

УДК 377.091

## К ВОПРОСАМ ПРИМЕНЕНИЯ МЕТОДОВ ИНТЕЛЛЕКТУАЛЬНОГО АНАЛИЗА ДАННЫХ В ЭЛЕКТРОННОМ ОБРАЗОВАНИИ

Якубжанова Д.К.<sup>1</sup>, Абдукаримова Х.<sup>1</sup>

<sup>1</sup> Самаркандский филиал Ташкентского университета информационных технологий имени Мухаммада ал-Хорезми, Самарканд, Узбекистан

yakubjanova55575@gmail.com

**Аннотация.** В статье обосновывается актуальность использования методов интеллектуального анализа данных для поддержки управления в электронном образовании процессе. Подробно излагаются основные методы интеллектуального анализа данных в электронном образовании, а также представлен обзор основных научных работ, посвященных данному направлению. Приведенные методы могут быть полезными в системах поддержки принятия решений на всех уровнях управления образовательным процессом.

**Ключевые слова:** интеллектуальный анализ данных, образовательный процесс, электронное образование, метод, прогнозирование, анализ образовательных данных.

### I. ВВЕДЕНИЕ

В интеллектуальном анализе данных алгоритм — это набор эвристики и вычислений, который создает на основе данных модель. Чтобы создать модель, алгоритм сначала анализирует предоставленные данные, осуществляя поиск определенных закономерностей и тенденций. Алгоритм применяет результаты этого анализа ко множеству итераций, чтобы подобрать оптимальные параметры для создания модели интеллектуального анализа данных. Затем эти параметры применяются ко всему набору данных, чтобы выявить пригодные к использованию закономерности и получить подробную статистику.

К основным задачам интеллектуального анализа данных относятся задача классификации, задача регрессии, задача прогнозирования, задачи кластеризации, задачи определения взаимосвязей. Одно из важнейших назначений методов интеллектуального анализа данных состоит в наглядном представлении результатов вычислений, что позволяет использовать инструментарий интеллектуального анализа данных людьми, не имеющими специальной математической подготовки.

Существует множество методов и алгоритмов интеллектуального анализа данных, таких как: искусственные нейронные сети, дерево решений, метод ближайшего соседа и k-ближайшего соседа, метод опорных векторов байесовские сети, линейная регрессия, корреляционно-регрессионный анализ, иерархические методы кластерного анализа, неиерархические методы кластерного анализа, в том числе алгоритмы k-средних и k-медианы, методы поиска ассоциативных правил, в том числе алгоритм

Apriori, метод ограниченного перебора, эволюционное программирование и генетические алгоритмы, разнообразные методы визуализации данных и множество других методов.

Основная часть аналитических методов, которые используются в технологии интеллектуального анализа данных — это известные математические методы и алгоритмы. Однако теперь появилась возможность их использования при решении тех или иных конкретных проблем, появившаяся благодаря новым возможностям технических и программных средств. Большинство методов интеллектуального анализа данных были разработаны в рамках теории искусственного интеллекта. Метод представляет собой норму или правило, определенный путь, способ, прием решений задачи теоретического, практического, познавательного, управленческого характера [1-10].

### II. ТЕОРЕТИЧЕСКИЕ ОСНОВЫ

В данном исследовании были проанализированы статьи, содержащие доказательную и экспериментальную базу по наиболее актуальным вопросам, касающимся применения методов и алгоритмов интеллектуального анализа данных в системе электронного обучения.

В статье [1] аргументирована актуальность применения методов интеллектуального анализа данных в образовательном процессе с целью поддержки принятия решений по его управлению. В работе предлагается программное приложение, которое обеспечивает проведение анализа данных с использованием сложных алгоритмов классификации.

В работе [2] авторы предлагают ряд методов интеллектуального анализа образовательных данных для оценки учебного поведения студентов в системе электронного обучения.

Они включают кластерный анализ, направленный на определение наиболее популярного временного порога выполнения задания за один сеанс работы, анализ и визуализацию данных для выделения основных параметров, способствующих эффективному прохождению курсов, а также V-кратную перекрестную проверку с применением статистической обработки, позволяющую группировать студентов по основным показателям их деятельности для определения зависимости между высоким процентом активности и успеваемостью. Анализ показал возможность оптимизации навигации по курсу и времени, затрачиваемого на выполнение заданий, при наличии информации о важных инструментах электронного обучения.

В статье [3] рассматриваются методы анализа образовательных данных и их отличия от методов интеллектуального анализа данных. Также анализируются направления применения анализа образовательных данных и преимущества, которые он дает всем участникам образовательного процесса.

В работе [4] рассмотрены различные методы анализа данных об успеваемости студентов. Основное внимание уделено прогнозированию академических достижений студентов по программам бакалавриата, а также изучению типичных прогрессий учащихся и их объединению с результатами прогнозирования. Для прогнозирования было использовано около 10 алгоритмов классификации.

Целью работы [5] является разработка основ технологии самооптимизации программных систем в составе e-learning. Она основывается на сформулированном и формализованном принципе рефлексивной адаптации программного обеспечения, применимого к широкому классу программных систем и основанного на выявлении новых знаний в поведенческой продукции системы.

В статье [6] рассмотрены особенности применения инструментов интеллектуального анализа данных и машинного обучения к небольшим наборам данных. Эта работа не является обзором популярных методов машинного обучения и интеллектуального анализа данных, однако собранный и представленный материал поможет читателю достичь удовлетворительных результатов при применении описанных вычислительных методов.

Согласно исследованию [10], в современной литературе по психологии и педагогике, особое внимание уделяется выявлению факторов, которые влияют на успех студентов в

учебе. Для анализа влияния исследуемых параметров на успеваемость, применяются статистические методы. Однако выводы, полученные в результате такого анализа, часто сталкиваются с проблемами, так как результаты тестирования конкретной выборки студентов оказываются неприменимыми даже для отдельных студентов этой же выборки. Такая ситуация требует выявления возможных факторов, которые ограничивают достоверность исследований, а также необходимости обратить внимание на специфику данных, подлежащих анализу, и объем выборок из общей совокупности, на которых этот анализ будет проводиться.

Применение методов и алгоритмов интеллектуального анализа данных позволяет эффективно обрабатывать и интерпретировать данные о поведении учеников. Эти методы включают машинное обучение, статистический анализ, анализ текстов и другие техники. Они помогают выявить скрытые паттерны, связи и закономерности в данных, которые можно использовать для прогнозирования поведения учеников.

Использование методов интеллектуального анализа данных для оценки поведения учеников в системе электронного обучения имеет ряд преимуществ. Во-первых, это позволяет автоматизировать процесс оценки и снизить нагрузку на преподавателей. Система может автоматически обрабатывать и интерпретировать информацию, предоставляя более объективные результаты. Во-вторых, применение интеллектуального анализа данных помогает получить более глубокое понимание поведения учеников и их потребностей, что приводит к эффективной персонализации образовательного процесса.

### III. ПРИМЕНЕНИЕ МЕТОДОВ

Использование методов анализа образовательных данных приносит значительные преимущества для различных участников образовательного процесса, включая обучающихся, их родителей, преподавателей и администрацию образовательных учреждений.

Для обучающихся такой подход позволяет адаптировать учебные материалы под их текущий уровень знаний и особенности обучения. Например, материалы представляются в виде текстов, видеороликов или графических материалов, а также студенты могут быстро выполнять практические задания, осваивая темы более эффективно. Кроме того, с помощью анализа образовательных данных можно получить

информацию для родителей о прогрессе в обучении и выявить возможные проблемы, которые можно успешно решить.

При применении методов интеллектуального анализа данных в сфере образования нужно учитывать определенные ограничения и анализировать возможные проблемы:

1. Определение необходимости сбора данных: Важно предварительно анализировать, какие данные необходимо накапливать. Это может быть сложной и дорогостоящей задачей. Важно понимать, какие данные подлежат сбору для дальнейшего анализа, а какие не стоит собирать на данном этапе, поскольку их анализ потребует значительных ресурсов.

2. Технические ограничения: Просто принять решение о необходимости сбора данных недостаточно. Важно решить проблемы хранения и управления накопленными данными. Они должны быть доступными, надежными и защищенными.

3. Информационная безопасность: Правила информационной безопасности и этические нормы должны учитываться при использовании образовательных данных. Контроль доступа к этим данным и информирование обучающихся о том, что предоставленная ими информация может быть использована для анализа, являются важными аспектами.

4. Специфика предметной области или образовательного учреждения: Исследование данных по реализации образовательных программ может быть прямо связано с предметной областью изучаемых дисциплин, местом проведения и временем их выполнения. Например, результаты студентов, обучающихся в ведущих высших учебных заведениях, могут отличаться от результатов обучающихся из учреждений другого уровня подготовки. Это требует определенного уровня экспертизы в конкретной сфере.

Преподаватели также получают значимую информацию, которая помогает улучшить содержание учебной программы и ее реализацию через дистанционное обучение. Это стимулирует студентов проявлять больший интерес к учебному процессу и способствует повышению их мотивации для достижения образовательных целей.

Администрация систем электронного обучения получает полезные данные для оценки эффективности отдельных дисциплин или всего образовательного процесса в целом. Это основа для оценки системы обучения и внесения улучшений путем разработки правил ре-

ализации образовательных программ, внедрения новых методов обучения или исключения неэффективных аспектов.

Анализ образовательных данных включает использование различных методов традиционного Data mining. Основные из них включают классификацию, кластеризацию, поиск связующих правил, поиск последовательных шаблонов и интеллектуальный анализ текстов. Все эти методы можно разделить на три большие группы.

Первая группа - прогнозирование. Целью прогнозирования является разработка модели, способной предсказывать значение интересующей переменной на основе значений независимых переменных. В случае, когда интересующая переменная имеет непрерывные значения, мы говорим о регрессии, то есть поиске связи между входными и выходной переменными. Если интересующая переменная имеет конечный набор дискретных значений или классов, то речь идет о задаче классификации, когда в зависимости от значений входных переменных выходная переменная относится к определенному классу. Прогнозирование предполагает наличие размеченных данных, то есть данных, в которых известны значения входных и выходной переменных.

Вторая группа - обнаружение структуры. Алгоритмы обнаружения структуры пытаются выявить в данных структуру, не используя предварительные представления о ней. Кластеризация является одним из наиболее известных методов обнаружения структуры данных. Кластеризация логически продолжает идею классификации, но отличается тем, что классы объектов изначально не определены, в то время как в классификации необходимо разметить часть данных перед созданием модели. Результатом кластеризации является разделение множества объектов на группы близких по какому-либо признаку объектов.

Третья группа - выявление взаимосвязей. Цель выявления взаимосвязей заключается в определении связей между переменными в наборе данных с большим числом переменных. Например, можно попытаться определить, какие переменные наиболее сильно связаны с интересующей переменной, или в каких парах связанных переменных связь наиболее сильная по сравнению с другими.

Прогнозирование основано на использовании и анализе исторических данных. Оно позволяет оценить будущие значения числовых показателей, а также заполнить пропущенные значения.

В математической статистике прогнозирование используется для прогнозирования будущих значений переменных на основе статистического анализа и моделей.

Нейронные сети также широко применяются для прогнозирования. Они используются для обработки больших объемов данных и построения сложных моделей прогнозирования.

Задачи Data Mining, связанные с прогнозированием, могут быть классифицированы по различным стратегиям. Одна из таких стратегий - это «обучение с учителем». В этом случае используются размеченные данные для обучения модели и выполнения прогнозирования, задачи, относящихся к этой категории, включают классификацию, оценку и прогнозирование.

«Обучение без учителя» - это другая стратегия, где модель обучается на размеченных данных и выполняет кластеризацию или другие задачи без использования явных меток классов.

Остальные задачи, которые не относятся к обучению с учителем или без учителя, могут быть классифицированы как «другие». Примеры таких задач включают поиск ассоциаций, обнаружение выбросов и временные ряды.

Интеллектуальный анализ данных в системах электронного обучения является критически важной задачей для извлечения, анализа и управления большими объемами данных. Прогнозирование академического поведения студентов позволяет определить, насколько хорошо отдельные лица и команды будут выполнять различные учебные задания.

Обработка больших данных включает в себя важные ресурсы данных и подходящие аналитические инструменты для создания эффективных подходов, основанных на данных. Эти подходы могут обеспечить раннюю обратную связь, информирование преподавателей и менеджеров, что может помочь улучшить показатели студентов, результаты формальной оценки, также могут быть использованы для выявления подверженных риску студентов, которые могут не сдать итоговые экзамены.

Процедуру прогнозирования можно упростить следующим образом:

1. Определение студентов, в том числе их уровень знаний, эрудиция, личностное соответствие, способность к обучению, начальные знания и другие аспекты.

2. Сбор и извлечение данных о студентах на первом шаге в наборе данных и выполнение предварительной обработки, уделяя особое внимание времени и усилиям, которые влияют на результаты обучения.

3. Интерпретация привычек обучения, чтобы усовершенствовать кривые прогнозирования.

Прогнозирование успеваемости может создать равные шансы для образования и однозначно повысить интерес учащихся.

Студентам обычно предлагают широкий ассортимент курсов и модулей, предоставляя им возможность выбирать как формальное, так и неформальное обучение на электронной обучающей платформе. Однако, сделать правильный выбор может стать реальной дилеммой. В этом случае, система рекомендаций курсов, основанная на анализе огромного объема данных, может оказаться полезной.

Преподаватели, опираясь на глубокое исследование и анализ образовательных данных, могут количественно оценить процесс обучения и статус студента. Они способны обращать внимание на корреляцию и определить причинно-следственные связи. Большие данные позволяют выявлять актуальные проблемы в системе образования, быстро и точно определять факторы влияния и стратегии вмешательства. Это открывает новые правила и значительно расширяет возможности обучения.

Кроме того, большие данные в сфере образования могут анализировать информацию, созданную студентами в процессе обучения. Они могут делать прогнозы относительно паттернов их обучения, а также уровня способностей, выявлять потенциальные проблемы в учебе и систематически улучшать модель обучения.

#### IV. ЗАКЛЮЧЕНИЕ

Анализ литературных данных подтверждает, что важным фактором повышения эффективности образования является разнообразие взаимодействий и обратная связь, адаптированные к конкретным курсам. Начальные знания и индивидуальные учебные планы помогают максимизировать результаты обучения с учетом различных стилей мышления студентов.

Использование больших данных позволяет исследователям в области образования делать выводы о целесообразности и преимуществах и недостатках с новых ракурсов. Это показывает, что каждое состояние обучения может быть замечено с использованием передовых методов, превосходящих традиционные подходы, и что каждый студент может сосредоточиться на своем индивидуальном пути к новым знаниям.

Однако анализ больших данных представляет определенные сложности в связи с характером их сбора, обработки, хранения, передачи, визуализации и обеспечения конфиденциальности информации. Под анализом больших данных понимается применение прогнозной аналитики и других передовых методов с целью извлечения полезной информации из большого объема данных.

Точность анализа больших данных помогает принимать более обоснованные решения. В итоге, принятие наилучших решений способствует повышению производственной эффективности, сокращению расходов и снижению рисков.

#### ЛИТЕРАТУРА

- [1] Горлушкина Н.Н., Коцюба И.Ю., Хлопотов М.В. Задачи и методы интеллектуального анализа образовательных данных для поддержки принятия решений // ОТО. 2015. №1.
- [2] Гущина О.М., Очеповский А.В. Интеллектуальный анализ данных в изучении учебного поведения студентов в системе электронного обучения. Виртуальное моделирование, прототипирование и промышленный дизайн // Материалы VI Международной научно-практической конференции. Том 2, Выпуск 6. 2020, Издательство: Тамбовский государственный технический университет.
- [3] Белоножко П.П., Карпенко А.П., Храмов Д.А. Анализ образовательных данных: направления и перспективы применения // Интернет-журнал «НАУКОВЕДЕНИЕ» Том 9, №4 (2017).
- [4] Н. А. Попова, Е. С. Егорова, Интеллектуальный анализ образовательных данных для прогноза успеваемости студентов вуза, Известия Кабардино-Балкарского научного центра РАН, 2023, выпуск 2, 18–29.
- [5] Бождай А.С., Евсеева Ю.И., Гудков А.А. Применение методов интеллектуального анализа данных для реализации рефлексивной адаптации в системах e-learning. Открытое образование. 2017;(4):13-20.
- [6] Аныш Х. Применение инструментов машинного обучения и интеллектуальный анализ данных в отношении баз данных с небольшим количеством записей. Advanced Engineering Research (Rostov-on-Don). 2021;21(4):346-363.
- [7] F. Matsebula, E. Mnkandla, "A big data architecture for learning analytics in higher education," IEEE africon: Science, Technology and Innovation for Africa, pp. 951-956, 2017.
- [8] J. F. H. Barril, Q. Tan, "Integrating privacy in architecture design of student information system for big data analytics," IEEE Conference on Cloud Computing and Big Data Analysis, pp. 139-144, 2017.
- [9] Якубжанова Д.К., Ахмеджонов А.Р. Роль дистанционного обучения в повышении качества образования в вузах // Science Time. 2016. №2 (26).
- [10] Сосницкий В.Н., Потанин Н.И., Шевелева Л.В. Проблемы статистического анализа средней успеваемости студентов. // Фундаментальные исследования. 2013. № 10-2. С. 316-320.

Поступила в редакцию 30.09.2023

**Цитирование:** Якубжанова Д.К., Абдукаримова Х. (2023). К вопросам применения методов интеллектуального анализа данных в электронном образовании. *Международный Журнал Теоретических и Прикладных Вопросов Цифровых Технологий*, 4(6), –С. 24-28.

#### ON THE APPLICATION OF METHODS INTELLIGENT DATA ANALYSIS IN ELECTRONIC EDUCATION

Yakubjanova D.K.<sup>1</sup>, Abdugarimova Kh.<sup>1</sup>

<sup>1</sup>Samarkand branch of the Tashkent University of Information Technologies named after Muhammad al-Khwarizmi, Samarkand, Uzbekistan, yakubjanova55575@gmail.com

**Abstract.** *The article substantiates the relevance of using data mining methods to support management in the electronic education process. The main methods of data mining in e-education are described in detail, and an overview of the main scientific works devoted to this area is presented. The presented methods can be useful in decision support systems at all levels of educational process management.*

**Keywords:** *data mining, educational process, e-education, method, forecasting, analysis of educational data.*