

Министерство образования Республики Беларусь

Учреждение образования
Белорусский государственный университет
информатики и радиоэлектроники

УДК 004.031.2 – 04.58

Гапутин
Всеволод Денисович

ПЛАТФОРМА УПРАВЛЕНИЯ И ОРГАНИЗАЦИИ МЕРОПРИЯТИЙ ДЛЯ
РАЗВИТИЯ КОРПОРАТИВНОГО ОБРАЗОВАНИЯ

АВТОРЕФЕРАТ

диссертации на соискание степени магистра технических наук
по специальности 1-31 80 10 «Теоретические основы информатики»

Научный руководитель

Сердюков Роман Евгеньевич

Кандидат технических наук, доцент
кафедры ИИТ

Минск 2020

КРАТКОЕ ВВЕДЕНИЕ

Обучающиеся стремятся выбрать такие курсы, которые скорее всего будут отвечать их интересам и возможностям. Это решение может сильно повлиять на их мотивацию и их будущие достижения, обычно не предлагается никакой поддержки, чтобы помочь им в этом решении. Неправильный выбор курса может привести к демотивации, в результате чего обучающийся может бросить учебу или не воспользоваться всей полнотой предоставленных возможностей. Следовательно, необходимо предоставить обучающимся необходимую поддержку, чтобы помочь им выбрать оптимальный набор курсов, в соответствии с их потребностями и интересам. Систему рекомендаций можно описать как набор инструментов и методик, используемых с целью предоставления предложений по различным объектам лицам, не обладающим достаточной компетенцией для оценки потенциально огромного числа имеющихся альтернатив. Рекомендации стали использоваться в электронной торговле, в сфере доступа к информации, развлечений и различных видов услуг, предоставляя предложения, которые эффективно сокращают большие информационные пространства, чтобы пользователи были направлены к тем событиям и услугам, которые наилучшим образом отвечают их потребностям и предпочтениям.

Рекомендация курсов непростая задача, так как рекомендация предмета должна быть не только интересной для учащегося, но схожа с его навыками и учитывает любые ограничения (например, предварительные условия предмета).

Современные рекомендационные системы в основном рекомендуют курсы в связи с их содержанием, не учитывая того, как этот курс влияет на общую успеваемость обучающихся. Кроме того, одни решения предлагают использование чрезмерно субъективных критериев при подготовке рекомендаций, в то время как другие требуют слишком большого участия учащихся.

Для достижения целей магистерской диссертации были поставлены следующие задачи:

1. Рассмотреть существующую литературу, источники и основные понятия, связанные с проблематикой систем рекомендаций.

2. Изучить применение интеллектуального анализа данных в образовательном контексте. Определить систему оценивания, на основе которой можно классифицировать системы рекомендаций в области сбора данных в сфере образования. Оценить современное состояние рекомендаций в образовательной среде и существующие методы моделирования навыков и успеваемости учащихся.

3. Описать применение сингулярного разложения матрицы в дополнение к пользовательской коллаборативной фильтрации. Описать использование модели ASAP-классификаторов с отсутствующими данными об учащемся, в качестве подготовительного шага к рекомендации.

4. Описать практический пример, основанный на реальных данных учащихся.

5. Оценить предложенный подход рекомендаций и соответствующие результаты.

ОБЩАЯ ХАРАКТЕРИСТИКА РАБОТЫ

Цель и задачи исследования

Целью диссертационной работы является исследование и описание эффективного алгоритма рекомендательной системы в образовательной области для решения задачи преугадывания предпочтений пользователей, на основании их предыдущего поведения, позволяющей создавать рекомендации с приемлемым уровнем релевантности в условиях большого числа курсов и учащихся при неполной информации об их оценках и предпочтениях.

Для достижения поставленной цели необходимо решить следующие задачи:

1. Изучить существующие подходы и методы построения рекомендательных систем;
2. Ознакомиться с существующими аналогичными решениями и программными продуктами и определить основные функциональные требования к системе;
3. Изучить существующие способы оценки точности и скоростирекомендационных алгоритмов;
4. Провести экспериментальное исследование разработанного программного продукта;
5. Описать алгоритм классификации данных.

Объектом исследования являются рекомендательные системы и сервисы в образовательной сфере.

Предметом исследования является математическое и программное обеспечение компьютерных систем для решения задачи рекомендательной системы, методы, фреймворки и алгоритмы построения классификации данных, оценка скорости, эффективности и точности существующих и разработанных алгоритмов.

Основной *гипотезой*, положенной в основу диссертационной работы, является возможность улучшения точности работы алгоритмов рекомендации в процессе практического использования системы. Данный подход, основанный на уже существующих алгоритмах коллаборативной фильтрации, позволяет проанализировать ключевые элементы, обобщить результаты и произвести сравнение релевантности различных моделей классификации, существующих на сегодняшний день.

Личный вклад соискателя

Результаты, полученные в диссертации, получены соискателем лично. Вклад научного руководителя Р. Е. Сердюков заключается в формулировке целей и задач исследования.

Новизна полученных результатов

Разработана система рекомендаций на основе пользовательской коллаборативной фильтрации, позволяющая генерировать персонализированные подборки потенциально интересных образовательных мероприятий с учетом предыдущего образовательного опыта пользователя.

Разработанная система направлена на развитие существующих систем рекомендаций, а также образовательных сервисов. Следует также отметить, что примеры, представленные в качестве примеров, в предлагаемом подходе являются под, а не искусственно созданными данными, что повышает эффективность информационной поддержки при освоении аспектов применения открытой семантической технологии на практике.

Положение, выносимое на защиту

Система рекомендации мероприятий с использованием сингулярного разложения матрицы и градиентного спуска матрицы оценок.

Структура и объём диссертации

Диссертация состоит из введения, четырёх глав, заключения, списка использованных источников и списка публикаций автора. В первой главе рассматривается существующая литература и основные понятия, связанные с проблематикой систем рекомендаций. Во второй главе описывается применение интеллектуального анализа данных в образовательном контексте. Также определяется система оценивания, на основе которой можно классифицировать системы рекомендаций в области сбора данных в сфере образования. Затем представляется современное состояние рекомендаций в образовательной среде и существующие методы моделирования навыков и

успеваемости учащихся. Завершается эта глава критическим взглядом на существующие работы, приближающиеся к рекомендациям в области образования. В третьей главе дается представление о сингулярном разложении матриц и описывается его применение в дополнение к подходу, основанному на использовании рекомендаций, основанных на пользователях. Также будет показано, как использовать модель ASAP-классификаторов, в которой отсутствуют данные об учащемся, в качестве подготовительного шага к рекомендации. В четвертой главе представлен практический пример, содержащий данные учащегося, где будет тестироваться система, а также объяснение того, как будет оцениваться данное предложение рекомендации и соответствующие результаты. Общий объем работы составляет 81 страниц, из которых основного текста 71 страниц, 30 рисунков, 8 таблиц, список использованных источников из 74 наименований на 6 страницах.

КРАТКОЕ СОДЕРЖАНИЕ РАБОТЫ

Во **введении** обоснована актуальность темы диссертационной работы, дана краткая характеристика исследуемых вопросов, обозначены актуальные задачи, решению которых посвящена диссертационная работа.

В **первой** главе описаны основные понятия рекомендательных систем, такие как основные проблемы рекомендаций, влияние социальных проблем, проблема обоснования оценки, оценку релевантности созданных рекомендаций, рассмотрение тип систем фильтрации рекомендаций.

Систему рекомендаций можно описать как набор инструментов и методик, используемых с целью предоставления предложений по различным объектам лицам, не обладающим достаточной компетенцией для оценки потенциально огромного числа имеющихся альтернатив.

Основная проблема рекомендательных систем заключается в том, что рекомендация, как правило, определяется не на всем пространстве оценок, а только на каком-то его подмножестве. Это означает, что полезность элементов, необходимо экстраполировать от известных до неизвестных оценок, либо с помощью эвристики, либо оценочных функций.

Системы рекомендаций предлагают три класса метрик точности рекомендательной системы:

1. *Метрики прогностической точности* измеряют, насколько предсказанные оценки системы рекомендаций близки к истинным оценкам пользователей;

2. *Метрики точности классификации* измеряют, сколько раз система имеет правильное суждение о том, является ли элемент хорошим для пользователя или нет;

3. *Метрики ранговой точности* измеряют способность алгоритма составлять упорядоченный список элементов, который соответствует потребностям пользователя.

Типы систем фильтрации рекомендаций:

1. *Контентная фильтрация;*
2. *Коллаборативная фильтрация;*
3. *Фильтрация, основанная на знаниях;*
4. *Фильтрация на основании сокращения размерности;*
5. *Гибридные фильтрации.*

Во второй главе описаны сбор данных об образовании, система оценок для классификации поведения рекомендательной системы, описание существующих систем и алгоритмов.

Можно использовать систему оценки для правильной классификации поведения рекомендательной системы при сборе образовательных данных в соответствии с фиксированным набором четко определенных параметров анализа:

1. *Рекомендуемые пункты.* Рекомендуемые пункты можно разделить на:
 1. *Учебные материалы;*
 2. *Обратная связь;*
 3. *Группы;*
 4. *Обязательные курсы.*
2. *Пользователи.* С учетом образовательного контекста можно разделить образовательную систему на три различных типа заинтересованных сторон:
 1. *Учащиеся;*
 2. *Обучающие;*
 3. *Разработчики курсов.*
3. *Метод рекомендаций.* Система рекомендаций может быть разбита по категориям в соответствии со своей методикой следующим образом:
 1. *Коллаборативная фильтрация ориентированная на пользователей;*
 2. *Предметная коллаборативная фильтрация;*
 3. *Контентная фильтрация;*
 4. *Фильтрация основанная на знаниях;*
 5. *Фильтрация основанная на сокращении размерности.*

4. Можно классифицировать каждую систему рекомендаций в соответствии с подходами по *добыче данных*, которые они используют для извлечения информации:

1. *Неконтролируемые техники;*
2. *Контролируемые техники.*

5. Выработка рекомендаций в образовательном контексте:

1. *Рекомендационные системы учебных материалов;*
2. *Рекомендационных системы основанные на обратной связи;*
3. *Рекомендационные системы основанные на группах;*
4. *Рекомендационные системы, основанные на курсы;*
5. *Моделирование учащихся и прогнозирование успеваемости.*

Представлены рисунки основных сервисов системы организации мероприятий корпоративного обучения, демонстрирующие базовые возможности работы системы.

В **третьей** главе рассматривается техника CRM и то, как она используется в рекомендации. Представлено подход по рекомендации курсов с использованием оценок учащихся, работа с неполными профилями оценок учащихся до выработки какой-либо рекомендации.

В рамках этой главы дается обоснование использования CRM, применение CRM к рекомендациям, градиентный подход. Описана методология рекомендаций применения коллаборативной фильтрации. Описана подготовка рекомендаций учащимся с неполными профилями оценок с использованием ASAP-классификаторов.

В **четвертой** главе на примере рассмотрено применение разработанного алгоритма коллаборативной фильтрации рекомендаций с использованием ASAP-классификаторов. В качестве примера использованы данные существующих пользователей. В результате обработки которых создаются потенциально релевантные выборки курсов. Также рассматривается использование ASAP-классификаторов, сравниваются различные вариации использования параметров, с которыми создаются рекомендации.

ЗАКЛЮЧЕНИЕ

Выбор курсов для обучения является одним из самых решающих шагов учащегося – сложность выбора оптимального набора курсов, отсутствие необходимой поддержки и консультаций, необходимых для принятия решения. В диссертации описывается использование известных методов рекомендаций курсов обучающимся. Описывается возможность сочетания успешных

рекомендательных методик, такие как сингулярное разложение матрицы и пользовательская коллаборативная фильтрация, чтобы исследовать оценки учащихся, объяснять и предсказывать оценки учащихся по рекомендуемым курсам.

В результате рассмотрения существующей литературы, источников и основных понятий были выделены наиболее актуальные проблемы существующих подходов, определены их преимущества и недостатки.

Была определена система оценивания, на основе которой были классифицированы системы рекомендаций в области сбора данных в сфере образования, оценено современное состояние систем рекомендаций в образовательной среде и существующие методы моделирования навыков и успеваемости учащихся.

Были описаны преимущества и недостатки применения сингулярного разложения матрицы в дополнение к пользовательской коллаборативной фильтрации и модели ASAP-классификаторов.

По итогу рассмотрения практического примера, основанного на реальных данных учащихся, был оценен предложенный подход рекомендаций и показаны соответствующие результаты на основании различных исходных данных. Данный подход основан на сочетании и адаптации двух наиболее успешных методик рекомендаций в образовательном контексте – использование разложения по единичному значению и пользовательская коллаборативная фильтрация. Найденные скрытые особенности используются для прогнозирования оценок курсов для каждого учащегося. Полученные результаты достигают низкого уровня средней арифметической ошибки на представленном примере. Подход позволяет распознать какие курсы являются релевантными для учащегося, так как рекомендации - это те курсы, по которым учащиеся ставят схожие оценки. Предложенный подход предоставляет персональные рекомендации разным учащимся, а не просто рекомендуются курсы, которые имеют лучшие оценки. Использование ASAP-классификаторов позволяет оценить данные о учащемся до процесса рекомендации, если профиль его оценок неполный. Полученные результаты лучше, чем некоторые базовые показатели.

Может оказаться так, что только подгруппа курсов будет доступна в различных доменах или что одни курсы сначала требуют окончания других. Необходимо изучить получение данных из различных источников информации для создания рекомендаций, специфичных для целевой области, с учетом некоторого количества академических факторов, которые помогают объяснить связи в отношении учащихся, поставивших оценки. Следует дальше изучать различные сценарии исследований, когда данные о целевом учащемся ограничены.

Для уточнения рекомендаций необходимо изучить содержательную составляющую, чтобы не только рекомендовать курсы, потенциально соответствующие навыкам учащихся, но и отвечающие их интересам.

Успеваемость учащегося связана с мотивацией, интересами, навыками, рабочим статусом, методами оценки, обучающими, а также любыми другими ресурсами, которые помогают ему во время обучения. Таким образом, имеется много дальнейших различных путей для изучения, будущие исследования могут быть объединены в одну систему рекомендаций курсов. Это лишь первый шаг в исследовании данной научной области и создании более точной рекомендательной системы.

Результаты экспериментальных исследований, приведенные в работе, представляют интерес при проектировании новых и модернизации существующих рекомендательных систем, но необходимо продолжать изучать развитие различных систем рекомендаций различных целевых областей.

Библиотека БГУИР