

MINISTERUL EDUCAȚIEI ȘI CERCETĂRII AL REPUBLICII MOLDOVA

Universitatea Tehnică a Moldovei

Facultatea Calculatoare Informatică și Microelectronică

Departamentul Ingineria Software și Automatică

Admis la ușinare
Șef de departament:
Fiodorov I. dr., conf.univ.

„___” ianuarie 2022

Sisteme de recomandare pentru servicii on-line
Рекомендательные системы для онлайн-сервисов

Teza de master

Student: _____ **Scutelnic Ilie, TI-201M**

Coordonator: _____ **Nastasenco Veaceslav, conf. univ.**

Consultant: _____ **Cojocarui Svetlana, asist. univ.**

Chișinău, 2022

Аннотация

В данной дипломной работе описываются принципы и механизмы работы рекомендательных систем. Технология, опираясь на статистические и демографические данные, а также на поисковые запросы, осуществляет выборку релевантных для пользователя объектов, таких как товары, кинофильмы, сторонние онлайн-сервисы и т.д. Первая глава посвящена описанию основных понятий, определению терминов, а также базовой теории в области рекомендательных систем. Далее будет осуществлен более подробный анализ существующих механизмов подборки рекомендаций и основных проблем, связанных с внедрением таких систем в онлайн-сервисы. Будут рассмотрены как устаревшие тривиальные алгоритмы, так и более продвинутые, основанные на кластеризации и матричных разложениях.

Целью данной работы является подробный анализ существующих механизмов, анализ и минимизация существующих проблем, связанных с ними.

Annotation

This thesis describes the principles and mechanisms of recommendation systems. The technology, based on statistical and demographic data, as well as on search queries, selects objects relevant to the user, such as products, movies, third-party online services, etc. The first chapter is devoted to a description of basic concepts, a definition of terms, and basic theory in the field of recommendation systems. Next, there will be a more detailed analysis of the existing mechanisms for picking recommendations and the main problems associated with the implementation of such systems in online services. Both outdated trivial algorithms and more advanced ones based on clustering and matrix decompositions will be considered.

The purpose of this paper is to analyze in detail the existing mechanisms, to analyze and minimize the existing problems associated with them.

Adnotare

Această teză descrie principiile și mecanismele sistemelor de recomandare. Tehnologia, bazată pe date statistice și demografice, precum și pe interogări de căutare, selectează elemente relevante pentru utilizator, cum ar fi produse, filme, servicii online de la terți etc. Primul capitol este dedicat descrierii conceptelor de bază, definirii termenilor și teoriei de bază în domeniul sistemelor de recomandare. Aceasta va fi urmată de o analiză mai detaliată a mecanismelor de recomandare existente și a principalelor probleme legate de implementarea unor astfel de sisteme în serviciile online. Vor fi luați în considerare atât algoritmi triviali învechiți, cât și algoritmi mai avansați, bazați pe clusterizare și descompunerea matricelor.

Scopul acestei lucrări este de a analiza în detaliu mecanismele existente și de a analiza și minimiza problemele existente asociate acestora.

СОДЕРЖАНИЕ

ВВЕДЕНИЕ	1
1 ОПРЕДЕЛЕНИЕ.....	2
1.1 Актуальность темы	2
1.2 Обзор и постановка задачи	3
1.3 Основные типы рекомендательных систем	4
1.4 Описание работы рекомендательных систем	6
1.5 Основные проблемы рекомендательных систем	8
1.6 Примеры известных рекомендательных систем	9
2 КОРРЕЛЯЦИОННЫЕ И ЛАТЕНТНЫЕ МОДЕЛИ	12
2.1 Основные подходы коллаборативной фильтрации.....	12
2.1.1 Определение и обозначения.....	13
2.1.2 Примеры.....	13
2.1.3 Модели коллаборативной фильтрации.....	16
2.2 Корреляционные модели	17
2.2.1 Модели, основанные на хранении данных.....	17
2.2.2 Функции близости.....	20
2.2.3 Задача восстановления пропущенного значения.....	21
2.3 Модели латентной семантики	22
2.3.1 Матричные разложения.....	23
2.3.2 Учёт дополнительных признаков данных.....	26
2.3.3 Оценивание качества рекомендаций.....	27
3 ГРАФОВЫЕ МОДЕЛИ.....	30
3.1 Основные понятия навигации в пространстве предпочтений.....	30
3.2 Модель пространства предпочтений и постановка задачи.....	35
3.2.1 Графовая модель при «холодном старте»	35
3.2.2 Графовая модель с полностью известной матрицей предпочтений	36
3.2.3 Графовая модель с частично известной матрицей предпочтений	38
3.3 Алгоритм нахождения максимальных элементов в графе предпочтений	40
3.4 Параллельный поиск максимальных элементов в графе предпочтений.....	45
3.5 Поиск максимально релевантных элементов элементов на графах	50
3.5.1 Постановка задачи	50
3.5.2 Локальная дискретная оптимизация со случайным поиском на графах	52
3.5.3 Построение функции внутреннего расстояния.....	54
3.6 Анализ работы системы релевантного выбора	59

ЗАКЛЮЧЕНИЕ.....	63
СПИСОК ИСПОЛЬЗОВАННОЙ ЛИТЕРАТУРЫ.....	64

ВВЕДЕНИЕ

В наши дни широкого ассортимента человек регулярно сталкивается с вопросом выбора. Возможность выбора на сегодняшний день существует практически в любом из видов услуг - выбор продукта в супермаркете, модели техники в магазине, фильма в кинотеатре и т.д. На сегодняшний день выбор продукции в той или иной сфере в большинстве случаев настолько широк, что отдельный потенциальный потребитель не имеет возможности изучить каждый вариант отдельно.

В то же время среди поставщиков услуг/товаров ведется непрекращающаяся конкуренция за клиента. Перед поставщиком стоит цель – осуществить реализацию своего товара максимальному числу потребителей в максимально возможном объеме. Владельцы традиционных торговых площадок используют традиционные инструменты для привлечения покупателей: мерчандайзинг, активная реклама товара/услуги и т.д. На онлайн-площадках активно набирают обороты внедренные в них рекомендательные системы, способные при наличии определённой информации о профиле пользователя определить объекты, имеющиеся в ассортименте, которые вероятнее всего его заинтересуют.

Сегодня сложно представить себе крупную онлайн-площадку, основанную на продажах или подписках без блока «Возможно Вас заинтересует». Практически все интернет-магазины, стриминговые платформы, крупные видеохостинги имеют такие блоки, содержимое которых стремится к релевантности для каждого пользователя. Цель данной работы – исследование и разработка эффективных алгоритмов рекомендательных систем в условиях большого числа пользователей и наличия широкой базы продукции, и определение основных проблем с ними связанных.

СПИСОК ИСПОЛЬЗОВАННЫХ ИСТОЧНИКОВ

1. Collaborative filtering. Режим доступа: https://en.wikipedia.org/wiki/Collaborative_filtering
2. Изучение принципа работы рекомендательных систем. Режим доступа: <https://neurohive.io/ru/osnovy-data-science/rekomendatelnyesistemy-modeli-i-ocenka/>
3. What are recommendation systems in machine learning. Режим доступа: <https://www.analyticssteps.com/blogs/what-are-recommendationsystems-machine-learning>
4. Бритвина Е.В. Исследование и разработка алгоритмов рекомендательных систем на основе графовых моделей данных. Ученая диссертация. 2015
5. Рекомендательные системы на практике. Режим доступа: <https://www.machinelearningmastery.ru/recommender-systems-inpracticecef9033bb23a/>
6. Коллаборативная фильтрация. Режим доступа: <http://www.machinelearning.ru/wiki/images/archive/9/95/20140413184117%21Voron-ML-CF.pdf>