

Научная статья

DOI: 10.15593/2224-9397/2023.2.07

УДК 004.62/37.012.4

**А.Э. Бойко, Т.В. Савицкая, Д.С. Лопаткин**Российский химико-технологический университет, им. Д.И. Менделеева,  
Москва, Российская Федерация**ПРИМЕНЕНИЕ ИНСТРУМЕНТОВ ИНТЕЛЛЕКТУАЛЬНОГО  
АНАЛИЗА ОБРАЗОВАТЕЛЬНЫХ ДАННЫХ ПРИ РАБОТЕ  
С СИСТЕМАМИ УПРАВЛЕНИЯ ОБУЧЕНИЕМ  
В ОРГАНИЗАЦИЯХ ВЫСШЕГО ОБРАЗОВАНИЯ**

Отечественные вузы всё более интенсивно используют системы управления обучением (LMS) для реализации смешанного формата обучения и построения дистанционных курсов; вместе с этим расширяются и возможности самих LMS в области сбора, хранения, и обработки образовательных данных, открывая потенциал для применения в образовательном процессе инструментов интеллектуального анализа данных, способствующих совершенствованию реализации как отдельных дисциплин, так и целых образовательных программ. **Цель:** осуществить анализ современных инструментов образовательной аналитики (LA) и кейсов использования интеллектуального анализа данных (EDM) при работе с системами управления обучением в организациях высшего образования. **Результаты:** в работе проанализированы отечественные и зарубежные публикации, посвященные технологии «process mining», полагающейся на сбор и анализ «цифровых следов» пользователей обучающих платформ, а также работы, касающиеся практического применения LA и EDM при реализации учебных дисциплин в вузах и на онлайн-курсах. Названы наиболее частые кейсы использования «цифровых следов» для определения уровня активности пользователей, их вовлеченности и паттернов поведения преуспевающих обучающихся. Определена граница понятий LA и EDM, проанализированы наиболее распространённые инструменты обеспечения индивидуальных образовательных траекторий, основанные на извлечении знаний из образовательных данных. Рассмотренные работы систематизированы по целям и областям применения, сложности реализации, а также классифицированы в соответствии с подходом аналитической компании Gartner, предусматривающей 4 вида аналитики: описательную, диагностическую, предиктивную (прогнозную), предписывающую. В работе также рассмотрен альтернативный подход к классификации инструментов EDM по объекту применения, выделяющий: аналитику обучения, прогностическую аналитику поведенческую аналитику и аналитику визуализации. Предложен вариант соотнесения указанных подходов. **Практическая значимость:** настоящий литературный обзор может послужить источником идей и основой для отбора релевантных кейсов образовательной аналитики в отечественных вузах.

**Ключевые слова:** образовательная аналитика, интеллектуальный анализ данных, цифровой след, система управления обучением, предиктивная аналитика, адаптивное обучение, техническое образование, высшее образование.

**A.E. Boyko, T.V. Savitskaya, D.S. Lopatkin**

D. Mendeleev University of Chemical Technology of Russia,  
Moscow, Russian Federation

## **EDUCATIONAL DATA MINING TOOLS APPLICATION IN HIGHER EDUCATION LEARNING MANAGEMENT SYSTEMS**

National universities are increasingly using learning management systems (LMS) to implement a mixed learning format and build online-courses; at the same time, the capabilities of the LMS themselves to collect, store, and process educational data are expanding, opening up the potential for the use of data mining tools in the educational process that contribute to the improvement of the implementation as certain courses, as well as entire educational programs. **Purpose:** to analyze modern learning analytics (LA) tools and cases of the use of educational data mining (EDM) when working with learning management systems in universities. **Results:** the paper analyzes national and foreign publications devoted to the "process mining" technology, which relies on the collection and analysis of LMS users "event logs", as well as works related to the practical application of LA and EDM in the implementation of academic disciplines in universities and online courses. The most frequent cases of using "event logs" to determine the level of user activity, their involvement and patterns of behavior of successful students are named. The boundary of the concepts of LA and EDM is defined, the most common tools for providing individual educational path based on the educational data mining are analyzed. The considered works are systematized by goals and fields of application, complexity of implementation, and also classified in accordance with the approach of the Gartner analytical company, which provides 4 types of analytics: descriptive, diagnostic, predictive, prescriptive. The paper also considers an alternative approach to classifying EDM tools by the object of application, distinguishing: learning analytics, predictive analytics, behavioral analytics and visualization analytics. A variant of the correlation of these approaches is proposed. **Practical relevance:** this literary review can serve as a source of ideas and a basis for selecting relevant cases of educational analytics in national universities.

**Keywords:** learning analytics, educational data mining, user logs, learning methodology system, predictive analytics, adaptive learning, engineering, higher education.

### **Введение**

Актуальность исследования обусловлена необходимостью практического внедрения учебной аналитики в образовательные процессы высших учебных заведений, с одной стороны; с другой – перспективой формирования единого подхода к дескриптивной и диагностической учебной аналитике, способной удовлетворить задачи системы менеджмента качества (СМК) по части представления данных тестирования для целей внутреннего и внешнего контроля качества образования. Кроме этого, ввиду интенсификации процессов построения дистанционных курсов отечественными вузами с использованием систем управления обучением (LMS «Moodle»), предложенный подход может быть использован ими для расширения возможностей мониторинга, представления и визуализации образовательных данных, связанных с различными формами тестирования обучающихся.

Организация сбора, хранения, мониторинга, обработки образовательных данных на уровне систем управления обучением (LMS) в современном мире способствует совершенствованию реализации отдельных учебных дисциплин или даже целых образовательных программ, выработке адаптивных подходов к обучению на основе мониторинга успеваемости студентов, их вовлеченности, а также является основой для построения индивидуальных образовательных траекторий. Ввиду существенного влияния учебной аналитики (learning analytics, LA) и интеллектуального анализа образовательных данных (educational data mining, EDM) на обеспечение качества образования в вузах анализ современных инструментов LA и EDM представляется важной исследовательской задачей для дальнейшего практического внедрения наиболее применяемых и перспективных кейсов в практику реализации отдельных технических дисциплин в вузах, а в перспективе и для расширения мониторинга и совершенствования отдельных образовательных программ.

Об актуальности внедрения инструментов «учебной аналитики» или «аналитики обучения» в образовательные процессы вуза немалое может сообщить само определение этого понятия. Это идентификация, сбор, анализ и представление данных об обучающихся и образовательном процессе с целью изучения обучения, оптимизации образовательного процесса и улучшения результатов обучения [1].

Близким к учебной аналитике и полагающимся на более сложные с технической точки зрения современные инструменты обработки данных понятием выступает «интеллектуальный анализ образовательных данных», ИАД (educational data mining, EDM), представляющий собой процесс нахождения закономерностей в больших наборах данных, включающий методы, находящиеся на пересечении искусственного интеллекта, машинного обучения, статистики и систем управления базами данных [2].

Разница между LA и EDM кроется в следующих различиях [3]:

– обнаружение закономерностей: исследователи в области EDM заинтересованы в автоматическом обнаружении, в то время как целью обучения в LA является использование человеческого суждения;

– источники: интеллектуальный анализ образовательных данных уходит корнями в компьютерные образовательные системы и процессы моделирования обучающихся (ответы, действия, активность, и т.д.).

Истоки аналитики обучения связаны с семантическими сетями, «интеллектуальными» образовательными программами и мерами воздействия на них, прогнозированием результатов обучения;

– адаптация и персонализация: EDM стремится к автоматической адаптации. LA скорее направлена на информирование преподавателей и студентов и расширение возможностей их взаимодействия;

– техники и методы: EDM использует больше методов классификации, кластеризации, байесовского моделирования, анализа взаимосвязей, обнаружения с помощью моделей и визуализации. С другой стороны, LA фокусируется на анализе социальных сетей, анализе настроений, анализе влияния, анализе дискурса, прогнозировании успеха учащихся, концептуальном анализе и «моделях формирования смысла».

### **1. Технология «process mining» как основа EDM и LA в современных образовательных системах**

Один из прогрессивных современных подходов интеллектуального анализа образовательных процессов – технология Process Mining (PM)<sup>1</sup> полагается на цифровые следы (logs, «логи» – цифровые следы), которые возникают благодаря повсеместной реализации образовательных процессов с использованием ИТ-систем. Из совокупности фиксирующихся цифровых следов возникают журналы событий, из которых можно извлечь полезную информацию и восстановить процесс, как он есть на самом деле. По мнению автора одной из наиболее значимых работ в этой области [4], «Process Mining» является важной частью науки о процессах («Process science») – более широкой дисциплины, которая объединяет знания из области информационных технологий и менеджмента для управления и улучшения операционными процессами. Рост интереса к РМ обусловлен крайне широким спектром его применения. Так же, как MS «Excel» применяется для сферы финансов, образования, спорта, медицины, и др., растущие возможности обработки данных сегодня обеспечивают реализацию «цифровых следов» в ИТ-системах практически для любых операционных процессов, поэтому РМ непременно способствует совершенствованию сквозных процессов. Всё больше разработчиков ПО из академической и коммер-

---

<sup>1</sup> Термином Process Mining обозначают комплекс методов и подходов, использующихся для интеллектуального анализа образовательных процессов на основе цифровых следов.

ческой сферы внедряют функционал РМ в свои продукты, поскольку он способен отслеживать выполнение любых процессов системы, давая возможность обнаруживать нарушения прямо в ходе их выполнения, а также определять узкие места, соответственно повышать производительность процессов, а также способен делать прогнозы и предоставлять рекомендации.

Помимо вышеуказанных преимуществ авторы [4] подчеркивают, что РМ технологически перспективны для реализации алгоритмов машинного обучения и интеллектуального анализа данных, что абсолютно необходимо для образовательных данных. Большинство исследований будут тому подтверждением.

Идею применения РМ в системах управления обучением в организациях высшего образования наиболее точно описывает следующее высказывание:

– «Преуспевающие студенты демонстрируют иные паттерны поведения (взаимодействия с обучающей системой) по сравнению с неуспевающими» (перевод) [5].

Для анализа этих паттернов исследователи наиболее часто обращаются к следующим «цифровым следам» обучающихся систем:

– активность на форумах: количество просмотренных на форуме сообщений (`discussion_viewed`), опубликованных и созданных сообщений (`post/discussion_created`), подписок на темы форума (`discussion_subscription_created`);

– просмотр образовательного контента (элементов или ресурсов курса): курс просмотрен (`course_viewed`), `course_module_viewed` (модуль курса просмотрен);

– работа с элементом курса «задание» (`assessable_uploaded`);

– общее количество кликов (`#of total clicks -> All the events`).

Хороший пример реализации аналитики, основанной на «цифровых следах», к примеру, в работах [6–9]. В работе [6] авторы описали создание сводных отчетов по цифровой активности отдельных обучающихся, где представлены: график активности пользователя, график по типам событий взаимодействия пользователя с разнообразными ресурсами курса и гистограмма предпочитаемых материалов. В связи с тем, что детали событий, содержащихся в логах системы, специфичны для каждого вида событий, авторы разделили их на группы: навигация пользователей, взаимодействие с видеоматериалами, взаимодей-

ствие с документами, ответы на вопросы, работа с различными курсами. Для извлечения структурированных логов системы авторы написали программный код с использованием библиотеки `psycopg2` [10], для хранения данных использовали СУБД PostgreSQL, используя в качестве ORM (Object-Relational Mapping – «виртуальной объектной базы данных») библиотеку SQLAlchemy [11], взаимодействуя с БД посредством интерфейса ORM. Разработанное авторами веб-приложение позволило просматривать активность отдельных обучающихся из платформы edX Университета ИТМО, а, кроме того, созданная информационная может послужить для разработки методов построения индивидуальных образовательных траекторий.

В работе [7] показано исследование интенсивности использования дистанционного курса за определенное время с помощью стандартного «Отчета о деятельности» (в нем были установлены временные фильтры на период реализации курса) и отчета «Участие в курсе» из LMS «Moodle» с помощью инструментов MS Excel (корреляционный анализ и поиск ассоциативных правил), по результатам которого было замечено, что наблюдается устойчивая тенденция к снижению учебной активности в течение всего семестра.

В отличие от предыдущей работы авторы получили обобщенные количественные характеристики по просмотрам вспомогательных материалов по разделам курса для всей учебной группы, а не для отдельных обучающихся, что, несмотря на меньшие трудозатраты для реализации, также может быть использовано для автоматизации рассылки сообщений обучаемым в связи с отсутствием просмотров конкретных вспомогательных материалов.

В работе [8] представлен профиль академической компетентности студента, сформированный на основе диагностических методик психологического толка, дополненный формами обратной связи обучающихся и данными об активности студентов из системы Moodle.

Сбор «цифровых следов» пользователей не ограничивается только LMS Moodle или EdX. Не менее ярким примером формирования отчетов о цифровой активности пользователей на основе логов является образовательная платформа «Юрайт», которая собирает и агрегирует сведения о статистике её использования, в том числе данные о следующих показателях [12]:

- время, которое пользователь провел на платформе за изучением курса или чтением учебника;
- количество пройденных пользователем тестовых заданий и заданий с открытым ответом и результаты их выполнения;
- количество просмотренных пользователем образовательных медиаматериалов (видео и аудио);
- создание и (или) изучение пользователем гибких курсов на платформе;
- успеваемость пользователя, имеющего авторизованную роль студента в группе пользователя и авторизованную роль преподавателя.

Актуальным направлением EDA также является анализ тональности (эмоциональной окраски) образовательных текстовых данных и задача классификации текстов комментариев, что прекрасно продемонстрировано в работе [13], где на основе скрапинга<sup>2</sup> отзывов с платформы «Stepik» сформирован датасет из 5721 русскоязычного отзыва по математике, программированию, биологии, химии и физике. Выполнено исследование русскоязычных отзывов из датасета на основе описательной статистики, частотного анализа униграмм и биграмм, сентиментного анализа<sup>3</sup> с помощью python-библиотеки «dostoevsky», продемонстрировавшего 74%-ную точность классификации по классам тональности на основе взвешенной метрики F1-score<sup>4</sup>.

Обеспечивают возможности организации такой аналитики инструменты сбора и анализа данных, интегрированные в образовательные платформы Moodle, Open edX, GetCourse, Coursera, национальный портал «Открытое образование» и др. Во всех указанных платформах так или иначе реализована технология Process Mining (PM), благодаря которой образовательные процессы оставляют цифровые следы (логи), которые по большей части и служат основой для обеспечения всех указанных видов аналитики.

Отметим, что EDA в настоящий момент в большей степени используется на таких коммерческих образовательных платформах, как

---

<sup>2</sup> Или скрепинг, или скрапинг (англ. web scraping) – это технология получения веб-данных путем извлечения их со страниц веб-ресурсов.

<sup>3</sup> Сентиментный анализ (анализ тональности) – это область компьютерной лингвистики, занимающаяся изучением эмоций в текстовых документах, в основе которой лежит машинное обучение.

<sup>4</sup> Среднее гармоническое значение между точностью модели (precision) и полнотой (recall).

Duolingo, Kahoot, Udemy, SoloLearn, SkyEng, и др., поскольку от качества адаптации образовательных материалов к обучающимся (по скорости подачи, сложности, объему, интенсивности, стилю общения) и персонализации обучения, а также наличия инструментов привлечения внимания и формирования привычки учиться, реализуемых через отслеживание действий и рекомендации (напоминания и рекомендации в «пуш-уведомлениях»<sup>5</sup>, инструменты «игрофикации», контент в социальных сетях и др.), зависят конкурентоспособность и целесообразность масштабирования платформ.

В случае же с применением EDA в образовательных программах государственных вузов их конкурентоспособность находится несколько в другой плоскости, и, что не менее важно, такие программы преследуют принципиально другие задачи, как то соблюдение требований Федеральных государственных образовательных стандартов (ФГОС), формирование компетентностной модели, обеспечение актуальности содержания образования современным отраслевым практикам и др.; кроме того, перечень дисциплин, реализуемых в рамках подготовки одной образовательной программы, может быть, сопоставим с общим количеством образовательных курсов, которые реализуются образовательными платформами.

В связи с этим в образовательных программах государственных вузов чаще применяется именно образовательная аналитика, которую в терминологии консалтинговой Гарнер (Gartner, Inc.) можно охарактеризовать как описательную (дескриптивную), диагностическую и предиктивную.

## **2. Характеристика основных видов образовательной аналитики в соответствии с подходом аналитической компании Gartner**

Описательная аналитика отвечает на вопрос «Что произошло?» и позволяет выявить проблему на основе простых данных: времени прохождения того или иного этапа курса, качества ответов и т. д [2].

Применительно к образовательным данным под описательной аналитикой наиболее часто понимают простые отчеты об оценках тестирования (включая срезы по группам обучающихся), времени прохождения курса, числе пользователей, просматривающих учебные эле-

---

<sup>5</sup> Это короткие сообщения, которые веб-ресурс рассылает своим подписчикам на компьютеры и мобильные устройства.



менты (или как-либо взаимодействующих с ними), количестве пользователей, получивших сертификат, анализе качества тестовых заданий, реализуемых на основе индекса сложности (трудности), дискриминации и расчета среднеквадратичного отклонения, как это сделано в работе [14]. Авторы подчеркивают, что применение LMS Moodle позволяет стандартизировать процедуру апробации теста, выявлять задания, требующие корректировки, и способствует повышению точности и созданию качественных тестов. Отметим, что в последних версиях LMS Moodle появилась цветовая индикация для обозначения вопросов, вероятно, требующих корректировки на основе статистики их использования, что упрощает совершенствование тестирования для педагогов, не глубоко погруженных в вопросы теории тестирования.

*Диагностическая аналитика* выявляет причину возникновения проблемы. Например, почему 70 % пользователей дали неверные ответы? Анализ покажет, что привело к такому результату: контент, способы его подачи, условия прохождения курса либо низкий базовый уровень пользователей [2].

Традиционно диагностическая аналитика реализуется статистическими методами, вероятно, поэтому в ГОСТ Р 59895-2021 «Технологии искусственного интеллекта в образовании» данное понятие не выделяют как связанное с использованием технологий искусственного интеллекта в образовании. Кроме этого часть функций, заложенная компанией Gartner в понятие «диагностическая аналитика», отражена в вышеназванном ГОСТе в термине «Описательная аналитика»: идентификация, сбор, анализ и представление данных об образовательном процессе и деятельности образовательной организации с целью выявления интерпретируемых зависимостей и закономерностей, понимания текущей ситуации и происходящих процессов. Таким образом, основываясь на формулировке «выявление интерпретируемых зависимостей и закономерностей», можно утверждать, что описательная и диагностическая аналитика, с точки зрения разработчиков национального стандарта, лежат в одной плоскости.

В качестве репрезентативного кейса реализации диагностической аналитики для оценки качества «контента» (теста) можно назвать работу [15], где исследовалась зависимость оценок экзамена от результатов тестирования по дисциплине «Информатика» на основе статистической зависимости экзаменационных оценок от оценок тестирования

с опорой на функцию линейной регрессии и критерий Фишера. В случае наличия статистической значимости можно сделать вывод о соответствии содержания тестирования содержанию экзаменационных вопросов, о склонности преподавателей оценивать студентов более «строго» или более «мягко». Кроме этого авторы затронули вопрос многократного тестирования обучающихся одним набором вопросов на разных потоках. В этом случае плохо успевающие студенты могут недобросовестно выполнить тест (списывание), как следствие, исследователи не обнаружили статистической значимости между результатами тестирования и экзаменационной оценкой.

*Предиктивная*, или прогнозная, аналитика, отвечая на вопрос «Что произойдет дальше?», позволяет спрогнозировать, насколько пользователи мотивированы закончить модуль или курс, выделить группу пользователей с высоким риском не завершить его освоение.

Наиболее распространенными методами для осуществления «умной» аналитики обучения на основе данных о студенческой успеваемости являются подходы, основанные на алгоритмах машинного обучения и применении нейронных сетей, которые вносят значительный вклад в высокое качество, точность и эффективность систем LA (learning analytics).

Среди методов машинного обучения, распространённых среди специалистов по анализу данных в любых областях, для задач LA можно выделить: метод опорных векторов<sup>6</sup>, линейную регрессию<sup>7</sup>, логистическую регрессию<sup>8</sup>, деревья решений<sup>9</sup>, «случайные леса»<sup>10</sup>, «случайные деревья»<sup>11</sup>. При реализации указанных методов исследователи нередко обращаются к таким решениям, как: Scikit-learn, Weka, Dataiku, Deductor, Statistica и др.

---

<sup>6</sup> C-Support Vector Classification. Scikit-learn documentation: <https://scikit-learn.org/stable/modules/generated/sklearn.svm.SVC.html>

<sup>7</sup> Ordinary least squares Linear Regression. Scikit-learn documentation: [https://scikit-learn.org/stable/modules/generated/sklearn.linear\\_model.LinearRegression.html](https://scikit-learn.org/stable/modules/generated/sklearn.linear_model.LinearRegression.html)

<sup>8</sup> Logistic Regression (aka logit, MaxEnt) classifier. Scikit-learn documentation: [https://scikit-learn.org/stable/modules/generated/sklearn.linear\\_model.LogisticRegression.html](https://scikit-learn.org/stable/modules/generated/sklearn.linear_model.LogisticRegression.html)

<sup>9</sup> Decision Trees (DTs). Scikit-learn documentation: <https://scikit-learn.org/stable/modules/tree.html>

<sup>10</sup> A random forest classifier. Scikit-learn documentation: <https://scikit-learn.org/stable/modules/generated/sklearn.ensemble.RandomForestClassifier.html>

<sup>11</sup> An ensemble of totally random trees. Scikit-learn documentation: <https://scikit-learn.org/stable/modules/generated/sklearn.ensemble.RandomTreesEmbedding.html>

В одной из важных работ о применении машинного обучения [16] авторы рассмотрели способы представления данных об успеваемости обучающихся (student academic performance, SAP) и их особенности, а также провели сравнительный анализ эффективности трех групп алгоритмов машинного обучения (machine learning, ML) для обработки данных SAP по классическим для ML метрикам: среднеквадратическое отклонение (mean-squared (MS) errors); корень из среднеквадратичного отклонения (root-mean-squared (RMS) errors); среднее абсолютное отклонение (mean absolute (MA) errors); среднее относительное отклонение (relative-squared (RS) errors); корень из среднего относительного отклонения (root-relative-squared (RRS) errors) и др.

Для различных форм данных SAP для каждой из указанных метрик были обучены и сопоставлены по эффективности модели, на основе чего даны рекомендации по использованию различных методов ML для следующих форматов данных:

– для данных SAP в форме абсолютных чисел (каждое аттестационное испытание имеет свое максимальное и минимальное значение в целых числах), наиболее эффективными методами являются метод опорных векторов и линейная регрессия;

– для данных SAP в форме относительных чисел (в процентах от максимума баллов) также наиболее эффективными методами являются метод опорных векторов и линейная регрессия;

– для данных SAP в форме обычных буквенных оценок (A/B/C/D/F) и подробных буквенных оценок (A+, A, A–, B+, B, B–, C+, C, C–) наиболее эффективными методами ML являются многослойный перцептрон и логистическая регрессия.

В работе [17] для прогнозирования итоговой оценки за курс используются не только SAP, но и данные о количестве времени, проведенного в системе, кликами по элементам курса (лекциям, конспектам и лабораторным работам). Авторы построили нейронную сеть из трёх скрытых слоев с использованием библиотеки TensorFlow Module: tf.keras<sup>12</sup> (для глубокого обучения) и добились точности предсказания итоговой оценки, равной, 91,7 %, по дисциплине «Основы программирования».

---

<sup>12</sup> Implementation of the Keras API, the high-level API of TensorFlow. TensorFlow documentation: [https://www.tensorflow.org/api\\_docs/python/tf/keras](https://www.tensorflow.org/api_docs/python/tf/keras).

В [18] был использован байесовский подход к мониторингу эффективности оценочных средств электронного образовательного курса. Для своевременного выявления студентов, имеющих высокий риск неаттестации в Сибирском федеральном университете (Красноярск) по дисциплине «Теория вероятностей и математическая статистика», была построена байесовская сеть, которая по текущим показателям учебной деятельности студентов прогнозирует успешность освоения дисциплины.

Байесовский классификатор в совокупности с индексом метакогнитивной осведомленности (Metacognitive Awareness Index, MAI) обучающихся был использован в [19] в качестве экспертной рекомендательной системы для формирования адаптивного пути обучения.

В [20] разработанная искусственная нейронная сеть позволила автоматически определять степень усвоения обучающимся той или иной темы учебного материала и выдавать рекомендации по повторному обучению, не прибегая к полному анализу результатов тестирования преподавателем вручную. В [21] на основе оценок по базовым дисциплинам первого учебного модуля (триместра) на примере 274 студентов первого курса бакалавриата направления «Прикладная математика и информатика» с помощью пакета «STATISTICA: Automated Neural Networks» была создана нейросетевая модель, позволяющая определять группу риска по успеваемости с ошибками первого и второго рода в 10, 46.

Одним из важных методов машинного обучения, находящихся на стыке диагностической и предиктивной аналитики, является метод кластеризации (нередко используется совместно с решающими деревьями), который используется как в процессе обучения для выявления «кластеров» неуспевающих студентов, так и по окончании обучения для проведения пост-аналитики с целью адаптации структуры курса, а также для идентификации отдельных факторов, влияющих на успеваемость [22–24]. В [22] кластерный анализ был выполнен путем применения алгоритма кластеризации  $k$ -средних к набору данных с последующим контент-анализом кластеров. Для определения оптимального числа кластеров была использована «средняя ширина силуэта» (Average Silhouette Width (ASW)). В набор данных вошли как данные SAP, так и данные по взаимодействию с элементами курса. В результате исследователи получили три кластера обучающихся, отличающихся оценками по всем элементам курса и количеству неуспе-

вающих студентов; провели сравнение медиан, мод, стандартного отклонения и частот прохождения тестов для самооценки по кластерам. В итоге исследователи диагностировали различный уровень понимания предметной области обучающихся и уровень активности при прохождении курса.

Таким образом, аналитика обучения определяет стратегию на этапах внедрения и оценки эффективности образовательного решения: описательная и диагностическая аналитика – это анализ того, что уже произошло, а прогностическая и предписывающая аналитика – прогноз и предписание будущих действий. В иерархии «уровней аналитики» данных предписывающая опирается на все «предшествующие» виды аналитики, и её реализация возможна только при наличии инструментов для обеспечения всех предшествующих видов аналитики.

К примеру, в работе [7] описаны «уровни персонафицированности цифровой среды»: третий, наивысший уровень, обеспечивается следующими средствами и задачами анализа данных: выявление скрытых от непосредственного наблюдения закономерностей для организации адаптивного обучения, аппарат машинного обучения, интеллектуального анализа данных, описательной статистики. Ключевое отличие наивысшего уровня «персонафицированности» от нижестоящих кроется в наличии аппарата машинного обучения, который в полной мере способен обеспечить LMS инструментами EDM.

Объединив классификацию Gartner и уровни «персонафицированности цифровой среды», получим графическое представление «иерархии аналитики» в образовательных системах (рис. 1).

На основе *предписывающей аналитики* принимаются решения о действиях после прохождения курса: как изменить курс или отдельный модуль, какие новые курсы предложить пользователю. В работе [8] со ссылкой на ряд исследований отмечается, что LMS Moodle, действительно, имеет широкие возможности и потенциал для использования аналитики обучения, но использование полученных моделей для обеспечения адаптивности и персонализированных образовательных маршрутов возможно только в ручном режиме.

Стоит уточнить, что на некоторых образовательных платформах не из сферы высшего образования в той или иной форме всё-таки присутствуют элементы предписывающей аналитики: наиболее распространённой формой её являются рекомендательные системы [25].

К примеру, на образовательных платформах «Stepik» и «Coursera» имеются разделы с рекомендуемыми к изучению курсами, сформированными на основе просмотренных разделов и страниц; также примерами реализации предписывающей аналитики являются уведомления о «риске отчисления» и системы напоминаний, диагностирующие низкую активность обучающего и генерирующие соответствующие напоминания о контрольных точках, сроках сдачи работ и элементы «игрофикации», нацеленные на повышение вовлеченности в образовательный процесс.



Рис 1. Взаимосвязь иерархии «уровней аналитики» и задач, решаемых на разных уровнях

Также существуют отдельные примеры практической реализации предиктивной аналитики на крупных образовательных платформах. В работе [26] исследователи предлагают вероятностную модель, основанную на исторических данных, которая используется для рекомендации персонализированных последовательностей уроков с целью помочь учащимся подготовиться к конкретным аттестационным испытаниям. Модель, основанная на алгоритмах Bayesian Knowledge Tracing (ВКТ), применялась на реальном датасете компании Knewton (специализируется на адаптивном обучении). На отечественном рынке образования также присутствуют примеры полностью адаптивных образовательных платформ, к примеру, онлайн-система адаптивного обучения математике «Platio», которая индивидуально подбирает траекторию

обучения в зависимости от уровня подготовки обучающегося и прогресса студента (диагностическое экспресс-тестирование, на основе которого при помощи генетического алгоритма создается цифровой двойник студента), в процессе выполнения заданий система оценивает прогресс обучающегося и корректирует траекторию обучения [27]. По мнению автора [28, 29], именно способность адаптироваться к модели обучаемого, к образовательному контенту, а также способность оказывать поддержку в решении задач определяют «уровень интеллектуальности» обучающих систем.

### **3. Классификация инструментов EDM по области применения**

Существует и другой подход к классификации инструментов EDM. В [30] авторы отобрали для анализа более 1200 публикаций по теме EDM, из которых 550 были включены в литературный обзор; они пришли к выводу о том, что работы в этой области в основном сосредоточены:

- на компьютерной аналитике обучения (computer-supported learning analytics, CSLA);
- компьютерной прогностической аналитике (computer-supported predictive analytics, CSPA);
- компьютерной поведенческой аналитике (computer-supported behavioral analytics, CSBA) и компьютерной аналитике визуализации (computer-supported visualization analytics, CSVA).

Приведенные в работе публикации также можно классифицировать в соответствии с предложенной группировкой. Такой подход не создаёт «иерархию» аналитических процессов, а сосредотачиваются на отдельных областях применения инструментов EDM под цели и задачи образовательных организаций, поэтому отдельные группы кейсов применения EDM не получается однозначно классифицировать в соответствии с подходом Garter ввиду большого количества публикаций, входящих в подгруппы, представленные на рис. 2, и их разнонаправленностью; поэтому на рисунке «стикером» 1 помечен вид аналитики, к которому наиболее тяготеют работы в представленной подгруппе, а стикерами 2 отмечены виды аналитики, которые также имеют место в этой области, но не являются приоритетными.



Рис. 2. Взаимосвязь направлений публикаций по теме EDM, предложенных в [30] с видами публикаций по Gartner



К примеру, по мнению авторов, группа работ, посвященных «отсеvu и вовлеченности», с подгруппами работ, посвященных направлениям «рефлексия и осознанность», «уровень удовлетворённости», «мотивация», относится к «прогностической аналитике» (CSPA), хотя указанные работы по большей части посвящены выстраиванию систем диагностики мотивации, уровня удовлетворённости и рефлексии, и уже потом прогнозированию показателей для оценки потенциального уровня отсева обучающихся.

Как видно на рис. 2, наибольший интерес исследователей находится в области прогнозной (предиктивной) аналитики, описательной (дескриптивной) и диагностической аналитики.

Рассмотренные в нашем исследовании работы также охватывают практически полный перечень подгрупп, представленных на рис. 2, за исключением вопросов совместного обучения и самообучения, которые наиболее актуальны для применения в гуманитарных дисциплинах.

Несмотря на приведенные в начале работы различия между LA и EDM, на практике довольно сложно отделить одно направление от другого ввиду схожести решаемых задач; самым существенным отличием можно назвать применение сложных алгоритмов и машинного обучения для оптимизации образовательного процесса и улучшения результатов обучения.

### **Заключение**

Бурное развитие образования с применением дистанционных образовательных технологий (ДОТ) последние несколько десятков лет и возрастающий спрос на массовые открытые онлайн-курсы способствуют внедрению инструментов EDM в образовательные процессы высших учебных заведений.

В Российской Федерации вузы активно строят дистанционные курсы на базе LMS Moodle. Россия входит в топ 10 стран мира по числу сайтов, использующих Moodle в официальных образовательных структурах (37 %). Для сравнения: в США – 34 %, в Италии – 32 % [6]. По данным команды Edutechnica, Moodle также входит в тройку самых распространенных в мире LMS: Blackboard, Moodle и Canvas [31]. По общему количеству зарегистрированных сайтов Moodle Российская Федерация находится на восьмом месте в мире [32].

Наиболее массовое предпочтение российскими вузами по реализации дистанционного и смешанного обучения в этой системе сохраняется уже в течение довольно продолжительного времени, что диктуется следующими аспектами среды: бесплатная модульная система электронного обучения с открытым кодом и подробной технической документацией, сильное сообщество (которое и развивает платформу, создавая новые модули и обеспечивая широкий функционал), полная кастомизация и локализация более чем на 100 языков, гибкая система статистики и отчетов [7]. Последнее особенно важно для реализации элементов CSLA, CSPA и CSVA в системах управления обучением.

Анализ публикаций по указанным направлениям показал, что во многих вузах страны уже функционирует описательная, диагностическая и предиктивная аналитика; а следующим шагом станет интеграция всех аналитических процессов в «дата-центрированную» систему принятия образовательных решений.

Наиболее распространёнными современными кейсами использования EDM в российских вузах являются: прогностическая аналитика в области совершенствования обучающих материалов (совершенствование сложности заданий, принятие решений, основанных на обратной связи обучающихся), прогностическая аналитика в области оценивания и мониторинга успеваемости обучающихся, отсева и оценки вовлечённости, в том числе с использованием машинного обучения и нейронных сетей; а также описательная и диагностическая аналитика в области моделирования поведения обучающихся на основе технологии РМ и «цифровых следов».

Наиболее перспективными для практического внедрения технологиями для РХТУ им. Д.И. Менделеева на текущем этапе будет описательная и диагностическая аналитика в области моделирования поведения обучающихся и дескриптивная/прогностическая аналитика в области оценивания и мониторинга успеваемости обучающихся по техническим дисциплинам, реализующимся для всех направлений подготовки.

Подобные «дата-центрированные» системы принятия образовательных решений, учитывающие уровень подготовки обучающихся, их интересы и стиль обучения, обеспечат широкие возможности мониторинга процесса обучения, высокую степень персонализации образовательного контента, что в значительной степени способствует повышению качества образования.

Таким образом, можно заключить, что отечественные организации высшего образования обширно применяют инструменты EDM при работе с системами управлением обучения, делая из Moodle и электронных информационно-образовательных сред полноценные интеллектуальные образовательные системы, которые характеризуются следующим [33]: переход от обучения, управляемого только системой, к обучению, управляемому как системой, так и обучаемым; более качественное консультирование и более глубокий анализ ответов обучающихся при тестировании и существенное упрощение подготовки тестовых заданий; более полный анализа усвоения материала и психофизического состояния обучаемого и, как следствие, более адекватная реакция системы; более полное использование арсенала современной дидактики, т.е. более эффективное осуществление процесса управления обучением.

### **Библиографический список**

1. Определение понятия «Учебная аналитика». ГОСТ Р 59895-2021. «Технологии искусственного интеллекта в образовании. Общие положения и терминология» [Электронный ресурс] // База ГОСТов allgosts.ru. – URL: [https://allgosts.ru/35/240/gost\\_r\\_59895-2021](https://allgosts.ru/35/240/gost_r_59895-2021) (дата обращения: 24.02.2023).
2. Ширинкина Е.В. Интеллектуальный анализ образовательных данных // Вестник Твер. гос. ун-та. Сер. Экономика и управление. – 2021. – № 3 (55). – С. 179–188. DOI: 10.26456/2219-1453/2021.3.179-188
3. Liñán L.C., Pérez Á.A.J. Education Data Mining and Learning Analytics: Differences, similarities and time evolution [Электронный ресурс] // RUSC Universities and Knowledge Society Journal. – 2015. – 12 (3). – P. 98–112. – URL: <https://doi.org/10.7238/rusc.v12i3.2515> (дата обращения: 24.02.2023).
4. Van Der Aalst W. Process mining: data science in action. – Heidelberg: Springer, 2016. – Vol. 2.
5. Rosário N.A.L. Predicting academic performance – a practical study using Moodle log data and sociodemographic traits. – 2022.
6. Романов А.А., Волчек Д.Г. Анализ данных о поведении пользователей в системах электронного обучения // Онтология проектирования. – 2020. – Т. 10, № 1 (35). – С. 100–111. DOI: 10.18287/2223-9537-2020-10-1-100-111

7. Тербушева Е.А., Пиотровская К.Р. Аналитический потенциал платформы Moodle для мониторинга качества персонафицированного обучения // Общество. Коммуникация. Образование. – 2021. – Т. 12. – № 4. – С. 19–34. DOI: 10.18721/JHSS.1

8. Морозов А.И. Аналитика обучения студентов при реализации массового онлайн-курса // Известия Волгоград. гос. пед. ун-та. – 2022. – № 5 (168). – С. 16–22.

9. Alignment of teacher's plan and students' use of lms resources. analysis of moodle logs // 2016 15th international conference on information technology based higher education and training (ithet). – IEEE, 2016. – P. 1–8.

10. The Python SQL Toolkit and Object Relational Mapper. Documentation [Электронный ресурс]. – URL: <https://docs.sqlalchemy.org/en/20/> (дата обращения: 24.02.2023).

11. Psycopg – PostgreSQL database adapter for Python [Электронный ресурс]. – URL: <https://www.psycopg.org/docs/>

12. Соглашение о сборе о цифровой активности пользователей [Электронный ресурс] // Официальный сайт ООО «Электронное издательство Юрайт». – URL: <https://urait.ru/info/publication-data-policy> (дата обращения: 24.02.2023).

13. Дюличева Ю.Ю. Датасет для анализа русскоязычных отзывов на MOOK, извлеченных с платформы Stepiк // Вопросы образования. – 2022. – № 4. – С. 298–321.

14. Кайдалова Л.В., Гуменникова Ю.В., Черницына Р.Н. Статистический анализ результатов тестирования по разделу «Линейная алгебра и аналитическая геометрия» в среде Moodle // Известия Самар. науч. центра Рос. акад. наук. Социальные, гуманитарные, медико-биологические науки. – 2019. – Т. 21, № 65. – С. 35–39.

15. Батасова В.С. Анализ объективности результатов тестирования знаний // Информатизация инженерного образования. – 2020. – С. 27–30.

16. Smart learning analytics: student academic performance data representation, processing and prediction / V.L. Uskov [et al.] // Smart Education and e-Learning 2020 / ed. V.L. Uskov, R.J. Howlett, L.C. Jain. – Singapore: Springer, 2020. – P. 3–18.

17. Технологии нейронных сетей для прогнозирования успеваемости обучения студентов в электронной информационно-образовательной среде вуза / В.И. Токтарова [и др.] // Цифровая гуманитаристика и технологии в образовании (DHTE 2022). – 2022. – С. 388–398.

18. Кустицкая Т.А. Байесовский подход к мониторингу эффективности оценочных средств электронного образовательного курса // Дистанционное обучение в высшем образовании: опыт, проблемы и перспективы развития. – 2020. – С. 164–165.

19. Agustianto K. et al. Design adaptive learning system using meta-cognitive strategy path for learning in classroom and intelligent tutoring systems // AIP Conference Proceedings. – AIP Publishing LLC, 2016. – Vol. 1755. – № 1. – P. 070012.

20. Попова Ю.Б., Яцынович С.В. Применение искусственной нейронной сети для определения степени усвоения учебного материала обучающимися // Новые горизонты – 2018: сб. материалов Белорус.-Китайск. молодеж. инновац. форума: в 2 т.; 15–16 ноября 2018 г. – Минск: Изд-во БНТУ, 2018. – Т. 2. – С. 154–159.

21. Русаков С.В., Русакова О.Л., Посохина К.А. Нейросетевая модель прогнозирования группы риска по успеваемости студентов первого курса // Современные информационные технологии и ИТ-образование. – 2018. – № 4. – С. 815–822.

22. Balaban I., Filipović D., Zlatović M. Post hoc identification of student groups: Combining user modeling with cluster analysis // Education and Information Technologies. – 2022. – P. 1–26.

23. Martinez Abad F., Chaparro Caso López A.A. Data-mining techniques in detecting factors linked to academic achievement // School Effectiveness and School Improvement. – 2017. – Vol. 28. – № 1. – P. 39–55.

24. A data mining approach to study the impact of the methodology followed in chemistry lab classes on the weight attributed by the students to the lab work on learning and motivation / M. Figueiredo [et al.] // Chemistry Education Research and Practice. – 2016. – Vol. 17, № 1. – P. 156–171.

25. Aher S.B., Lobo L. Combination of machine learning algorithms for recommendation of courses in E-Learning System based on historical data // Knowledge-Based Systems. – 2013. – Vol. 51. – P. 1–14.

26. Reddy S., Labutov I., Joachims T. Latent skill embedding for personalized lesson sequence recommendation: arXiv preprint arXiv:1602.07029. – 2016.

27. Цифровой репетитор Plario. Первая в России онлайн-система адаптивного обучения математике [Электронный ресурс]. Официальный сайт платформы plario.ru. – URL: <https://plario.ru/> (дата обращения: 24.02.2023).

28. Попова Ю.Б. От LMS к адаптивным обучающим системам // Системный анализ и прикладная информатика. – Минск: Изд-во Белорус. нац. техн. ун-та, 2019. – № 2. – Р. 58–64.

29. Попова Ю.Б. Интеллектуальная составляющая автоматизированной системы обучения cats // Образовательные технологии и общество. – Казань: Изд-во Казан. нац. исслед. технолог. ун-та, 2019. – Т. 22, № 4. – С. 24–37.

30. Aldowah H., Al-Samarraie H., Fauzy W.M. Educational data mining and learning analytics for 21st century higher education: A review and synthesis // Telematics and Informatics. – 2019. – Vol. 37. – Р. 13–49.

31. LMS Data – Spring 2018 Updates | edutechnica [Электронный ресурс]. – URL: <https://edutechnica.com/2018/03/04/lms-data-spring-2018-updates/> (дата обращения: 24.01.2023).

32. Moodle statistics. World Moodle registration map [Электронный ресурс] // Официальный сайт «Moodle.org». – URL: <https://stats.moodle.org/> (дата обращения: 02.05.2023).

33. Голенков В.В. Интеллектуальные обучающие системы и виртуальные учебные организации // Дистанционное обучение – образовательная среда XXI века: материалы междунар. науч.-метод. конф.; Минск, 18–20 декабря 2001 г. – Минск: Бестпринт, 2001. – С. 21–26.

## References

1. Opređenje poniatia «Uchebnaia analitika». GOST R 59895-2021. «Tekhnologii iskusstvennogo intellekta v obrazovanii. Obshchie polozheniia i terminologiiia» [Definition of the concept of "Educational analytics". GOST R 59895-2021. "Artificial intelligence technologies in education. General provisions and terminology"]. Baza GOSTov allgosts.ru, available at: [https://allgosts.ru/35/240/gost\\_r\\_59895-202](https://allgosts.ru/35/240/gost_r_59895-202) (accessed 24 February 2023).

2. Shirinkina E.V. Intellektual'nyi analiz obrazovatel'nykh dannykh [Educational data mining]. *Vestnik Tverskogo gosudarstvennogo universiteta. Ekonomika i upravlenie*, 2021, no. 3 (55), pp. 179-188. DOI: 10.26456/2219-1453/2021.3.179-188

3. Liñán L.C., Pérez Á.A.J. Education Data Mining and Learning Analytics: Differences, similarities and time evolution. *RUSC Universities and Knowledge Society Journal*, 2015, 12 (3), pp. 98-112, available at: <https://doi.org/10.7238/rusc.v12i3.2515> (accessed 24 February 2023).

4. Van Der Aalst W. Process mining: data science in action. Heidelberg: Springer, 2016, vol. 2.

5. Rosário N.A.L. Predicting academic performance - a practical study using Moodle log data and sociodemographic traits, 2022.

6. Romanov A.A., Volchek D.G. Analiz dannykh o povedenii pol'zovatelei v sistemakh elektronnoho obucheniia [User behavior data analysis in e-learning systems]. *Ontologiya proektirovaniia*, 2020, vol. 10, no. 1 (35), pp. 100-111. DOI: 10.18287/2223-9537-2020-10-1-100-111

7. Terbusheva E.A., Piotrovskaiia K.R. Analiticheskii potentsial platformy Moodle dlia monitoringa kachestva personifitsirovannogo obucheniia [Analytical potential of the LMS Moodle for monitoring the quality of personification]. *Obshchestvo. Kommunikatsiia. Obrazovanie*, 2021, vol. 12, no. 4, pp. 19-34. DOI: 10.18721/JHSS.1

8. Morozov A.I. Analitika obucheniia studentov pri realizatsii massovogo onlain-kursa [Analytics of students' education in the process of the implementation of Massive Open Online Course]. *Izvestiia Volgogradskogo gosudarstvennogo pedagogicheskogo universiteta*, 2022, no. 5 (168), pp. 16-22.

9. Alignment of teacher's plan and students' use of lms resources. analysis of moodle logs. *2016 15th international conference on information technology based higher education and training (ithet)*. IEEE, 2016, pp. 1-8.

10. The Python SQL Toolkit and Object Relational Mapper. Documentation, available at: <https://docs.sqlalchemy.org/en/20/> (accessed 24 February 2023).

11. Psycopg - PostgreSQL database adapter for Python, available at: <https://www.psycopg.org/docs/> (accessed 24 February 2023).

12. Soglashenie o sbore o tsifrovoi aktivnosti pol'zovatelei [User Digital Activity Collection Agreement]. Ofitsial'ny sait OOO "Elektronnoe izdatel'stvo Iurait", available at: <https://urait.ru/info/publication-data-policy> (accessed 24 February 2023).

13. Diulicheva Iu.Iu. Dataset dlia analiza russkoiazychnykh otzyvov na MOOK, izvlechenykh s platformy Stepik [Dataset for Analysis of Russian-Language Reviews on MOOCs Extracted from Stepik]. *Voprosy obrazovaniia*, 2022, no. 4, pp. 298-321.

14. Kaidalova L.V., Gumennikova Iu.V., Chernitsyna R.N. Statisticheskii analiz rezul'tatov testirovaniia po razdelu «Lineinaia algebra i analiticheskaia geometriia» v srede Moodle [Statistical analysis of testing

results by section «Linear algebra and analytical geometry» in the Moodle environment]. *Izvestiia Samarskogo nauchnogo tsentra Rossiiskoi akademii nauk. Sotsial'nye, gumanitarnye, mediko-biologicheskie nauki*, 2019, vol. 21, no. 65, pp. 35-39.

15. Batasova V.S. Analiz ob"ektivnosti rezul'tatov testirovaniia znaniia [Analysis of the objectivity of knowledge testing results]. *Informatizatsiia inzhenerenogo obrazovaniia*, 2020, pp. 27-30.

16. Uskov V.L. et al. Smart learning analytics: student academic performance data representation, processing and prediction. *Smart Education and e-Learning 2020*. Ed. V.L. Uskov, R.J. Howlett, L.C. Jain. Singapore: Springer, 2020, pp. 3-18.

17. Toktarova V.I. et al. Tekhnologii neironnykh setei dlia prognozirovaniia uspevaemosti obucheniia studentov v elektronnoi informatsionno-obrazovatel'noi srede vuza [Neural Network Technologies for Predicting Student Learning Achievement within eLearning Environment of the HEI]. *Tsifrovaia gumanitaristika i tekhnologii v obrazovanii (DHTE 2022)*, 2022, pp. 388-398.

18. Kustitskaia T.A. Baiesovskii podkhod k monitoringu effektivnosti otsenochnykh sredstv elektronного obrazovatel'nogo kursa [Bayesian approach to monitoring the effectiveness of evaluation tools of an electronic educational course]. *Distantcionnoe obuchenie v vysshem obrazovanii: opyt, problemy i perspektivy razvitiia*, 2020, pp. 164-165.

19. Agustianto K. et al. Design adaptive learning system using meta-cognitive strategy path for learning in classroom and intelligent tutoring systems. *AIP Conference Proceedings*. AIP Publishing LLC, 2016, vol. 1755, no. 1, 070012 p.

20. Popova Iu.B., Iatsynovich S.V. Primenenie iskusstvennoi neironnoi seti dlia opredeleniia stepeni usvoeniia uchebnogo materiala obuchaiushchimisia [The use of an artificial neural network to determine the degree of assimilation of educational material by students]. *Novye gorizonty - 2018. Sbornik materialov Belorusko-Kitaiskogo molodezhnogo innovatsionnogo foruma; 15-16 November 2018*. Minsk: Belorusskii natsional'nyi tekhnicheskii universitet, 2018, vol. 2, pp. 154-159.

21. Rusakov S.V., Rusakova O.L., Posokhina K.A. Neurosetevaia model' prognozirovaniia gruppy riska po uspevaemosti studentov pervogo kursa [A neural network model for predicting a risk group based on the progress of first-year students]. *Sovremennye informatsionnye tekhnologii i IT-obrazovanie*, 2018, no. 4, pp. 815-822.



22. Balaban I., Filipović D., Zlatović M. Post hoc identification of student groups: Combining user modeling with cluster analysis. *Education and Information Technologies*, 2022, pp. 1-26.

23. Martinez Abad F., Chaparro Caso López A.A. Data-mining techniques in detecting factors linked to academic achievement. *School Effectiveness and School Improvement*, 2017, vol. 28, no. 1, pp. 39-55.

24. Figueiredo M. et al. A data mining approach to study the impact of the methodology followed in chemistry lab classes on the weight attributed by the students to the lab work on learning and motivation. *Chemistry Education Research and Practice*, 2016, vol. 17, no. 1, pp. 156-171.

25. Aher S.B., Lobo L. Combination of machine learning algorithms for recommendation of courses in E-Learning System based on historical data. *Knowledge-Based Systems*, 2013, vol. 51, pp. 1-14.

26. Reddy S., Labutov I., Joachims T. Latent skill embedding for personalized lesson sequence recommendation: arXiv preprint arXiv:1602.07029, 2016.

27. Tsifirovoi repetitor Plario. Pervaia v Rossii onlain-sistema adaptivnogo obucheniia matematike [Digital tutor Plario. Russia's first online adaptive math learning system]. Ofitsial'nyi sait platformy plario.ru, available at: <https://plario.ru/> (accessed 24 February 2023).

28. Popova Iu.B. Ot LMS k adaptivnym obuchaiushchim sistemam [From LMS to adaptive trainings systems]. *Sistemnyi analiz i prikladnaia informatika*. Minsk: Belorusskii natsional'nyi tekhnicheskii universitet, 2019, no. 2, pp. 58-64.

29. Popova Iu.B. Intellektual'naia sostavliaiushchaia avtomatizirovannoi sistemy obucheniia cats [The intellectual component of the automated cats training system]. *Obrazovatel'nye tekhnologii i obshchestvo*. Kazan': Kazanskii natsional'nyi issledovatel'skii tekhnologicheskii universitet, 2019, vol. 22, no. 4, pp. 24-37.

30. Aldowah H., Al-Samarraie H., Fauzy W.M. Educational data mining and learning analytics for 21st century higher education: A review and synthesis. *Telematics and Informatics*, 2019, vol. 37, pp. 13-49.

31. LMS Data - Spring 2018 Updates edutechnica, available at: <https://edutechnica.com/2018/03/04/lms-data-spring-2018-updates/> (accessed 24 January 2023).

32. Moodle statistics. World Moodle registration map. Ofitsial'nyi sait "Moodle.org", available at: <https://stats.moodle.org/> (accessed 02 May 2023).

33. Golenkov V.V. *Intellectual'nye obuchaiushchie sistemy i virtual'nye uchebnye organizatsii [Smart educational systems]. Distantcionnoe obuchenie - obrazovatel'naia sreda XXI veka. Materialy mezhdunarodnoi nauchno-metodicheskoi konferentsii; Minsk, 18-20 December 2001.* Minsk: Bestprint, 2001, pp. 21-26.

### **Сведения об авторах**

**Бойко Александр Эдуардович** (Москва, Российская Федерация) – ассистент кафедры «Менеджмент и маркетинг» Российского химико-технологического университета им. Д.И. Менделеева (125047, Москва, Миусская площадь, 9, e-mail: boiko.a.e@muctr.ru).

**Савицкая Татьяна Вадимовна** (Москва, Российская Федерация) – доктор технических наук, профессор, профессор кафедры «Кибернетика химико-технологических процессов» Российского химико-технологического университета им. Д.И. Менделеева (125047, Москва, Миусская площадь, 9, e-mail: savitskaia.t.v@muctr.ru).

**Лопаткин Дмитрий Станиславович** (Москва, Российская Федерация) – кандидат экономических наук, доцент, заведующий кафедрой «Менеджмент и маркетинг» Российского химико-технологического университета им. Д.И. Менделеева (125047, Москва, Миусская площадь, 9, e-mail: lopatkin.d.s@muctr.ru).

### **About the authors**

**Alexander E. Boyko** (Moscow, Russian Federation) – Assistant of the Department of Management and Marketing of D. Mendeleev University of Chemical Technology of Russia (1250479, Moscow, 9, Miusskaya Square, e-mail: boiko.a.e@muctr.ru).

**Tatiana V. Savitskaya** (Moscow, Russian Federation) – Doctor of Technical Sciences, Professor Department of Cybernetics of Chemical and Technological Processes of the D. Mendeleev University of Chemical Technology of Russia (1250479, Moscow, 9, Miusskaya Square, e-mail: savitskaia.t.v@muctr.ru).

**Dmitry S. Lopatkin** (Moscow, Russian Federation) – Ph. D. in Economics Sciences, Associate Professor, Head of the Department of Management and Marketing of D. Mendeleev University of Chemical Technology of Russia (1250479, Moscow, 9, Miusskaya Square, e-mail: lopatkin.d.s@muctr.ru).

Поступила: 03.05.2023. Одобрена: 26.05.2023. Принята к публикации: 01.09.2023.

**Финансирование.** Исследование не имело спонсорской поддержки.

**Конфликт интересов.** Авторы заявляют об отсутствии конфликта интересов.

**Вклад авторов.** Все авторы сделали эквивалентный вклад в подготовку статьи.

Просьба ссылаться на эту статью в русскоязычных источниках следующим образом:

Бойко, А.Э. Применение инструментов интеллектуального анализа образовательных данных при работе с системами управления обучением в организациях высшего образования / А.Э. Бойко, Т.В. Савицкая, Д.С. Лопаткин // Вестник Пермского национального исследовательского политехнического университета. Электротехника, информационные технологии, системы управления. – 2023. – № 46. – С. 151–177. DOI: 10.15593/2224-9397/2023.2.07

Please cite this article in English as:

Boyko A.E., Savitskaya T.V., Lopatkin D.S. Educational data mining tools application in higher education learning management systems. *Perm National Research Polytechnic University Bulletin. Electrotechnics, information technologies, control systems*, 2023, no. 46, pp. 151-177. DOI: 10.15593/2224-9397/2023.2.07