

Управление учебным процессом в высшей школе при смешанном и дистанционном формате обучения

А.А. Ивлиева, А.М. Бершадский, А.Ф. Валько

Пензенский государственный университет

Аннотация: В работе актуализируется стратегический подход к поддержанию качества успеваемости студентов. В статье «Большие данные» представляются как подход к организации данных, получаемых в ходе учебного процесса. Интеллектуальный анализ рассматривается как инструмент анализа Больших данных. Также авторами рассмотрены подходы, повышающие эффективность обучения, которые позволяют прогнозировать успеваемость учащихся. Смешанное обучение, как интеграция в учебный процесс лучших возможностей очного и дистанционного обучения, представлено основой для последующего сбора Больших данных для анализа. Рассмотрено применение методов интеллектуального анализа данных для разработки моделей прогнозирования поведения и успеваемости учащихся. Авторами предложено моделирование учащихся – как ключевая концепция интеллектуального анализа данных в образовании, которая относится к качественному представлению поведения учащихся. Описаны алгоритмы интеллектуального анализа данных, классификация и кластеризация представлены как наиболее частые задачи. Алгоритм CART дерева решений, представлен как один из эффективных методов интеллектуального анализа данных, необходимых для прогнозирования успеваемости учащихся на основе онлайн-активности, которая хранится в файле журнала Moodle LMS. В завершение сделан вывод, что личный контакт опытного педагога со студентом в учебном процессе невозможно заменить на инструменты, предоставляемые Moodle и аналогичными системами, но в связи с ростом угроз, вызванных пандемией, актуальность дистанционного образования приобретает все большее значение.

Ключевые слова: качество успеваемости студентов, большие данные, интеллектуальный анализ, моделирование учащихся, Алгоритм CART дерева решений, Moodle LMS, дистанционное образование.

Успеваемость студентов является главным критерием оценки деятельности любого высшего учебного заведения, как в России, так и за рубежом.

Как правило, в высших учебных заведениях, именно итоговые оценки используются для определения общего уровня успеваемости студентов. Оценки, выставляемые студентам, основываются на структуре курса, количестве набранных баллов за весь курс и итоговой оценке за экзамен. Оценка, как стимулирующий фактор, важна для поддержания успеваемости учащихся и в конечном итоге определяется эффективностью учебного

процесса [1]. Анализируя успеваемость студентов, можно оптимизировать стратегическую программу на весь период обучения в ВУЗе.

В настоящее время существует множество методик оценки успеваемости студентов. Безусловно, размерность данных, получаемых при анализе успеваемости студентов в текущем семестре и предыдущих учебных периодах, может быть отнесена к так называемым Большим данным. Инструментом для анализа Больших данных служит интеллектуальный анализ данных, который является одним из самых популярных методов анализа успеваемости учащихся в высшей школе. В последнее время интеллектуальный анализ данных широко применяется в сфере образования многих стран [2]. Интеллектуальный анализ данных - это процесс, используемый для извлечения полезной информации и шаблонов из огромной базы данных [3]. Извлекаемая из баз данных информация и шаблоны могут быть использованы как для прогнозирования успеваемости учащихся, так и для осуществления управляющих воздействий на сам учебный процесс. В результате это может помочь преподавательскому составу и учебному отделу обеспечить более эффективный подход к обучению студентов. Студенты, в свою очередь, могут поднять свой уровень знаний и оптимизировать свою учебную деятельность. Все это позволит администрации вуза улучшить показатели по успеваемости, оптимизировать свои учебные планы и осуществлять эффективные управленческие воздействия на образовательную среду. Таким образом, применение методов интеллектуального анализа данных может быть сфокусировано на определенных требованиях к различным объектам образовательной среды.

Прогнозирование успеваемости учащихся с целью предотвращения негативных сценариев в плане низкого уровня успеваемости как отдельно взятого студента, группы, факультета, так и всего ВУЗа в целом является предметом интенсивных исследований в области интеллектуального анализа

данных в образовании. Средний балл или оценки по заданиям, тестам, лабораторным и курсовым работам, посещаемости, а также данные, определяющие специфические характеристики учащихся, такие, как пол, возраст и семейное положение и семейные обстоятельства, личное поведение учащихся, убеждения, мотивация и стратегии обучения являются атрибутами, которые часто используются исследователями при прогнозировании успеваемости учащихся [4].

Одним из подходов, повышающих эффективность обучения, является смешанное обучение, которое обычно определяется как интеграция в учебный процесс лучших возможностей очного и дистанционного обучения [5]. Традиционные методы обучения в аудитории могут сочетаться с мультимедийным контентом и доступом в Интернет, это создает гибридную методологию обучения, которая поддерживает взаимодействия в образовательной среде в любое время и в любом месте. Системы управления обучением, такие, как Moodle или Blackboard за рубежом, используются образовательными учреждениями для повышения эффективности таких аспектов обучения, как: создание и редактирование учебного контента, совместная работа и общение, мониторинг и администрирование обучения. Встроенные инструменты образовательной среды содержат: редакторы HTML кода, инструменты веб-публикации, средства для подготовки заметок и заданий, записи и загрузки видео-лекций, размещения подсказок и полезных URL-ссылок, создания презентаций, организации проектных групп и общения с каждым студентом, который нуждается в их помощи.

В курсах смешанного обучения студенты используют онлайн-материалы, чтобы заранее организовать свою траекторию обучения, в то время как преподаватели используют учебное время для изучения причин и решения проблем, которые могут возникать у отдельных студентов и групп в целом, а также заниматься подготовкой и организацией практических

занятий. Роль преподавателя, в настоящее время, смещается с роли поставщика знаний к роли некоего гида, помогающего студентам осуществлять постановку учебных целей и задач, организуя их обучение, корректируя направление их обучения при овладении содержанием учебного курса [6]. Существует множество подходов к смешанному обучению, это обеспечивает большую гибкость при определении приоритетов физических и сетевых элементов учебной среды в соответствии с различными стилями обучения и базой знаний. Комбинируя различные инструменты системы управления обучением, преподаватели могут в большей или меньшей степени вмешиваться в процесс обучения, сочетая различные формы обучения в зависимости от скорости усвоения материала студентами.

Опираясь на информацию, полученную из интерактивных мероприятий, таких, как совместная работа студентов, преподаватели могут на раннем этапе выявить проблемные ситуации и разработать целенаправленные воздействия на процесс обучения. Четкие инструкции и более глубокое понимание концепций могут быть предоставлены либо с использованием учебных часов на занятиях, либо с использованием функций онлайн-обучения [7]. Оба метода обучения могут использоваться для поддержки, вовлечения и эффективной мотивации студентов как с низким уровнем успеваемости и понимания материала, так и очень способных студентов [7]. Как показали исследования, учащиеся курсов смешанного обучения удовлетворены и мотивированы удобством, независимостью и персонализацией, которые обеспечивает онлайн-обучение. По мере того, как учебные учреждения все чаще внедряют смешанное обучение, методы интеллектуального анализа данных становятся мощным средством для анализа возросшего количества данных об успеваемости учащихся. Это позволяет сократить время, необходимое для диагностики проблемных мест в образовательном процессе по каждому предмету.

В связи с постоянным ростом количества электронных учебных материалов в образовании, методы интеллектуального анализа данных все чаще применяются в образовании для разработки моделей прогнозирования поведения и успеваемости учащихся [8]. В связи с необходимостью извлечения полезной информации из больших наборов данных и выявления скрытых закономерностей и тенденций, интеллектуальный анализ данных - также называемый обнаружением знаний в базах данных (KDD), использует различные методы статистического и машинного обучения для анализа данных и облегчения принятия решений. Следуя этому утверждению, веб-сайт сообщества Educational Data Mining, educationdatamining.org, определяет интеллектуальный анализ данных в образовании следующим образом: «Анализ данных в образовании - это новое направление, занимающееся разработкой методов исследования уникальных типов данных, которые поступают из образовательных учреждений. Также данное направление изучает то, как использовать вышеупомянутые методы, чтобы лучше понимать студентов и условия, в которых они учатся» [9]. Классификация, кластеризация, интеллектуальный анализ ассоциативных правил, регрессия и обнаружение с помощью моделей - это обычные методы интеллектуального анализа данных, используемые большинством исследователей при анализе собранных данных. Моделирование учебного поведения студентов, прогнозирование успеваемости, прогнозирование отсева и удержания, идентификация учащихся, испытывающих трудности, являются основными целями интеллектуального анализа данных в образовательной среде [10].

Моделирование учащихся - ключевая концепция интеллектуального анализа данных в образовании, относящаяся к качественному представлению поведения учащихся, которое, в свою очередь, может использоваться при принятии учебных решений [11]. Модели могут отражать содержание знаний, навыки решения проблем, стили обучения, эмоции и отношения

учащихся, опыт или уверенность в себе, ошибки и заблуждения, результаты действий или промежуточные результаты [12]. В смешанной среде обучения необработанные данные из тщательно отобранных онлайн-взаимодействий студентов сначала проверяются на реальных управленческих действиях, а затем используются для создания прогнозной модели. Различные комбинации показателей, таких, как IP-адреса, просмотры страниц контента, количество пройденных тестов, время выполнения задачи, средняя продолжительность сеанса, сообщения, которыми обмениваются, просматриваемые сообщения, комментарии, а также вклад в создание контента могут использоваться для изучения формируемого уровня успеваемости.

В структуре процесса интеллектуального анализа присутствуют этапы сбора и предварительная обработка данных. Процесс сбора больших объемов данных является важнейшим этапом интеллектуального анализа, потому что правильный выбор атрибутов определяет в целом концепцию дальнейшего анализа. После определения целей и требований интеллектуального анализа некоторые данные могут быть удалены или могут быть добавлены новые вычисленные атрибуты для обеспечения значимых результатов интеллектуального анализа данных. Такое извлечение признаков очень часто требует преобразования данных, закодированных в различных сложных форматах базы данных, в форматы, более подходящие для алгоритмов интеллектуального анализа данных. Предварительная обработка данных также включает этап очистки данных, на котором выявленные ошибки проверяются и исправляются, а нулевые или отсутствующие значения обрабатываются соответствующим образом. Данные извлекаются из разных источников и объединяются в единый многомерный формат для обработки. На этапе обработки данных точные характеристики из набора структурированных данных анализируются с помощью математических

алгоритмов, которые могут быть выбраны на основе типов данных их взаимосвязей и тенденций. Именно их модель стремится уловить интеллектуальный анализ данных [13].

Алгоритмы интеллектуального анализа данных используют четыре класса отношений между данными: классы, кластеры, ассоциации и последовательные шаблоны. Выбор используемого алгоритма зависит от таких факторов, как производительность, которая может быть достигнута в конкретной области приложения в зависимости от уровня точности результатов или понятности модели. Классификация и кластеризация являются наиболее частыми задачами интеллектуального анализа данных на данных больших размерностей [14].

Классификация - это контролируемый процесс обучения, в котором алгоритм учится на предоставленных примерах и создает модель, которая отображает новые, невидимые экземпляры в предопределенных классах. Предварительные знания включены в каждый пример в виде пары из входного вектора атрибутов (характеристик данных) и значения (метки), указывающей класс, к которому принадлежит этот пример. Ближайшие соседи, деревья решений, классификаторы на основе правил, искусственные нейронные сети, машины опорных векторов и байесовский алгоритм - это базовые классификаторы, используемые для классификации новых данных.

Кластеризация, напротив, представляет собой процесс обучения без учителя, который назначает записи данных кластерам, уже содержащим элементы [15]. Эти элементы больше похожи на новую запись, чем на элементы, содержащиеся в других кластерах. Сходство между элементами измеряется функциями расстояния, такими, как евклидово или косинусное расстояние. Расстояние определяет качество кластеризации: кластеры высокого качества имеют высокое внутрикластерное сходство и низкое межкластерное сходство. Иерархическая кластеризация, кластеризация К-

средних и кластеризация на основе плотности являются основным подходом к кластеризации данных в исследованиях [16].

Одним из наиболее ярких примеров алгоритмов интеллектуального анализа данных является CART (classification and regression trees) – аббревиатура, обозначающая методы классификации и регрессии с использованием дерева решений [17]. Это методика обучения, основанная на деревьях решений, которая возвращает классификационные или регрессионные деревья. CART используется именно для задач классификации, о которых говорилось выше. Одним из преимуществ данного алгоритма является легкость в интерпретации и понимании [18].

Алгоритм CART дерева решений, как один из эффективных методов интеллектуального анализа данных, можно использовать для прогнозирования успеваемости учащихся на основе онлайн-активности, которая хранится в файле журнала Moodle LMS. Журналы Moodle включают действия, выполненные преподавателем, такие, как оценка студента или загрузка файла, связанного с лекцией, а также события, связанные с действиями студентов, такие, как публикация в блоге группы или отправка задания [18]. Необходимые данные можно извлечь из системного файла журнала и преобразовать в таблицу Excel, с помощью макросов VBA использовать на этапе предварительной обработки для расчета количества просмотренных файлов, вкладов в контент, усилий по выполнению тестов и обмена сообщениями для каждого отдельного студента. После предварительной обработки данных файл Excel, включающий вышеупомянутые атрибуты и идентификаторы учащихся, можно объединить с другим файлом Excel, содержащим уникальные идентификаторы учащихся, соответствующие оценки и прочие активности.

Безусловно, личный контакт опытного педагога со студентом в учебном процессе невозможно заменить на инструменты, предоставляемые

Moodle и аналогичными системами, но в связи с ростом угроз, вызванных пандемией, актуальность дистанционного образования приобретает все большее значение. Получение инструментов, позволяющих проводить качественный анализ учебного процесса и эффективно управлять удаленным обучением, оказывая управленческие воздействия как в целом на учебные потоки, так и на отдельных студентов, является перспективным направлением в области информационных технологий в образовании.

Литература

1. Бортник Б.И., Стожко Н.Ю., Судакова Н.П., Оценка компетенций: формализация и формалистика // Электронный научный журнал «Современные проблемы науки и образования». 2017. № 4. URL: science-education.ru/ru/article/view?id=26693.

2. Пиотровская К.Р., Тербушева Е.А., Интеллектуальный анализ данных и развитие научно-исследовательских компетенций бакалавров // Электронный научный журнал «Современные проблемы науки и образования». 2017. № 1. URL: science-education.ru/ru/article/view?id=26084.

3. Певченко, С. С. Методы интеллектуального анализа данных // Молодой ученый. 2015. № 13 (93). С. 167-169. URL: moluch.ru/archive/93/20875.

4. Долотова Р.Г., Долотов А.Е., Применение информационных образовательных ресурсов для повышения эффективности самостоятельной работы студентов // Электронный научный журнал «Современные проблемы науки и образования». 2016. № 2. URL: science-education.ru/ru/article/view?id=24382.

5. Абрамян Г.В., Катасонова Г.Р., Особенности организации дистанционного образования в вузах в условиях самоизоляции граждан при вирусной пандемии // Электронный научный журнал «Современные

проблемы науки и образования». 2020. № 3. URL: science-education.ru/ru/article/view?id=29830.

6. Листков В.Ю., Белоусова Э.В. Перспективы развития образовательных технологий. Технологии в образовании–2018: материалы Международной научно-методической конференции. 23–30 апреля 2018 г. / ЧОУ ВО Центросоюза РФ «СибУПК». – Новосибирск, 2018. – 21-25 с.

7. Терещенко В.А., Качество вузовского образования и некоторые пути его повышения. Технологии в образовании–2018: материалы Международной научно-методической конференции. 23–30 апреля 2018 г. / ЧОУ ВО Центросоюза РФ «СибУПК». – Новосибирск, 2018. – 156-167 с.

8. Белоношко П.П., Карпенко А.П., Храмов Д.А., Анализ образовательных данных: направления и перспективы применения // Интернет-журнал «Науковедение». 2017. Том 9, №4. URL: naukovedenie.ru/PDF/15TVN417.pdf.

9. Educational Data Mining URL: educationdatamining.org (дата обращения: 05.11.2020).

10. Kabakchieva, D., Predicting student performance by using data mining methods for classification // Cybernetics and Information Technologies. 2013. Vol. 13 (1). pp. 61-72.

11. Kisimov V., Kabakchieva, D., Naydenov, A., Stefanova, K., Agile elastic desktop corporate architecture for big data // Cybernetics and Information Technologies. 2020. Vol. 20(3). pp. 15-31.

12. Troussas, C., Chrysafiadi, K., Virvou, M., Personalized tutoring through a stereotype student model incorporating a hybrid learning style instrument // Education and Information Technologies, 2020. URL: link.springer.com/article/10.1007/s10639-020-10366-2 (дата обращения: 05.11.2020)

13. Tzaban, H., Guy, I., Greenstein-Messica, A., Rokach, L., Shapira, B. SIGIR Product Bundle Identification using Semi-Supervised Learning 2020 - Proceedings of the 43rd International ACM SIGIR Conference on Research and Development in Information Retrieval, 2020, pp. 791-800.

14. Mitchell, T. M. (1997). Machine Learning. New York: WCB/McGrawHill.

15. Хабр: Введение в машинное обучение. URL: habr.com/ru/post/448892/ (дата обращения: 05.11.2020).

16. Xu D., Tian Y. A comprehensive survey of clustering algorithms // Ann. Data Sci. 2015. V. 2, № 2. Pp. 165–193.

17. StatSoft: Деревья классификации URL: statsoft.ru/home/textbook/modules/stclatre.html (дата обращения: 12.05.21)

18. Дубив Н.В., Соловьева И.В., Обратная связь «преподаватель - студент» на платформе moodle при обучении иностранному (немецкому) языку // Электронный научный журнал «Современные проблемы науки и образования». 2019. № 2. URL: science-education.ru/ru/article/view?id=28715.

References

1. Bortnik B.I., Stozhko N.YU., Sudakova N.P., Elektronnyj nauchnyj zhurnal «Sovremennye problemy nauki i obrazovaniya», 2017. №4. URL: science-education.ru/ru/article/view?id=26693.

2. Piotrovskaya K.R., Terbusheva E.A., Elektronnyj nauchnyj zhurnal «Sovremennye problemy nauki i obrazovaniya», 2017. №1. URL: science-education.ru/ru/article/view?id=26084.

3. Pevchenko, S. S., Molodoj uchenyj 2015. №13 (93). pp. 167-169. URL: moluch.ru/archive/93/20875/.

4. Dolotova R.G., Dolotov A.E., Elektronnyj nauchnyj zhurnal «Sovremennye problemy nauki i obrazovaniya», 2016. №2. URL: science-education.ru/ru/article/view?id=24382.

5. Abramyan G.V., Katasonova G.R., Elektronnyj nauchnyj zhurnal «Sovremennye problemy nauki i obrazovaniya», 2020. №3. URL: scienceeducation.ru/ru/article/view?id=29830.
 6. Listkov V.YU., Belousova E.V. Mezhdunarodnaya nauchno-metodicheskaya konferenciya: trudy. Novosibirsk, 2018, pp. 21-25.
 7. Tereshchenko V.A. Mezhdunarodnaya nauchno-metodicheskaya konferenciya: Trudy. Novosibirsk, 2018, pp. 156-167.
 8. Belonozhko P.P., Karpenko A.P., Internet-zhurnal «Naukovedenie». 2017. Volume 9, №4. URL: naukovedenie.ru/PDF/15TVN417.pdf.
 9. Educational Data Mining. URL: educationdatamining.org (date accessed: 05.11.2020)
 10. Kabakchieva, D., Cybernetics and Information Technologies. 2013. Vol. 13 (1). pp. 61-72.
 11. Kisimov V., Kabakchieva, D., Naydenov, A., Stefanova K. Cybernetics and Information Technologies. 2020. Vol. 20(3). pp. 15-31
 12. Troussas, C., Chrysafiadi, K., Virvou, M., Education and Information Technologies, 2020. URL: link.springer.com/article/10.1007/s10639-020-10366-2 (date accessed: 05.11.2020)
 13. Product Bundle Identification using Semi-Supervised Learning Tzaban, H., Guy, I., Greenstein-Messica, A., Rokach, L., Shapira, B. SIGIR 2020 - Proceedings of the 43rd International ACM SIGIR Conference on Research and Development in Information Retrieval, 2020, pp. 791-800 .
 14. Mitchell, T. M. 1997. Machine Learning. New York: WCB/McGrawHill.
 15. Habr: Vvedenie v mashinnoe obuchenie. [Habr: Introduction to Machine Learning]. URL: habr.com/ru/post/448892/ (date accessed: 05.11.2021)
 16. Xu D., Tian Y. Ann. Data Sci. 2015. V. 2, No 2. pp. 165–193.
-



17. StatSoft: Derev'ya klassifikacii. [StatSoft: Classification Trees]. URL: statsoft.ru/home/textbook/modules/stclatre.html (date accessed: 12.05.21).

18. Dubiv N.V., Solov'eva I.V., Elektronnyj nauchnyj zhurnal «Sovremennye problemy nauki i obrazovaniya», 2019. №2. URL: science-education.ru/ru/article/view?id=28715.