

СИСТЕМА РЕКОМЕНДАЦИИ В ОБЛАСТИ ОБРАЗОВАНИЯ И КУРСОВ

Халецкий А. М.

Кафедра интеллектуальных информационных технологий, Белорусский государственный университет информатики и радиоэлектроники
Минск, Республика Беларусь
E-mail: unkingip@mail.ru

В данной публикации рассказывается о системе, которая упрощает выбор курсов учащихся. Объясняется метод применения образовательного интеллектуального анализа, сбор данных об образовании, который начинается с предварительной обработки необработанных данных, которые разделены на категории. Описан сам метод сбора данных для извлечения значимой информации, которая лежит в основе рекомендации, а так же оценка созданной системы.

ВВЕДЕНИЕ

Успех data mining при применении к нескольким видам бизнеса имеет свои доказательства, благодаря своей способности извлекать, открывать и представлять в интерпретируемой форме интересные и новые закономерности из больших объемов информации. Применение этих методов в отношении образовательного контекста в последнее время и обусловлено последними достижениями в области образовательных технологий: в настоящее время можно извлекать большие массивы данных из многих источников образовательной информации. С развитием Интернета в настоящее время возможно распространение образования через веб-сервисы: в настоящее время можно использовать системы электронного обучения, которые обеспечивают онлайн-обучение, системы управления обучением, которые облегчают общение между преподавателями и учащимися, а также существуют интеллектуальные системы обучения и образовательные гипермедийные системы, которые направлены на адаптацию к конкретным потребностям бизнеса, предоставляя наилучшие ресурсы для улучшения знаний сотрудников.

I. СБОР ДАННЫХ ОБ ОБРАЗОВАНИИ

Educational data mining начинается с предварительной обработки необработанных данных (собранных из многих источников), применяя наиболее подходящие методы извлечения данных, а затем осуществляет их дальнейшую обработку, в результате которой получается структурированная информация, способная лучше понять процесс обучения в области образования. Можно разделить EDM на одиннадцать основных категорий следующим образом:

1. Анализ и оценка;
2. Обратная связь для преподавателей, оказывающих поддержку;
3. Рекомендации для обучающихся;
4. Прогнозирование оценок обучающихся;
5. Моделирование обучающихся;

6. Обнаружение нежелательного поведения учащихся;
7. Группировка обучающихся;
8. Социальный сетевой анализ;
9. Разработка планов;
10. Построение учебных программ;
11. Составление расписания.

В связи с успешной применимостью как сбора образовательных данных, так и обучения с использованием веб-технологий, возрос интерес к применению систем рекомендаций в этой конкретной области. Таким образом, в условиях образования уже существует большое число систем рекомендаций, которые отличаются от других рекомендуемых лиц (например, рекомендуемых продукты) по различным целям, которые они пытаются достичь. Тем не менее, необходимо выявить особенности систем рекомендаций, ориентированных на образование, чтобы обеспечить развитие и оценку.

II. МЕТОД СБОРА ДАННЫХ

подавляющее большинство систем DM будет использовать один (или более) из известных различных подходов. Можно классифицировать каждую систему рекомендаций в соответствии с подходами по добыче данных, которые они используют для извлечения значимой информации, которая лежит в основе их рекомендаций. Можем разделить ее на две основные категории:

1. Неконтролируемые техники. Эта категория определяется набором методик, которые не используют никаких внешних знаний для руководства процессом сбора знаний при поиске любой закономерности в данных. Другими словами, в процессе обучения не даются правильные результаты. Наиболее популярными примерами этого класса методик являются правила кластеризации и ассоциации;
2. Контролируемые техники. В отличие от категории, описанной выше, этот набор методик может использовать знания о данных для создания точной модели, которая может корректно классифицировать экзем-

плярные области в процессе анализа. Эти знания представляют собой не более чем метку или класс для каждого известного экземпляра. Таким образом, после обучения созданная модель сможет обобщать классификацию для каждого нового экземпляра. В этой группе выделяются несколько методов, таких как деревья принятия решений, байесовские сети и нейронные сети;

III. МЕТОДОЛОГИЯ И РЕКОМЕНДАЦИЯ КУРСОВ

Предлагается система, которая может рекомендовать набор курсов, которые максимально повышают вероятность хорошей академической успеваемости студентов с учетом их предыдущих оценок. Данное предложение исследует исторические оценки студентов, используя сингулярные значения декомпозиции вершины совместного подхода к фильтрации, чтобы захватить скрытые латентные особенности, которые могут объяснить оценки, полученные обучающимися. Был выбран принцип SVD (сингулярное разложение матриц), его можно использовать для обнаружения скрытых скрытых функций, которые могут объяснить результаты, полученные учащимися. Также рассматривается проблема неполноты профилей студентов на момент рекомендации, используя классификаторы ASAP. Классификаторы ASAP по-другому представляют те данные, которые обычно используются в классификации. Пропустив изучение того, как можно обрабатывать данные для достижения более высокой точности классификации что является основной проблемой, изучаемой в этой области эта методика подходит к проблеме классификации частично наблюдаемых экземпляров, когда классификаторы могут быть обучены полностью наблюдаемым экземплярам.

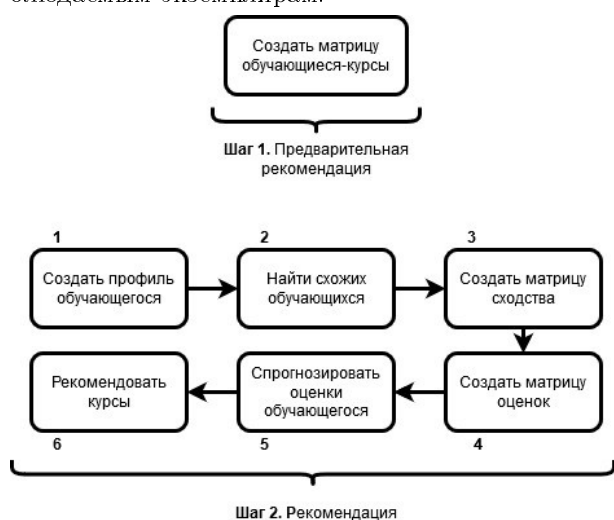


Рис. 1 – Основные этапы процесса подготовки рекомендаций

IV. ОЦЕНКА СОЗДАННОЙ СИСТЕМЫ

Больше внимания уделяется оценке точности прогнозирования оценок. Таким образом, если достигается низкая погрешность в предсказаниях, то можно с уверенностью рекомендовать курсы. Учащиеся могут выбрать более 200 различных курсов. Для всех курсов можно выставить оценку от 1 до 5. В проводимых исследованиях систем рекомендаций используется несколько видов мер для оценки успешности рекомендаций. Однако правильный способ оценки каждой системы в значительной степени зависит от целей и сферы применения системы. Рекомендательный процесс можно разделить на две различные проблемы: предсказание оценок и решение о том, какие курсы рекомендовать с учетом вычисленных прогнозов. Общее качество рекомендаций, независимо от метода, используемого для их получения, в значительной степени зависит от качества прогнозов. То есть, чем меньше ошибка на рекомендуемых курсах, тем лучше будет качество выдаваемых рекомендаций.

V. ЗАКЛЮЧЕНИЕ

Выбор курсов для обучения является одним из самых решающих шагов учащегося. Студентам все еще сложно выбрать для себя оптимальный набор курсов, и в большинстве случаев у них нет необходимой поддержки и консультаций, необходимых для принятия решения. В данной публикации мы показываем, что можем использовать известные методы рекомендаций, чтобы рекомендовать студентам курсы, которые не только интересны им, но и наиболее схожи с их навыками.

VI. СПИСОК ЛИТЕРАТУРЫ

1. Messina, E. R. Measuring the Performance and Intelligence of Systems / E. R. Messina, A. M. Meystel // N. Engl. . – 2001. – Vol. 342, № 982. – P. 164.
2. Системы искусственного интеллекта. Учебное пособие. / М. В. Бураков . – СПб.: Инкарт, 2010. – 440 с.
3. Флах, П. Машинное обучение. Наука и искусство построения алгоритмов, которые извлекают знания из данных / пер. с англ.// Вест. аритмол. – 2015. – С. 400.
4. Самообучающиеся системы. / С. И. Николенко, А. Л. Тулупьев // – МЦНМО, 2009. – 288 с.
5. Vialardi, C. Recommendation in higher education using data mining techniques / C. Vialardi, J. Bravo // N. Engl. – May 2009., – P. 190–199/
6. Gutierrez, F. Providing grades and feedback for student summaries by ontology-based information extraction / F. Gutierrez, D. Dou // N. Engl. – September 2012., – P. 1722–1726/