

Министерство образования Республики Беларусь
Учреждение образования
Белорусский государственный университет
информатики и радиоэлектроники

УДК 004.032.26:655.42

Юргель
Владислав Юрьевич

Модели и методы построения и работы с каталогом книжных изданий

АВТОРЕФЕРАТ
на соискание академической степени
магистра технических наук

по специальности 1-40 80 05 – Математическое и программное обеспечение
вычислительных машин, комплексов и компьютерных сетей

Научный руководитель
Серебряная Л.В.
к.т.н., доцент

Минск 2020

КРАТКОЕ ВВЕДЕНИЕ

В современном мире рост количества доступных книжных изданий каждый год значительно увеличивается, что приводит к необходимости автоматизации формирования и поиска аналогичных книжных изданий на основе сформированных пользовательских предпочтений, что позволит моментально выполнять поиск книжных изданий и предлагать пользователю обновить свой текущий каталог для дальнейшего приобретения.

Наиболее эффективной областью для выполнения автоматизации обработки и предложения рекомендаций является область рекомендательных систем и методы распознавания текстов нейронной сетью повсеместно с компьютерной лингвистикой.

Для составления рекомендаций используется довольно немалое количество методов, чаще всего, когда требуется предсказать пользовательскую рекомендацию, для этого используются оценки пользователя и похожих пользователей. Такие методы уже давно применяются не только в сфере книжных изданий, но и в сфере кино и музыки.

Задача рекомендательной системы – проинформировать пользователя о товаре, который ему может быть наиболее интересен в данный момент времени. Клиент получает информацию, а сервис зарабатывает на предоставлении качественных услуг. Услуги – это не обязательно прямые продажи предлагаемого товара. Сервис также может зарабатывать на комиссионных или просто увеличивать лояльность пользователей, которая потом выливается в рекламные и иные доходы.

В зависимости от модели бизнеса рекомендации могут быть его основой, как, а могут быть просто удобным дополнительным сервисом (как, например, в каком-нибудь интернет-магазине одежды), призванным улучшить потребительский опыт и сделать навигацию по каталогу более удобной.

Персонализация онлайн-маркетинга – очевидный тренд последнего десятилетия. По оценкам McKinsey, 35% выручки Amazon или 75% Netflix приходится именно на рекомендованные товары и процент этот, вероятно, будет расти. Рекомендательные системы – это про то, что предложить клиенту, чтобы сделать его счастливым.

ОБЩАЯ ХАРАКТЕРИСТИКА РАБОТЫ

Цель и задачи исследования

Целью диссертационной работы является исследование моделей и методов обработки книжных изданий с дальнейшим применением методов для поиска на основе пользовательских рекомендаций.

Для достижения поставленной цели необходимо решить следующие задачи:

Изучить существующие модели и методы обработки и построения каталога книжных изданий.

Исследовать различные методы и модели обработки и построения каталога книжных изданий.

Разработать архитектуру программного средства обработки и построения каталога книжных изданий.

Объектом исследования являются пользовательские предпочтения, т.е. проанализированные и обработанные данные о предпочитаемых пользователем книжных изданиях.

Предметом исследования являются модели, методы и программные средства обработки и формирование поисковых запросов на основе пользовательских предпочтений.

Основной *гипотезой*, положенной в основу диссертационной работы, является разработка модуля, с помощью которого можно применить исследованные методы для рекомендательной системы, а также использовать атрибуты проанализированных книжных изданий в качестве объекта для поиска по рекомендациям. Данный модуль может помочь эффективно находить похожие элементы используя текущий набор атрибутов из пользовательских рекомендаций, что позволит ускорить процесс поиска и обработки заранее отфильтрованной информации.

Личный вклад соискателя

Результаты, приведенные в диссертации, получены соискателем лично. Вклад научного руководителя Л.В. Серебряной заключается в формулировке целей и задач исследования.

Опубликованность результатов диссертации

По теме диссертации опубликовано 3 печатных работы, две из которых в журнале «Вестник науки и образования».

Структура и объем диссертации

Диссертация состоит из введения, общей характеристики работы, четырех глав, заключения, списка использованных источников, списка публикаций автора и приложений.

В первой главе представлен анализ предметной области, выявлены основные существующие проблемы в рамках тематики исследования, показаны направления их решения. Вторая глава посвящена анализу существующих решений и предметной области, их достоинства и недостатки. В третьей главе предложена практическая реализация ПС для обработки и построения каталога книжных изданий. В четвертой главе предоставлена оценка полученных результатов и сравнение с аналогами.

Общий объем работы составляет 64 страницы, из которых основного текста – 45 страниц, 24 рисунка на 12 страницах, 5 таблиц на 4 страницах, список использованных источников из 10 наименований на 1 странице и 1 приложение на 10 страницах.

КРАТКОЕ СОДЕРЖАНИЕ

Во введении рассмотрено современное положение и роль рекомендательных систем в социально-экономической сфере, а также приведена проблема, которая приводит к необходимости их рассмотрения.

В первой главе приведен анализ аналогов с каталогом книжных изданий, а также указаны их недостатки в плане реализации рекомендательной системы для пользовательских предпочтений. Помимо это, также сформулирована и корректно поставлена задача,

В первой главе сформулирована и конкретно поставлена задача, дан обзор ее предметной области, подробно освещены моменты, с которыми можно столкнуться при решении данной задачи, а именно: проблемы рекомендательных систем и формирование удовлетворительных результатов для пользовательских рекомендаций.

Во второй главе рассмотрены модели и методы, обязательные для реализации рекомендательной системы, приведены их достоинства и недостатки, а также произведен выбор модели и метода в качестве дальнейшего исследования.

В третьей главе приведена реализация программного средства формирования каталога книжных изданий, рассмотрена его архитектурная часть и реализация основных модулей, в частности, реализация модуля рекомендательной системы на основе выбранной модели и метода.

В четвертой главе дана оценка полученным результатам исследования и реализации программного средства, а также их сравнение с аналогами, приведенными в первой главе.

В заключении приведены основные достигнутые результаты и возможные будущие шаги, направленные на развитие области.

ЗАКЛЮЧЕНИЕ

В работе была затронута тема исследования и реализации рекомендательной системы для формирования пользовательских рекомендаций на основе его предпочтений используя метода коллаборативной фильтрации. Для того, чтобы как можно более качественно решить поставленную задачу, были выполнены следующие шаги:

1. Проведен детальный анализ предметной области, а именно: исследованы научные статьи и литература по рекомендательным системам, исследовано количество книжных изданий и их потребность в рекомендациях для пользователей.

2. Исследованы существующие модели и методы для реализации рекомендательных систем.

3. Выбраны модель и метод для тщательного исследования и использования в качестве основы рекомендательной системы.

4. Предложена структура базы данных, необходимая для хранения и накопления данных.

5. Предложена реализация программного средства и необходимых модулей, в частности для работы с каталогом книжных изданий и пользовательских рекомендаций на основе предпочтений.

6. Произведена оценка результатов исследования и реализации рекомендательной системы, а также ее сравнение с аналогами.

7. Сделан вывод о том, какая модель наилучшим образом решает поставленную задачу.

8. Произведена оценка качества рекомендательной системе на основе метрики RMSE: 0,921352.

Поставленная задача была достигнута с использованием методов, позволяющих выполнять кластеризацию пользователей по определенным критериям, выявлять среднюю оценку кластера и, на ее основе, предлагать рекомендации пользователю. Несмотря на хорошее качество решения задачи, в настоящее время есть методы, которые позволяют улучшить качество модели, например, рекуррентные нейронные сети (Recurrent Neural Network; RNN).

СПИСОК ОПУБЛИКОВАННЫХ РАБОТ

1-А. Юргель, В.Ю. Методы построения и работы с каталогом книжных изданий. / В.Ю. Юргель // Вестник науки и образования. – 2019. – № 23 (77). Часть 1. – с. 10-11.

2-А. Юргель В.Ю. Сложности моделирования естественного языка. // В.Ю. Юргель // Вестник науки и образования. – 2019. – № 23 (77). Часть 1. – с. 11-12.

3-А. Юргель В.Ю. Исследование и сравнение нейронных сетей для классификации текста. / В.Ю. Юргель // 55-я юбилейная научная конференция аспирантов, магистрантов и студентов. – 2019. – с. 173-174.

Библиотека БГУМР