов.

Современные рекомендательные системы используют передовые технологии, включая машинное обучение, нейронные сети и гибридные методы, что позволяет анализировать большие объемы данных и выявлять сложные зависимости. Однако, несмотря на широкое распространение данных систем, многие интернет-магазины сталкиваются с проблемами, такими как низкая точность рекомендаций, сложность интеграции и недостаточная адаптивность к изменяющимся предпочтениям пользователей. Это обуславливает актуальность разработки эффективного рекомендательного сервиса, способного решить указанные проблемы.

**Степень разработанности проблемы**. Проблема разработки рекомендательных систем затрагивается как в отечественной, так и в зарубежной литературе. Среди основных подходов выделяют контентную и коллаборативную фильтрацию, методы на основе знаний, а также современные технологии, такие как нейронные сети и гибридные модели. В работах Исламовой С.А., Ковалева И.Ю., Беспалова М.О., Куликова К.Г. [ссылка], рассматриваются различные аспекты построения рекомендательных систем, их преимущества и ограничения.

На рынке программных решений представлены такие платформы, как Ozon, Wildberries, Яндекс.Маркет и другие, которые успешно используют рекомендательные технологии для повышения эффективности продаж. Однако анализ существующих решений показывает, что многие из них требуют доработки или адаптации под специфику конкретного интернет-магазина.

**Объектом** исследования является процесс разработки рекомендательного сервиса для интернет-магазина.

**Предметом** исследования выступают методы и алгоритмы машинного обучения, применяемые для построения персонализированных рекомендаций, методы, технологии проектирования и разработки сервиса.

**Целью** выпускной квалификационной работы является проектирование и разработка рекомендательного сервиса для интернет-магазина на основе современных методов машинного обучения.

Для достижения поставленной цели были решены следующие **задачи**:

1. Рассмотреть понятие и виды рекомендательных систем.
2. Провести анализ существующих методов и алгоритмов, применяемых в рекомендательных системах.
3. Выполнить анализ данных для обучения модели.
4. Разработать алгоритм рекомендаций на основе выбранных методов.
5. Реализовать программный прототип рекомендательного сервиса.

*Задачи должны быть максимально близко по содержанию к задачам календарного плана, в итоговом варианте нужно приблизить.*

**Информационной базой** исследования послужили научные статьи, учебные пособия, материалы конференций, а также открытые датасеты и данные реальных интернет-магазинов.

**Практическая значимость** работы заключается в том, что разработанный рекомендательный сервис позволит интернет-магазину повысить удовлетворенность клиентов, увеличить продажи и оптимизировать бизнес-процессы за счет автоматизации персонализации контента.

**Структура и объем работы.** Общий объем работы составляет 85 страниц, 23 рисунка, 5 таблиц. Выпускная квалификационная работа включает введение, две главы, заключение, список литературы, приложение.

Во введении обоснована актуальность темы исследования, описана степень разработанности проблемы, определены объект, предмет исследования, сформулирована цель и задачи исследования, определены методы и практическая значимость исследования, описана структура работы.

Первая глава включает три пункта. В первом пункте раскрывается понятие рекомендательной системы, рассматриваются ее основные виды. Во втором пункте описаны математические методы и методики разработки рекомендательного сервиса. В третьем пункте проведен подробный обзор существующих программных решений, в частности, Ozon, Wildberries.

Вторая глава включает в себя … пунктов.

В заключении приведены основные результаты работы.

Список литературы содержит перечень источников, используемых при написании работы.

В приложениях приведен код разработанных программных решений.

1. **ПРЕДПРОЕКТНЫЙ АНАЛИЗ ОБЪЕКТА АВТОМАТИЗАЦИИ**

**1.1. Понятие и классификация рекомендательных систем**

В современных условиях развития сферы торговли большую популярность получили интернет-магазины. Их преимущество состоит в том, что подобный формат организации торговли позволяет уменьшить численность персонала, исключить аренду и кассовое обслуживание.

С возрастанием ассортимента предлагаемых производителями товаров и количества онлайн-магазинов покупателю становится всё сложнее находить необходимые товары среди большого их множества и разнообразия. В данном контексте на первый план выходят рекомендательные системы, которые приобретают особую актуальность, поскольку они позволяют, исходя из предпочтений потребителя, подобрать нужный для него товар.

Рекомендательные системы являются эффективным инструментом для кинотеатров, электронных библиотек и конечно же интернет-магазинов, позволяющим решать задачи увеличения доходности бизнеса. Благодаря использованию интеллектуальных систем и алгоритмов, анализа поведения пользователя, такие системы могут предлагать товары, которые наиболее интересны для конкретного пользователя и удовлетворяют его потребности.

В отечественной литературе авторы дают различные определения рекомендательных систем.

Рекомендательная система – это сервис или программа, цель которой – определение объектов, интересующих пользователей, и предоставление этих объектов. Благодаря анализу предпочтений, основываясь на предположениях, система выдает результаты, которые могут быть полезны пользователю [Исламова, С. А. Обзор рекомендательных систем / С. А. Исламова, Н. В. Липатникова // Modern Science. – 2019. – № 4-3. – С. 174-176.].

Система рекомендаций – это метод, используемый для предоставления пользователю оптимальных предложений по выбору конкретного элемента или товара [Овечкин А. В. ПРОЕКТИРОВАНИЕ РЕКОМЕНДАТЕЛЬНОЙ СИСТЕМЫ НА ОСНОВЕ НЕЙРОННОЙ СЕТИ // StudNet. 2022. №4.].

Рекомендательные системы можно разделить на несколько видов в зависимости от принципов их работы:

1. Контентные рекомендательные системы.
2. Коллаборативные рекомендательные системы.
3. Рекомендательные системы на основе знаний.
4. Гибридные рекомендательные системы.

Рассмотрим подробнее основные классы систем.

1. Контентные рекомендательные системы. Система формирует рекомендацию на основе поведения пользователя, то есть предлагаются объекты, похожие на те, что он уже искал. Степень похожести оценивается по признакам содержимого объектов (подбираются объекты со схожими характеристиками) [Исламова, С. А. Обзор рекомендательных систем / С. А. Исламова, Н. В. Липатникова // Modern Science. – 2019. – № 4-3. – С. 174-176.].

Принцип работы данного класса систем в том, что система анализирует характеристики товаров и находит схожие с теми товарами, которые интересуют пользователя (рисунок 1). Его предпочтения система фиксирует через взаимодействия, такие как просмотры, покупки, оценки и отзывы.

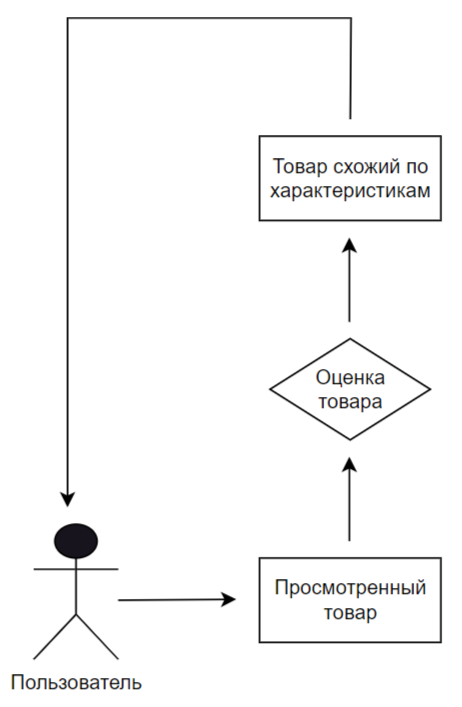


Рисунок 1 – Схема работы контентной рекомендательной системы

1. Коллаборативные рекомендательные системы.

Основная идея коллаборативных систем, использующих данный тип фильтрации, заключается в использовании известных предпочтений (оценок) групп пользователей для прогнозирования неизвестных предпочтений другого пользователя (рисунок 2). Другими словами, главной идеей является заинтересованность похожих людей в одних и тех же объектах. Рекомендации базируются на автоматическом сотрудничестве группы пользователей и на выделении тех пользователей, которые демонстрируют схожие предпочтения или шаблоны поведения [Исламова, С. А. Обзор рекомендательных систем / С. А. Исламова, Н. В. Липатникова // Modern Science. – 2019. – № 4-3. – С. 174-176.].

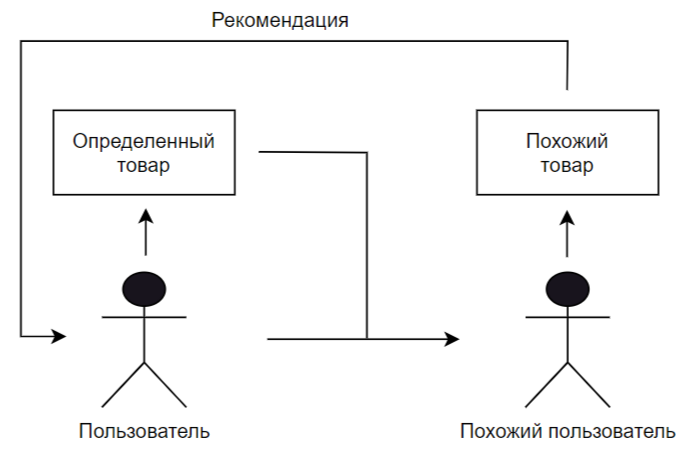


Рисунок 2 – Схема работы коллаборативной рекомендательной системы

В коллаборативной рекомендательной системе выделяют два метода фильтрации [Сидорова, М. Н. Рекомендательные системы как инструмент маркетинга / М. Н. Сидорова, А. А. Сидоров // Вестник Академии управления и производства. – 2021. – № 2-2. – С. 96-105.]:

– user-based, основанный на данных пользователей с похожими интересами;

– item-based, основанный на выборе товаров и услуг, ориентируясь на то, чем покупатель интересовался ранее.

Подход item-based по смыслу похож на user-based, но основан на сходстве товаров, а не предпочтений пользователей.

1. Рекомендательные системы на основе знаний. Этот тип систем основывается на знаниях о предметной области и схож с методом контентной фильтрации, однако используется более глубокий анализ объектов. Связь между объектами выстраивается не по простым критериям похожести, а по взаимосвязанности различных групп товаров. Схема работы данной системы представлена на рисунке 3 [Исламова, С. А. Обзор рекомендательных систем / С. А. Исламова, Н. В. Липатникова // Modern Science. – 2019. – № 4-3. – С. 174-176.].

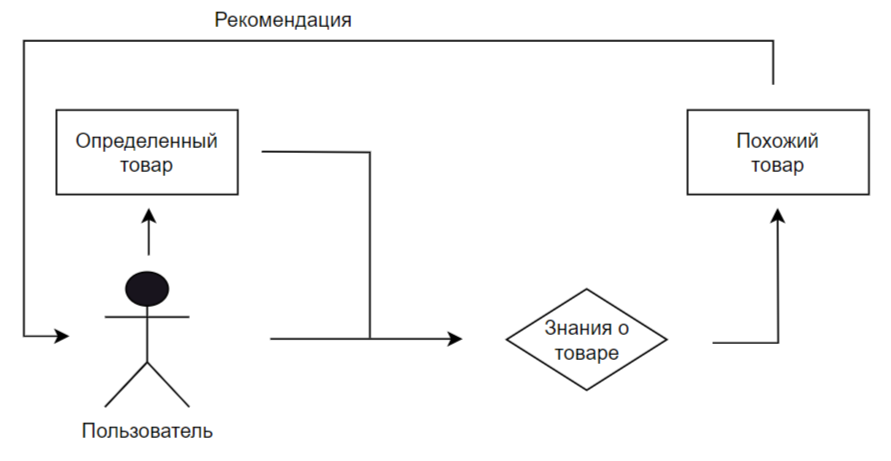


Рисунок 3 – Схема работы рекомендательной системы на основе знаний

1. Гибридные рекомендательные системы. Гибридные рекомендательные системы строятся с использованием комбинации подходов контентного и коллаборативного. Такие системы направлены на то, чтобы компенсировать недостатки одного подхода внедрением другого [Сорокин, А. С. Исследование и разработка алгоритмов рекомендации и ранжирования обучающих мультимедийных материалов / А. С. Сорокин, И. Б. Зарубин].

Рассмотрим основную идею, лежащую в основе разработки гибридных систем. Допустим, похожий пользователь купил интересующую данного пользователя модель ноутбука в синем корпусе вместе с чехлом и сумкой такого же цвета. Тогда в рекомендуемых товарах будет указана модель ноутбука в том же цвете, несколько таких же чехлов и сумок (рисунок 4).

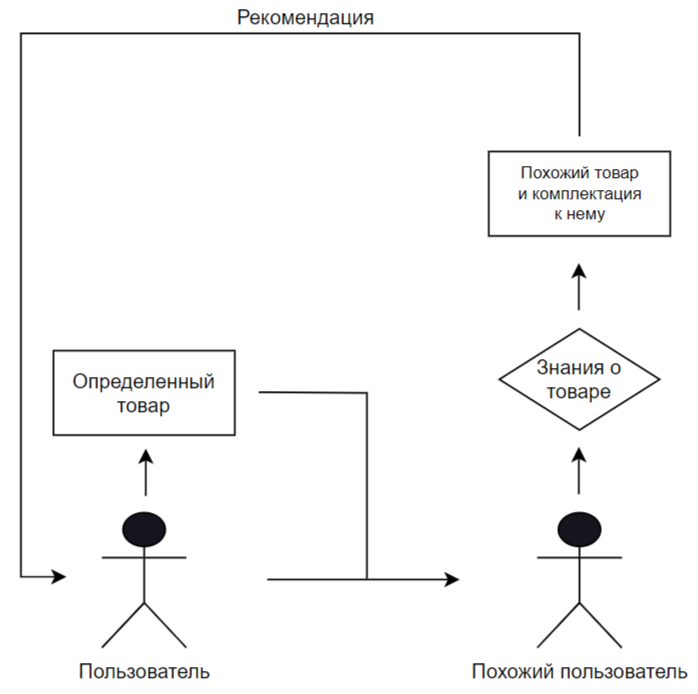


Рисунок 4 – Схема работы гибридной рекомендательной системы

Каждый тип рекомендательных систем, рассмотренных выше, имеет свои преимущества для решения конкретной задачи. Применение систем позволяет улучшить персонализированные рекомендации для пользователей, а также значительно увеличить маржинальность бизнеса.

**1.2. Методология разработки рекомендательных систем**

Разработка рекомендательных систем – это процесс, основанный на применении современных алгоритмов, методов анализа данных и машинного обучения. В контексте интернет-магазинов задача состоит в создании модели, которая будет анализировать предпочтения пользователей и рекомендовать подходящие товары, увеличивая таким образом эффективность продаж.

Разработка рекомендательного сервиса интернет-магазина включает реализацию следующих этапов:

1). Сбор данных. На данном этапе накопление информации проводится для создания некого профиля пользователей и моделей для прогнозирования на основе их предпочтений и поведения. Системе необходимо собрать как можно больше информации о пользователе, иначе она будет работать некорректно [Разуваев Кирилл Андреевич, Гринберг Ханна Эдуардовна, Маслова Александра Сергеевна, Веинский Владимир Андреевич, Милютин Алексей Валерьевич АНАЛИЗ СОВРЕМЕННЫХ ПОДХОДОВ В ПРОЕКТИРОВАНИИ РЕКОМЕНДАТЕЛЬНЫХ СИСТЕМ].

2). Предварительная обработка данных. Предобработка данных – неотъемлемая часть машинного обучения, включающая в себя подготовку и очистку данных перед их использованием. Качество данных оказывает непосредственное влияние на точность и производительность моделей. Недостатки в данных, такие как пропуски, дубликаты и выбросы, могут привести к неправильным выводам и ухудшению результатов обучения. [Хузяханов, Е. И. Эффективные методы предобработки данных: как подготовить данные для машинного обучения]. С предобработкой сталкиваются все современные системы машинного обучения. Ее эффективные методы улучшают качество входных данных, позволяют оптимизировать обучение моделей, что на финальном этапе приводит к улучшению точности прогнозируемых значений.

3). Выбор модели формирования рекомендаций.

4). Выбор инструментов и технологий разработки.

5). Интеграция в платформу.

В последнее время в качестве инструментария разработки моделей рекомендательных широко используются методы машинного обучения и нейронные сети. Нейронные сети представляют собой математические модели, построенные на аналогии с биологической нервной системой человека, которые состоят из нейронов, объединенных в слои. В отличие от классических методов, которые часто предполагали линейные зависимости, нейронные сети способны моделировать нелинейные отношения, что позволяет им быть более гибкими и точными в прогнозировании. Каждый нейрон принимает входные данные, проводит над ними операции и передает результат следующему слою. Глубокие нейронные сети с большим количеством слоев позволяют моделировать сложные взаимосвязи и достичь высокой точности прогнозирования.

Глубокое обучение – совокупность широкого класса методов, основанных на искусственных нейронных сетях. Преимущество глубоких нейронных сетей –  возможность улавливать малейшие зависимости в данных, благодаря чему они нашли широкое применение в задачах построения рекомендаций, в том числе на основе сессий[https://www.sberbank.ru/ru/person/kibrary/experts/obzor-metodov-postroeniya-rekomendatelnykh-sistem-na-osnove-sessij].

Различают разные виды нейронных сетей [https://www.sberbank.ru/ru/person/kibrary/experts/obzor-metodov-postroeniya-rekomendatelnykh-sistem-na-osnove-sessij].

1). Рекуррентные нейронные сети (RNN). Главная особенность рекуррентных нейронных сетей заключается в наличии внутреннего состояния скрытых слоев. Стандартные RNN обновляют скрытое состояние , используя следующую функцию обновления:

,

где – гладкая ограниченная функция, – входные данные слоя.

Проблема базовых RNN – неспособность удерживать в памяти длинных цепочек последовательностей. Это связано с проблемой затухающего/взрывающегося градиента. Решить эту проблему призваны сети долгой краткосрочной памяти (Long Short-Term Memory, LSTM).

2). Сети долгой краткосрочной памяти (Long Short-Term Memory, LSTM). В LSTM сетях была успешно решена проблема исчезновения или взрыва градиентов в процессе обучения с использованием метода обратного распространения ошибки. LSTM сеть контролируется рекуррентными вентилями, известными как «вентили забывания», которые позволяют ошибкам распространяться назад во времени через потенциально бесконечное количество виртуальных слоев. Этот механизм обеспечивает обучение LSTM сети с возможностью сохранения информации о тысячах и даже миллионах временных интервалов в прошлом. Таким образом, преимущество LSTM перед RNN в прогнозировании заключается в их способности более эффективно обрабатывать долгосрочные зависимости и сохранять информацию на протяжении длительных временных интервалов, что позволяет им давать более точные и надежные прогнозы.

3). Графовые нейронные сети (GNN). Они способны улавливать более сложные зависимости между объектами в сессии благодаря ее представлению в качестве ориентированного графа. Структура графа позволяет хранить информацию о различных взаимосвязях между пользователями и объектами. После чего GNN итеративно агрегирует информацию о соседях каждой вершины и обновляет ее [Ниничук М. М., Намиот Д. Е. ОБЗОР МЕТОДОВ ПОСТРОЕНИЯ РЕКОМЕНДАТЕЛЬНЫХ СИСТЕМ НА ОСНОВЕ СЕССИЙ]. В зависимости от того, какая информация и как представляется в виде графа, можно выделить следующие типы графовых нейронных сетей для рекомендаций:

– GNN, основанные на информации о взаимодействиях пользователей и объектов. Информация представляется в виде двудольного графа, первая часть вершин представляет пользователей, вторая – объекты.

– GNN, дополненные информацией о социальных взаимосвязях. Помимо графа с информацией о взаимодействии пользователей и объектов имеется граф социальных взаимодействий, характеризующих взаимоотношения пользователей между собой. Другой вариант – имеется единственный гетерогенный граф, одновременно содержащий всю описанную ранее информацию.

– GNN, дополненные информацией из базы знаний. В данных сетях известна некоторая информация об объектах, представленная либо в виде отдельного графа, либо как отдельный тип рёбер в гетерогенном графе. Обычно используется в рекомендациях, которые связывают объекты на основе их некоторых атрибутов.

Как отмечалось выше, важным этапом формирования рекомендаций является выбор методов и технологий разработки.

В данном исследовании сервис будет выполнен в виде онлайн-консультанта интернет-магазина. Для этого будут использоваться инструменты и технологии HTML, CSS, PHP, SQL, ChatGPT, а также Python.

HTML – это Hypertext Markup Language, язык гипертекстовой разметки. Он позволяет размечать содержимое так, чтобы сделать его более понятным для устройств и пользователей [Фрэйн Бен Отзывчивый дизайн на HTML5 и CSS3 для любых устройств. 3-е изд].

CSS – каскадные таблицы стилей, мощный механизм управления представлением отдельных документов или их наборов. Он позволяет управлять визуальной составляющей сайта и интернет ресурса [Мейер Э. CSS – каскадные таблицы стилей.].

PHP – распространённый язык программирования общего назначения с открытым исходным кодом, который сконструировали специально для веб-разработки и который встраивают в HTML-разметку [https://www.php.net/manual/ru/introduction.php]. Среди всех языков программирования PHP уверенно входит в первую десятку, а среди языков веб-программирования является безусловным лидером. На базе этого языка создаются социальные сети, системы управления содержимым сайта, корпоративные сайты, блоги, интернет-магазины и многие другие сервисы. Также данный язык программирования отлично поддерживает работу с базами данных, что является его преимуществом.

SQL (Structured Query Language – язык структурированных запросов) является языком реляционных баз данных, которые уже на протяжении многих лет заслуженно пользуются признанием у разработчиков информационных систем. Сочетание простоты и наглядности основных понятий с их строгим математическим обоснованием обеспечили широкое распространение реляционных баз данных. Данный язык позволяет быстро и четко работать с базами данных [Кара-Ушанов, В.Ю. К21 SQL — язык реляционных баз данных : учебное пособие / В.Ю. Кара-Ушанов.— Екатеринбург : Изд-во Урал. ун-та, 2016.— 156 с.].

ChatGPT – чат-бот с генеративным искусственным интеллектом, разработанный компанией OpenAI, способный работать в диалоговом режиме, поддерживающий запросы на естественных языках. Принцип работы системы заключается в общении с пользователем, генерированием текстов на разных языках, включая русский, которые относятся к различным областям [Рупасов, К. Р. Нейросеть chatgpt: революция в мире искусственного интеллекта / К. Р. Рупасов, Е. А. Тесленко, С. А. Пионтковская // Вестник науки. – 2024. – Т. 1, № 7(76). – С. 586-592. – EDN ERWSQU.]. Эта технология позволит общаться с пользователем без привлечения для этого человека.

Выше перечисленные технологии обеспечат создание веб-интерфейса чата с онлайн-консультантом. СhatGPT позволит не только автоматизировать общение с пользователем, но и будет являться центром связи между рекомендательным модулем и модулем веб-интерфейса (поисковым).

В настоящее время в литературе появляется все больше исследований в области разработки рекомендательных систем. Рассмотрим подробнее подходы отдельных авторов.

В статье Ковалева И. Ю. и Бондаренко В. И. сравниваются алгоритмы коллаборативной рекомендательной системы для интернет-магазинов по их эффективности [Ковалев, И. Ю. Подходы к построению рекомендательных систем для интернет-магазинов / И. Ю. Ковалев, В. И. Бондаренко]. Авторы использовали репозиторий Microsoft Recommenders, который предоставляет набор утилит для разработки и тестирования моделей. MuvieLens dataset был выбран как набор данных. Для оценки эффективности алгоритмов рассматривались такие метрики, как MAP (средняя точность), precision (точность), recall (полнота) и nDCG (нормализованная дисконтированная совокупная прибыль). В процессе работы были протестированы ALS, SVD, SAR, NCF, FastAI, BPR, BiVAE и LightGCN алгоритмы. В результате тесты показали, что SAR хорошо подходит для первоначальной выборки из-за своих показателей скорости и качества. Алгоритм BPR более эффективны для финальной стадии рекомендаций, так как является вторым по точности предсказаний после BiVAE, но выигрывает у лидера за счет скорости обучения и скорости предсказания. Таким образом, можно полагать, что модели SAR и BiVAE являются наиболее эффективными алгоритмами, позволяющими создать рекомендации, которые соответствуют личным предпочтениям каждого пользователя.

В работе Беспалова М. О. и Косникова С. Н. описываются результаты применения методов машинного обучения в экономическом моделировании и прогнозировании, а также в разработке рекомендательных систем для электронной торговли [Беспалов, М. О. Исследование и реализация рекомендательных алгоритмов в сфере торговли видеоиграми / М. О. Беспалов, С. Н. Косников // Экономика в социокультурном пространстве современности: проблемы, решения, прогнозы : Материалы X Международной научно-практической конференции, Владимир, 23 ноября 2023 года. – Владимир: АРКАИМ, 2023. – С. 14-22]. Авторы рассматривают такие подходы, как контентно-ориентированные методы, коллаборативная фильтрация и нейронные сети. Для реализации моделей использовался датасет о видеоиграх и пользователях, обработка данных реализовалась с помощью библиотеки Pandas. В своем исследовании Беспалов М.О. и Косников С.Н. применяют метод TF-IDF для анализа описаний видеоигр, который использовался для нахождения сходства между играми и генерируемыми рекомендациями. Авторы применяли коллаборативную фильтрацию для анализа предпочтений пользователей. Матрица схожести, которая строилась с помощью косинусного расстояния, позволяла подбирать рекомендации, основанные на поведении похожих игроков. На основе методов библиотеки LightFM протестированы матричные модели, для оценки качества построенных моделей использовались метрики средней абсолютной и среднеквадратической ошибок. В результате, наиболее успешной конфигурацией, которая показала самую высокую точность, была модель с функцией потерь WARP. Также в работе была разработана нейронная сеть, архитектура которой включала несколько слоев с различным числом нейронов и использовала методы снижения размерности данных. Для обучения нейросетевой модели авторы использовали методы библиотек Sklearn и PyTorch. Путем сравнения двух подходов было выяснено, что нейронная сеть превосходит модель матричной факторизации по качеству рекомендаций, но уступает в скорости обучения.

В статье авторов Малынова А.А. и Прохорова И.В. приводятся этапы разработки чат-бота на основе ChatGPT для системы рекомендаций концертных мероприятий [Малынов, А. А. Разработка чат-бота на базе ChatGPT для системы рекомендации концертных мероприятий]. Архитектура включает в себя: чат-бот, службу аутентификации, службу обогащения пользовательского контекста и программный интерфейс ChatGPT. Основной системой взаимодействия с пользователем является интерфейс чат-бота, а служба аутентификации обеспечивает хорошую безопасность системы. Для улучшения рекомендаций используется пользовательский контекст, в который входит информация о предыдущих посещениях. Ядром является ChatGPT, который обрабатывает запросы пользователей и предоставляет персонализированные рекомендации. Важную роль в работе всей системы играет подготовка входных данных. Следующим этапом создается структура для каждого запроса, которая включает инструкции для модели, список мероприятий, информацию о пользователе и сам запрос. Это позволяет модели эффективно анализировать данные и предоставлять точные рекомендации. После проведенных экспериментов авторами сделан вывод, что в большинстве случаев чат-бот успешно справляется с задачей предоставления релевантных мероприятий, учитывая контекст и предпочтения пользователей. Результаты исследования обосновывают тот факт, что использование технологий ChatGPT может значительно улучшить качество рекомендаций, а также повысить удобство и эффективность взаимодействия с системой [Малынов, А. А. Разработка чат-бота на базе ChatGPT для системы рекомендации концертных мероприятий / А. А. Малынов, И. В. Прохоров].

В работе [Куликов, К. Г. Применение нейросетей в рекомендательных системах / К. Г. Куликов, В. В. Бугаков, Д. К. Реутов] Куликова К.Г., Бугакова В.В. и Реутова Д.К. основное внимание уделено современным архитектурам нейронных сетей и их компонентам, включая эмбеддинг-слои, рекуррентные ячейки, полносвязные слои и блоки трансформеров. Авторы утверждают, что основными преимуществами DL-моделей являются способность использовать дополнительные данные и их гибкость в работе с несколькими задачами одновременно. Одним из важнейших этапов разработки является использование эмбеддинг-слоев, которые преобразуют идентификаторы пользователей и товаров в плотные векторы, что позволяет оценивать вероятность взаимодействия между ними. Данный этап схож с методом матричной факторизации, но опережает его за счет использования глубоких архитектур. Авторы также рассматривают метаданные, такие как пол, возраст, местоположение пользователя, а также цена и категория товаров, которые позволяют учитывать больше факторов при генерации рекомендаций. В работе отдельное внимание уделяется рекомендациям на основе сессий. Данный подход использует последовательность действий пользователя в рамках одной сессии, таких как просмотр товаров, для прогнозирования наиболее релевантных предложений. Для анализа авторы используют рекуррентные нейронные сети. В заключении можно сказать, что рассмотренная технология с каждым годом только укрепляет свои позиции, и уже стала стандартной практикой, однако, в перспективе развитие рекомендательных систем будет только ускорятся, что позволит составлять полное представление о потребностях и желаниях каждого пользователя[Куликов, К. Г. Применение нейросетей в рекомендательных системах / К. Г. Куликов, В. В. Бугаков, Д. К. Реутов].

Краткий обзор авторских подходов к решению проблемы разработки рекомендательных систем представлен в таблице 1.

Таблица 1 – Краткий обзор авторских подходов к решению проблемы разработки

рекомендательных систем

|  |  |  |  |
| --- | --- | --- | --- |
| № | Авторы | Методы | Подход создания  рекомендательной системы |
| 1 | Ковалев И. Ю.  Бондаренко В. И.  [Ковалев, И. Ю. Подходы к построению рекомендательных систем для интернет-магазинов / И. Ю. Ковалев, В. И. Бондаренко] | Методы ALS, SVD, SAR, NCF, FastAI, BPR, BiVAE  и LightGCN. | Коллаборативная  фильтрация |
| 2 | Беспалов М. О.  Косников С. Н.  [Беспалов, М. О. Исследование и реализация рекомендательных алгоритмов в сфере торговли видеоиграми] | Метод TF-IDF, модель  матричной факторизации, методы снижения  размерности данных. | Контентная фильтрация,  коллаборативная фильтрация и нейронные сети |
| 3 | Малынов А. А.  Прохоров И. В.  [Малынов, А. А. Разработка чат-бота на базе ChatGPT для системы рекомендации концертных мероприятий] | Технологии ChatGPT. | Нейронные сети |
| 4 | Куликов К. Г.  Бугаков В. В.  Реутова Д. К.  [Куликов, К. Г. Применение нейросетей в рекомендательных системах. – Курск: ЗАО «Университетская книга», 2024. – С. 62-69] | Методы создания  мультимодальных  эмбеддингов, разработка рекомендаций на основе сессий, рекуррентные нейронные сети. | Нейронные сети |
| 5 | Чжан Ч.  Афанасьев Г. И.  [Чжан, Ч. Основные технологии и перспективы эволюции персонализированных рекомендательных систем / Ч. Чжан, Г. И. Афанасьев // E-Scio. – 2022. – № 4(67). – С. 309-320] | Динамические модели  интересов пользователей, моделирование объектов рекомендации. | Контентная фильтрация |
| 6 | Леонова О. С.  Коломойцева И. А.  [Леонова, О. С. Разработка рекомендательной системы на основе контентно- ориентированной фильтрации для книжного интернет-магазина] | Методы построения  матрицы схожести и  косинусного расстояния. | Контентная фильтрация |

Проанализировав данные таблицы 1, можно сделать вывод, что отечественные специалисты используют различные подходы в разработке рекомендательных систем. В основном в их работах преобладают контентная и коллаборативная фильтрация. Многие авторы выделяют гибридные системы, как самую эффективную фильтрацию. Перспективным направлением разработки рекомендаций считается использование нейронных сетей. При разработке авторы использовали такие современные технологии, как HTML и CSS, Python, его библиотеки и базы данных.

**1.3. Обзор программных решений в предметной области**

С каждым днем все больше и больше крупнейших интернет-магазинов внедряют в свои среды рекомендательные системы. Их применение позволяет пользователям тратить меньше времени на поиск нужного товара среди их огромного количества. Рассмотрим одни из самых популярных интернет-магазинов, которые используют в своем функционале рекомендательные системы.

Одним из самых популярных и крупнейших сервисов считается Ozon. Его рекомендательная система, работает на основе анализа данных о пользователях, товарах и их взаимодействии. Она анализирует множество источников данных, включая историю покупок и просмотров каждого пользователя, а также товары, добавленные в корзину или избранное. В результате все эти данные помогают сформировать представление о пользователе и определить его предпочтения. Помимо этого, система анализирует поведение пользователей, куда входят частота посещений, поисковые запросы, клики на товары и время, проведенное на странице. Важную роль играют данные о самих товарах, включая их описание, характеристики, отзывы, рейтинг и цену [https://docs.ozon.ru/legal/terms-of-use/site/algorithms/recomendation-algorithms/].

Ozon применяет несколько подходов к построению рекомендаций:

– контентно-ориентированный подход;

– коллаборативный подход;

– гибридный подход.

Для обработки данных и предсказания предпочтений система использует технологии машинного обучения, включая нейронные сети и алгоритмы ранжирования. Предсказание производится в несколько этапов [https://docs.ozon.ru/legal/terms-of-use/site/algorithms/recomendation-algorithms/]:

1. Вычисляются все фичи (свойства, от английского features) каждого отдельного товара, которые могут повлиять на факт продажи. Каждая фича – это вещественное число. Например, фичей может быть рейтинг товара или его цена.
2. К набору фичей применяется модель машинного обучения, предсказывающая вероятность совершения действия на основе значений фичей. В качестве модели Ozon использует один из методов машинного обучения – градиентный бустинг деревьев решений.
3. Модель прогнозирует число, которое используется для оценки вероятности.

Главной целью моделирования является наиболее точное предсказание вероятности взаимодействия на основании фичей этого товара. Фичи товара содержат информацию о том, какими свойствами обладали товары, приобретённые покупателями в прошлом, и какие из этих свойств имеет ранжируемый товар. Модель на основании данных из прошлого предсказывает покупательское поведение в будущем. Также при прочих равных условиях покупатели предпочитают товары с большим количеством отзывов. Чем больше у товара отзывов, тем выше предсказанная моделью вероятность. При этом разные фичи имеют разную предсказательную силу и по-разному влияют на вероятность покупки. Модель обучается на данных о зависимости между фичами товаров и итоговыми покупками при помощи специального алгоритма. Многие фичи рассчитываются на основе покупательского поведения: покупатели Ozon влияют на положение товаров в выдаче рекомендаций своими действиями. Используются и графовые модели, которые строят связи между пользователями и товарами, помогая системе понять, какие товары покупают вместе или какие клиенты имеют схожие предпочтения [https://docs.ozon.ru/legal/terms-of-use/site/algorithms/recomendation-algorithms/].

Рекомендательная система Ozon учитывает сезонные факторы и тренды. Например, в преддверии праздников она может активно предлагать подарки, а летом – сезонные товары, такие как туристическое снаряжение. Важным элементом является и обработка поисковых запросов. Интеграция поиска с рекомендациями позволяет выводить результаты, релевантные не только запросу, но и общему профилю пользователя. Для повышения эффективности работы система постоянно оптимизируется, проводятся A/B тесты, позволяющие оценить влияние изменений в алгоритмах на ключевые метрики, такие как клики, конверсии и средний чек. Оценка обратной связи от пользователей помогает улучшать качество рекомендаций. Использование облачных технологий обеспечивает обработку огромных объемов данных в реальном времени, что позволяет быстро адаптировать рекомендации под действия пользователей.

Таким образом, рекомендательная система Ozon представляет собой сложную и многоуровневую структуру, основанную на современных технологиях анализа данных и машинного обучения. Она помогает клиентам находить нужные товары быстрее и удобнее, создавая персонализированный опыт и повышая эффективность продаж.

Еще одним примером интернет-магазина, использующего технологии разработки рекомендаций, можно считать «СберМегамаркет». В его рекомендательной системе сбор информации о предпочтениях пользователей происходит способом логирования – ведения технических записей о действиях пользователей на платформе и сохранения таких записей на внутреннем хранилище для возможности последующего использования. К данным о предпочтениях относятся такие сведения, как просмотр товаров и приобретенные товары.

Систематизация и анализ сведений о предпочтениях пользователей происходит, исходя из цели создания рекомендательной модели и рекомендаций, которые необходимо получить (например, подбор наиболее популярных товаров среди пользователей, которые показывают схожие сценарии поведения на платформе). После определения цели происходит подбор требуемых для реализации цели данных о предпочтениях, на основе которых будет выстроена рекомендательная модель. Выделяются фичи: например, количество заказов пользователя с данным товаром при наличии скидки на товар, количество покупок пользователем товаров со скидкой, предыдущая цена товара, количество заказов с данным товаром и другие. После систематизации данных к сведениям о предпочтениях пользователей и фичам применяются методы машинного обучения. Рекомендательная модель формирует набор параметров, описывающих зависимости между входными данными (предпочтениями пользователей и фичами) и ответом (финальной рекомендацией). Таким образом, рекомендательная модель производит оценку вероятности того, что пользователь совершит определённое действие в интернет-магазине (например, добавит товар в корзину, купит товар).[ <https://kuper.ru/sp/recommendations>]

В практике организации интернет-торговли использует рекомендательные системы и сервис «Литрес». Сервис применяет рекомендательные технологии, чтобы предсказать заинтересованность пользователя в услуге (книги, аудиокниги, подкасты) на основе предпочтений других пользователей, похожих на данного. Первым этапом разработки рекомендаций является сбор и анализ данных пользователей и их действий. После препроцессинга исходных данных создается разряженная матрица с ID пользователей в строках и ID элементов в столбцах. Весь алгоритм строится на выполнении одного из методов (моделей) матричной факторизации. Матричная факторизация использует модель Alternating least squares (ALS). Эта модель изучает бинарную цель взаимодействия каждого пользователя с каждым объектом, но взвешивает каждое бинарное взаимодействие по доверительному значению уверенности в этом взаимодействии пользователя/объекта. В неявной реализации используются значения разреженной матрицы для представления достоверности. При невозможности определения рекомендаций пользователю с помощью матрицы в качестве рекомендаций используются самые покупаемые объекты на языке пользователя за определенный период. [https://www.litres.ru/company/recommendations/]. Рекомендательные технологии, используемые в сервисе, обрабатывают данные на основе совершенных пользователем действий.  Примерами таких действий могут быть: посещение страницы / экрана, клик по элементу, прокрутка, покупка книги, чтение книги, оценка книги и другие.  Хранение данных осуществляется на серверах cервиса.

Еще одним популярным сервисом в нашей стране является платформа Wildberries. Данный сервис использует рекомендательные технологии для формирования персональных и неперсональных подборок товаров, из которых пользователь может выбрать наиболее подходящие. Для формирования рекомендаций используются такие математические методы как [https://legal.wildberries.ru/rules-for-the-use-of-recommendation-technologies/country/ru/lang/ru/]:

– нейронные сети,

– ансамбли решающих деревьев,

– коллаборативные фильтрации,

– градиентный бустинг,

– классические модели машинного обучения.

К основным этапам формирования рекомендаций на платформе Wildberries можно отнести:

1. Формирование набора данных о последней активности пользователя.
2. Предсказание последующих действий пользователя на основе полученных данных с использованием различных алгоритмов.
3. Формирование подходящих рекомендаций на основе объединения результатов работы нескольких алгоритмов.
4. Обработка и сохранение сформированных рекомендаций на основании схожести характеристик товаров и информации о пользователе.

Рекомендательную систему использует и платформа Яндекс.Маркет. Рекомендации Яндекс.Маркета помогают покупателям выбрать подходящие товары и сэкономить время на их поиске[https://yandex.ru/legal/recommendations/ru/#market]. Этому способствуют такие данные как: история поиска товаров, история покупок, просмотры и клики, товары из корзины и избранного, любимые бренды и т.д. В основе рекомендаций лежат такие математические методы как: коллаборативная фильтрация, контентная фильтрация, графовые модели, градиентный бустинг, нейронные сети.

В таблице 1 приведены результаты сравнительного анализа функционала рекомендательных систем торговых платформ.

Таблица 2 – Сравнительный анализ торговых платформ

|  |  |
| --- | --- |
| Торговая платформа | Методы и подходы к разработке рекомендательных систем |
| Ozon | Контентная, коллаборативная, гибридная фильтрация, градиентный бустинг, графовые модели. |
| СберМегамаркет | Машинное обучение, логирование, анализ предпочтений. |
| Литрес | Матричная факторизация (ALS). |
| Wildberries | Нейросети, ансамбли деревьев, градиентный бустинг,  коллаборативная фильтрация. |
| Яндекс.Маркет | Коллаборативная и контентная фильтрация, графовые модели, бустинг, нейронные сети. |

Обзор сервисов позволяет сделать вывод, что все они стремятся к улучшению рекомендательных систем, большинство из них используют интеллектуальные методы такие, как методы искусственного интеллекта и машинного обучения, которые способствуют увеличению прибыли компаний.