

АНАЛИЗ ОБРАЗОВАТЕЛЬНЫХ ДАННЫХ В ВЫСШИХ УЧЕБНЫХ ЗАВЕДЕНИЯХ

Скиба И. Г., Тарасюк И. С., Нестеренков С. Н.

Отдел информационных технологий, кафедра программного обеспечения информационных технологий,
Белорусский государственный университет информатики и радиоэлектроники
Минск, Республика Беларусь

E-mail: {i.skiba, i.tarasiuk, s.nesterenkov}@bsuir.by

Основной инструмент для поддержки принятия управленческих решений, направленных на повышение качества обучения в высшем учебном заведении, это технология анализа образовательных данных. Такой анализ пробует извлечь закономерности из данных, производимых в процессе обучения, таких как, например, успеваемость, посещаемость или данных из систем дистанционного образования. Он включает в себя алгоритмы машинного обучения и статистические методы для получения информации об академической успеваемости студентов и прогнозирования дальнейшей динамики на протяжении всего учебного процесса.

ВВЕДЕНИЕ

Развитие систем электронного обучения и внедрения информационных технологий в образовательный процесс привело к возникновению больших объемов данных. Главная цель анализа образовательных данных (АОД) – это повышение качества образования и подготовка качественных специалистов [1]. В процессе анализа и применения образовательных данных выделяют четыре вида пользователей, вовлеченных в процесс анализа образовательных данных:

- Обучающиеся, которые заинтересованы в повышении своего уровня успеваемости, осознанном выборе профессиональной сферы.
- Преподаватели, заинтересованные в разбиении обучающихся на группы, подбора самого оптимального метода обучения, а также в улучшении учебного курса.
- Исследователи, заинтересованные в объективной оценке эффективности учебного процесса, разработке современных методов анализа образовательных данных.
- Администрация вуза, заинтересованная в повышении качества управленческих решений для влияния на учебный процесс, а также наблюдении динамики после проведения соответствующий мероприятий.

I. Цели

Основной целью методов интеллектуального анализа данных является преобразование данных в значимую информацию о процессе обучения и поиск закономерностей для лучшего принятия решений относительно обучения. Процесс применения АОД обычно состоит из пяти следующих этапов (см. рис. 1):

1. Предварительная обработка данных.
2. Выявление закономерностей в данных тем или иным методом data mining.
3. Проверка (валидация) обнаруженных закономерностей.

4. Применение найденных закономерностей для прогнозирования будущих событий в обучающей среде.
5. Использование построенных прогнозов при поддержке принятия решений и выработки образовательной политики.



Рис. 1 – Процесс применения методов АОД

II. МЕТОДЫ ТРАДИЦИОННОГО DATA MINING

Для обеспечения высокого качества анализа больших массивов образовательных данных нужно применять соответствующие аналитические инструменты, в том числе эффективные алгоритмы, способные выявить основные зако-

номерности в созданных массивах. В АОД широко используются методы традиционного Data mining, основными из которых являются: классификация, кластеризация, поиск связующих правил, поиск последовательных шаблонов, а также интеллектуальный анализ текстов. Все эти методы можно разделить на три большие группы.

III. ПРОГНОЗИРОВАНИЕ

Цель прогнозирования заключается в том, чтобы разработать модель, которая будет предсказывать значение интересующей величины по значениям, которые приобретают независимые переменные. Прогнозирование успеваемости — одна из старейших и наиболее популярных задач АОД, для решения которой используются самый широкий спектр методов: нейронные сети; байесовские сети; системы, основанные на правилах; методы регрессионного и корреляционного анализа. Методы регрессионного и корреляционного анализа широко применяются для выявления и описания зависимостей между случайными величинами по экспериментальным данным и основываются на теории вероятности и математической статистике [2].

IV. ОБНАРУЖЕНИЕ СТРУКТУРЫ

Алгоритмы обнаружения структуры пытаются выявить в данных структуру, без использования каких-либо заведомо известных представлений о ней. Наиболее известной группой подобных алгоритмов являются алгоритмы кластеризации. Кластеризация является логическим продолжением идеи классификации. Особенность ее состоит в том, что классы объектов изначально не предопределены, тогда как при классификации прежде чем перейти к созданию модели необходимо разметить некоторое подмножество данных. Результатом кластеризации является разбиение множества объектов на группы (кластеры) близких друг к другу по ряду признаков или свойств [3].

V. ВЫЯВЛЕНИЕ ВЗАИМОСВЯЗЕЙ

Цель выявления взаимоотношений состоит в том, чтобы установить некоторые взаимосвязи между переменными в множестве данных с большим числом переменных. Например, можно попытаться определить, какие переменные сильнее всего связаны с интересующей или в какой из пар связанных переменных связь сильнее, чем в других. Выявление взаимоотношений в АОД чаще всего используется в форме поиска связующих правил или поиска последовательных шаблонов [4].

VI. ИСТОЧНИКИ ДАННЫХ

Для исследования данных можно использовать данные внутренних информационных систем и систем дистанционного обучения. В част-

ности, можно использовать такие данные внутренних систем, как посещаемость, успеваемость, данные о распределении и другие доступные данные, позволяющие исследовать отношение студента к образовательному процессу. Например, в Белорусском государственном университете информатики и радиоэлектроники на данный момент активно используется интегрированная информационная система «БГУИР: Университет», в которой есть возможность выставления оценок и ведение учета посещаемости[5].

VII. ЗАКЛЮЧЕНИЕ

Методы обработки больших данных используются для управления высшим учебным заведением [6]. Образовательная среда, которая использует методы АОД, позволит адаптировать изучаемый курс для студентов под их возможности. Преподавателям АОД дает возможность получить информацию, которая позволит улучшить содержание курса, настроить процесс обучения для каждого учащегося. Так, на основе данных о качестве и количестве ошибок, которые совершают студенты, выявляют причины возникновения данных ошибок и вносят изменения в курс. Это позволяет администрации оценить эффективность результатов обучения, а также разработать комплекс решений для повышения уровня образовательного процесса за счет использования имеющихся ресурсов.

VIII. СПИСОК ЛИТЕРАТУРЫ

1. Трубоч, Г. Г. Использование анализа данных в образовании / Г. Г. Трубоч // 77-я научная конференция студентов и аспирантов Белорусского государственного университета. — Минск, 2020. — С. 119-122.
2. Дополнительные главы математики. Статистический анализ / Н. И. Кривцова, О. Е. Мойзес // Томск: Изд-во Томского политехнического университета. — 2015.
3. Баранников, К. А. Методология анализа больших данных в образовании (системно-методологический подход, основанный на анализе образовательных данных, поиска стратегии принятия управленческих и организационно-педагогических решений в образовании) / К. А. Баранников, С. М. Лесин // Народное образование. — 2020. — №. 2 (1479). — С. 81-90.
4. Белоношко П. П. Анализ образовательных данных: направления и перспективы применения / П. П. Белоношко, А. П. Карпенко, Д. А. Храмов // Вестник евразийской науки. — 2017. — Т. 9. — №. 4 (41). — С. 57.
5. Нестеренков, С. Н. Интегрированная информационная система как средство автоматизации управления образовательным процессом в учреждениях высшего образования / С. Н. Нестеренков, Т. А. Рак, О. О. Шатилова // Информационные технологии и системы 2017 (ИТС 2017) = Information Technologies and Systems 2017 (ITS 2017) : материалы междунар. науч. конф. (Республика Беларусь, Минск, 25 октября 2017 года) — Минск : БГУИР, 2017. — С. 212.
6. Нестеренков, С. Н. Применение больших данных в электронном образовании. / С. Н. Нестеренков, М. И. Макаров, Н. В. Ющенко, А. Д. Радкевич // Пятая Международная научно-практическая конференция «BIG DATA and Advanced Analytics. BIG DATA и анализ высокого уровня», — Минск, 2019.