

DOI: 10.32517/0234-0453-2022-37-4-54-63

АНАЛИЗ ОБРАЗОВАТЕЛЬНЫХ ДАННЫХ ВЗАИМОСВЯЗИ УСПЕШНОСТИ ОБУЧЕНИЯ И ПОВЕДЕНИЯ СТУДЕНТОВ В ЦИФРОВОЙ ОБРАЗОВАТЕЛЬНОЙ СРЕДЕ ВУЗА

В. И. Токтарова¹ ✉, О. Г. Попова¹¹ *Марийский государственный университет, г. Йошкар-Ола, Республика Марий Эл, Россия*

✉ toktarova@yandex.ru

Аннотация

Цифровизация является одним из ключевых факторов развития современной системы образования. Использование в образовательном процессе цифровых сквозных технологий ведет к накоплению большого количества данных о поведении и результатах обучения студентов в цифровой образовательной среде, анализ которых позволит повысить эффективность и качество их обучения. В статье проанализирована взаимосвязь между успешностью обучения и поведением студентов в цифровой образовательной среде вуза на основе методов образовательной аналитики. При проведении исследования были использованы теоретические, эмпирические и математические методы. Приведены определение образовательных данных, их целевое назначение. Описаны группы методов анализа образовательных данных, рассмотрены примеры применения моделей на их основе в образовательной сфере. Разработана и продемонстрирована модель взаимосвязи итоговой оценки студента и его активности / поведения при обучении на примере онлайн-курса по программированию, размещенного на платформе LMS Moodle. Анализ зависимости проводился как по отдельным показателям, так и по их совокупности. В результате были выделены действия студента, в наибольшей степени влияющие на итоговую оценку. Модель может быть применена для прогноза успешности обучения студентов в цифровой образовательной среде на основе их активности и образовательных результатов с целью повышения эффективности управления процессом обучения в вузах.

Ключевые слова: методы анализа данных, цифровая образовательная среда, анализ образовательных данных, регрессионный анализ, образовательные данные, студент, вуз.

Для цитирования:

Токтарова В. И., Попова О. Г. Анализ образовательных данных взаимосвязи успешности обучения и поведения студентов в цифровой образовательной среде вуза. *Информатика и образование*. 2022;37(4):54–63. DOI: 10.32517/0234-0453-2022-37-4-54-63

AN ANALYSIS OF EDUCATIONAL DATA ON THE CORRELATION BETWEEN LEARNING SUCCESS AND STUDENTS' BEHAVIOR IN THE UNIVERSITY DIGITAL EDUCATIONAL ENVIRONMENT

V. I. Toktarova¹ ✉, O. G. Popova¹¹ *Mari State University, Yoshkar-Ola, The Mari El Republic, Russia*

✉ toktarova@yandex.ru

Abstract

Digitization is one of the key factors in the development of the modern educational system. Using end-to-end digital technologies in the educational process leads to the accumulation of a large amount of data on students' behavior and learning outcomes in the digital educational environment. The analysis of these data will enable educators to improve the efficiency and quality of students' learning. The article analyzes the correlation between learning success and students' behavior in the university digital educational environment employing the methods of educational analytics. The research relied on theoretical, empirical, and mathematical methods. We provided a definition of educational data and explain their intended purpose. Next, we described the methods used to analyze educational data and gave examples of applying models based on them in the education sector. We developed and demonstrated a model of the correlation between students' final grades and their activity during university training using the example of an online programming course hosted on the Moodle LMS platform. Impact analysis was carried out both for individual indicators and for their combination. As a result, the research has identified a student's actions that create the most significant impact on the final score. The model can be used to predict students' learning success in the digital educational environment based on their activity and learning outcomes in order to improve the managerial effectiveness of the learning process at universities.

Keywords: data analysis methods, digital learning environment, educational data analysis, regression analysis, educational data, student, higher education institution.

For citation:

Toktarova V. I., Popova O. G. An analysis of educational data on the correlation between learning success and students' behavior in the university digital educational environment. *Informatics and Education*. 2022;37(4):54–63. (In Russian.) DOI: 10.32517/0234-0453-2022-37-4-54-63

1. Введение

Массовая цифровизация, происходящая в современном мире, коснулась и сферы образования. Суть цифровой трансформации состоит в эффективном и гибком применении новейших сквозных цифровых технологий в образовательной области с целью перехода к персонализированному и ориентированному на результат процессу обучения. Значительное число инициатив и планов относительно цифровизации образования нашло отражение в федеральном проекте «Кадры для цифровой экономики» национальной программы «Цифровая экономика Российской Федерации» [1], Стратегии развития информационного общества в Российской Федерации [2], согласно которой развитие дистанционных образовательных технологий и электронного обучения является обязательной составляющей формирования информационного пространства знаний.

В настоящее время для организации электронного обучения многие высшие учебные заведения используют системы управления обучением (Learning Management System, LMS) — платформы или программные приложения, предназначенные для интеграции инструментов обучения, администрирования, управления и распространения образовательных и информационных материалов, формирования аналитики и отчетности [3]. LMS предоставляют рабочее пространство для обучения, обмена информацией и коммуникации между участниками образовательного процесса. К преимуществам их использования можно отнести и накопление информации о поведении, траектории и результатах обучения как каждого студента, так и группы в целом.

Одной из популярных LMS, используемой в российских вузах, является Moodle (*Modular Object-Oriented Dynamic Learning Environment*) — модульная объектно-ориентированная динамическая среда обучения с открытым исходным кодом, используемая для создания веб-сайтов поддержки учебного процесса и дистанционного образования [4]. Основными преимуществами LMS Moodle являются: наличие встроенной системы разработки курсов, которая способна реализовывать большое количество педагогических технологий; формирование персонального пространства для каждого пользователя; реализация индивидуальной образовательной траектории; функциональность и открытость исходного кода; наличие подробной документации и др. Среда позволяет получить информацию для анализа образовательных данных: о времени доступа к курсу, длительности обучения на курсе / модуле / теме, сообщениях на форумах и в чатах, выполненных заданиях и др. [5].

Образовательные данные — это информация об учебных результатах обучающихся, развитии образовательных систем, условиях образования (образовательных программах, особенностях образовательной среды, образовательных методиках и др.) [6].

Все участники образовательного процесса заинтересованы в его максимальной эффективности [7]. Обучающиеся стремятся улучшить свою успеваемость, для этого им нужно получать рекомендации и своевременную обратную связь по усвоению учебных дисциплин, построению индивидуальных образовательных траекторий. Преподавателям важно понимание поведенческих, когнитивных, социальных аспектов субъектов учебного процесса для выбора и применения определенных методов и приемов обучения. Развитие электронного обучения предъявляет новые требования к организации и повышению качества учебно-методического процесса.

Образовательные данные могут применяться для принятия организационно-педагогических и управленческих решений в образовании. Для этого необходимы сбор и анализ большого количества данных об образовательном процессе на всех этапах обучения [8]. Анализ образовательных данных (*Educational Data Mining*) — это направление исследований, которое связано с применением методов интеллектуального анализа данных, машинного обучения и статистики к информации, производимой образовательными учреждениями. Анализ образовательных данных (АОД) позволяет выявить результативность обучения студентов на каждом этапе и сделать его эффективнее. Выделяют следующие цели анализа образовательных данных [9]:

- прогнозирование поведения обучающегося посредством создания модели обучающегося, включающей информацию о его знаниях, способностях, поведении, мотивации и других индивидуальных характеристиках;
- разработка новых способов представления учебной информации, соответствующих выявленным познавательным способностям студентов;
- изучение взаимодействия «преподаватель — студент», обеспечение обратной связи, рекомендаций и др.

Большое количество работ в области анализа образовательных данных посвящено анализу логов (журналов регистрации действий пользователя) LMS для поиска взаимосвязи между показателями, чаще всего влияния различных показателей на успеваемость. На онлайн-курсах студенты сами управляют своим временем, выполняя задания тогда, когда им удобно. При этом они самостоятельно должны организовать учебный процесс так, чтобы выполнить задания вовремя и при этом максимально усвоить учебный материал, что, к сожалению, эффективно получается не у всех студентов. В этом случае по результатам анализа данных журнала LMS можно судить о причинах неудачного или успешного прохождения курса студентом и прогнозировать его дальнейшее поведение. Такие данные могут быть использованы как для поддержки студентов в ходе их обучения, так и для корректировки учебного курса.

Доказано, что поведение студентов в рамках обучения на онлайн-курсах влияет на результаты обу-

чения. Так, в работе [10] авторами была вычислена корреляция между поведением студентов в электронной образовательной среде и его эффективностью. В работе использованы такие показатели, как пол и возраст студентов, курс, специальность, уровень подготовки, фактическое время обучения, количество публикаций, оценки за домашние задания и за экзамен, итоговая оценка. На основе методов корреляции авторы пришли к заключению, что каждое действие обучающихся влияет на эффективность их обучения. Важным фактором повышения результативности обучения на рассматриваемых онлайн-курсах стало выполнение домашнего задания, поэтому были предложены рекомендации уделить большее внимание самостоятельной работе, степени сложности заданий и контролю за их выполнением.

В работе [11] исследователи группировали студентов по продолжительности активности в электронной среде и по типам действий, учитывались поведение слушателей на онлайн-курсе, их успеваемость, завершенность / незавершенность курса. К данным о поведении на курсе относились вид используемого учебного ресурса (лекция, видеолекция), выполнение практических заданий, посещение форума и др. Для поиска модели, учитывающей длительность активности, применялись методы кластеризации. Выяснилось, что у обучающихся, успешно освоивших курс, наблюдались более частые и продолжительные сеансы работы на протяжении всего курса. Кроме того, было обнаружено, что распределение сеансов является значимым показателем наличия у студентов сформированных навыков, связанных с управлением временем.

Авторы исследования [12] прогнозировали успеваемость студентов в смешанной учебной среде, основываясь на следующих показателях: общее проведенное время в системе, частота входа в учебную среду, регулярность интервала посещения, количество посещений, количество публикаций на форумах. Анализ был проведен при помощи множественного линейного регрессионного анализа и метода Random Forest. Была доказана зависимость используемых моделей от типа смешанного обучения. В случае курса смешанного обучения, основанного на активных обсуждениях субъектов обучения в форумах и чатах, достижения студентов можно трактовать с помощью модели линейного множественного регрессионного анализа. Но в случае курса смешанного обучения, построенного на основе лекций (ознакомление с теоретическим материалом, самостоятельное выполнение и загрузка заданий), модель линейного множественного регрессионного анализа неприемлема для прогнозирования. Исследователи пришли к выводу, что модели прогнозирования должны строиться в зависимости от типов смешанного обучения.

Таким образом, цель нашего исследования — выявление взаимосвязей между успешностью обучения и поведением студентов в цифровой образовательной среде вуза на основе множественной регрессионной модели.

2. Методы

Методологическую базу исследования составляет комплекс методов: теоретические (анализ нормативных источников, системный, структурно-функциональный и сравнительно-сопоставительный анализ; контент-анализ, синтез, систематизация, дедукция, индукция, прогнозирование); эмпирические (наблюдение, анкетирование, тестирование, экспертная оценка, педагогический эксперимент); математические (корреляционный анализ, регрессионный анализ, методы математической статистики, математическое моделирование, Data Mining).

Основные методы анализа образовательных данных можно разделить на три группы [13]:

- *методы прогнозирования*: построение модели, предсказывающей значение зависимой переменной по значениям независимых переменных. К таким методам относятся алгоритмы классификации (если выходная переменная имеет конечный дискретный набор значений), регрессии (если выходная переменная имеет непрерывные значения). Например, исследуя доступ к учебным материалам и результатам студентов, можно предсказать, какую оценку будет иметь другой студент со схожей активностью на курсе;
- *методы обнаружения структуры*, к которым относятся алгоритмы кластеризации. Кластеризация является продолжением идеи классификации, но ее особенность состоит в том, что классы изначально не определены. К примеру, в результате кластеризации студенты могут быть объединены в группы по каким-либо схожим признакам;
- *методы выявления взаимосвязей*: построение модели, устанавливающей взаимосвязи между переменными. К данной группе методов относятся корреляционный анализ, поиск связующих правил и последовательных шаблонов. Например, можно выявить связь между посещением онлайн-занятий и успешным прохождением курса.

Исследованиям в области анализа данных посвящены работы многих зарубежных и отечественных ученых. Среди них С. И. Барцев, П. П. Белоножко, Н. Вапник, П. Дж. Вербос, А. И. Галушкин, А. Н. Горбань, Ю. И. Журавлев, А. П. Карпенко, М. Мински, А. Б. Новиков, В. А. Охонин, С. Пайперт, Н. И. Потанин, К. В. Рудаков, В. Н. Сосницкий, Б. Уидроу, О. А. Фиофанова, Дж. Хопфилд и др. Большой вклад в разработку подходов к решению проблем анализа образовательных данных внесли Р. Бейкер, П. Л. Брусиловский, С. Вентура, К. А. Вилкова, У. С. Захарова, Л. С. Лисицына, Х. Ромеро, И. Д. Рудинский и др. Анализу успеваемости при использовании цифровой образовательной среды посвящены работы Л. В. Легашева, Д. И. Парфенова, А. Б. Шухмана, J. H. Kim, L. Liu, H. Ma, Y. Park, J. Song, J. Yao и др.

3. Результаты

Переход образования в цифровой формат привел к существенным изменениям организации образовательного процесса [19], где одним из факторов успешного обучения студента является управление образовательным процессом: мониторинг работы студентов в цифровой образовательной среде и электронных образовательных платформах, корректировка траекторий обучения на основе анализа цифрового следа и др. [20].

В рамках проведенного экспериментального обучения нами были использованы данные о действиях и результатах обучения студентов Института цифровых технологий (ИЦТ) по курсу «Программирование», размещенному в электронной информационно-образовательной среде вуза (<https://elearning.marsu.ru/>) (рис. 1). Выбор данной дисциплины обосновывается тем, что программирование является базовым курсом для обучения студентов направлений подготовки ИЦТ. Эффективность и качество усвоения учебного материала курса влияют на освоение студентами программы последующих дисциплин, а также их профессиональную успешность в дальнейшем. На основе результатов анализа выбранного курса можно прогнозировать успешность обучения студентов на курсах и по другим дисциплинам со схожей структурой.

Для изучения взаимосвязи между итоговой оценкой по курсу и поведением студента на курсе были использованы следующие переменные: идентификационный номер студента, итоговая оценка, количество обращений к форуму, количество переходов к чату, общее время, проведенное в системе (в мин), количество кликов по теории каждой из пяти тем, количество кликов по каждому заданию, оценки за все задания. Количество кликов по форуму и чату были объединены в один показатель, так как на данном курсе названные элементы выполняли идентичную роль. Всего в эксперименте приняли участие 231 студент.

Процесс анализа образовательных данных (АОД) состоял из следующих этапов [9]:

- 1) предварительная обработка данных;
- 2) выявление закономерностей в данных;
- 3) проверка обнаруженных закономерностей;
- 4) применение найденных закономерностей для прогнозирования;
- 5) использование полученных прогнозов для принятия решений и выработки образовательной политики.

На первом этапе АОД была проведена предварительная обработка данных, после которой в рассмотрении остались показатели 226 студентов. Были рассчитаны их описательные статистики (табл.).

Здесь основные действия студента — это клики по элементам курса, т. е. переходы к определенному разделу курса в электронной среде. Так, клики по теории — открытие студентом онлайн-лекции, содержащей теорию по определенной теме. Студент может обращаться к лекции как в начале изучения темы, так и при выполнении практического задания. Клик по заданию — переход к элементу курса, который содержит учебное задание по теме. Например, студент может открыть задание для ознакомления, выполнения и / или последующей самопроверки. Всего в данном курсе пять тем. Количество кликов по теории и заданиям каждой темы рассмотрено отдельно. Из таблицы видно, что в среднем каждый студент провел в системе 420,99 мин. За задания по темам 2 и 5 нет оценки «2» ни у одного студента, при этом количество кликов по заданию и теории по этим темам существенно не отличается от количества кликов по другим темам (в среднем 7–11 раз студенты открывали лекцию по каждой теме и 8–10 раз переходили к заданию).

Визуализация данных позволяет сделать быстрый обзор показателей, лежащих в основе разведочного анализа. Из данной матрицы корреляции признаков (рис. 2) видно, что корреляция между признаками существует. Наиболее сильное влияние на итоговую оценку оказывают оценки за задания

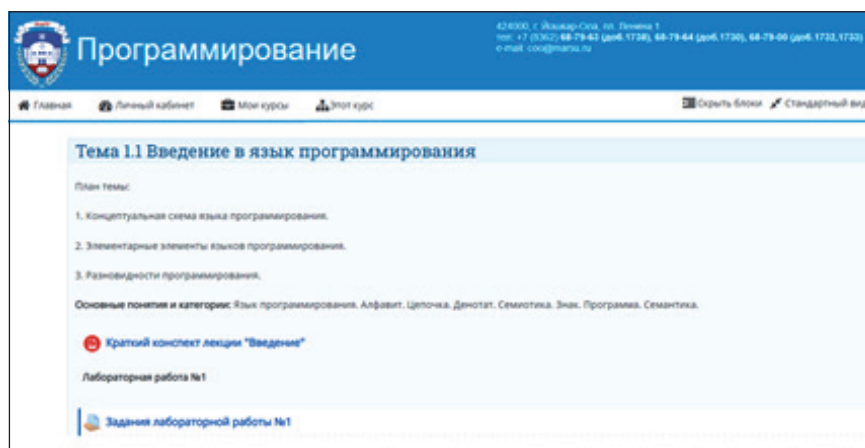


Рис. 1. Фрагмент курса «Программирование», размещенного в электронной информационно-образовательной среде вуза

Fig. 1. Fragment of the “Programming” course posted in the electronic information educational environment of the university

Описательные статистики

Descriptive statistics

	Среднее	Стандартное отклонение	Минимальное	25 %	50 %	75 %	Максимальное
Итоговая оценка	4.319	0.680	2.800	4.000	4.600	4.800	5.000
Клики по форуму и чату	6.628	3.610	0.000	4.000	7.000	10.000	17.000
Время в системе (в мин)	420.987	191.879	123.300	216.600	428.900	568.400	852.100
Клики по теории темы 1	10.522	4.101	1.000	8.000	10.000	14.000	25.000
Клики по заданию темы 1	8.487	5.702	1.000	3.000	8.000	14.000	20.000
Оценка по теме 1	4.310	0.778	2.000	4.000	4.000	5.000	5.000
Клики по теории темы 2	7.788	5.697	1.000	4.000	6.000	9.000	25.000
Клики по заданию темы 2	8.310	5.217	1.000	4.000	8.000	11.000	30.000
Оценка по теме 2	4.345	0.727	3.000	4.000	4.000	5.000	5.000
Клики по теории темы 3	9.310	4.841	0.000	6.000	8.000	11.000	25.000
Клики по заданию темы 3	7.372	3.949	1.000	4.000	7.000	10.000	18.000
Оценка по теме 3	4.310	0.778	2.000	4.000	4.000	5.000	5.000
Клики по теории темы 4	7.929	3.766	1.000	6.000	8.000	9.000	18.000
Клики по заданию темы 4	8.442	4.026	2.000	5.000	9.000	10.000	18.000
Оценка по теме 4	4.265	0.789	2.000	4.000	4.000	5.000	5.000
Клики по теории темы 5	9.814	3.527	3.000	7.000	9.000	13.000	17.000
Клики по заданию темы 5	10.106	4.115	2.000	7.000	10.000	12.000	22.000
Оценка по теме 5	4.363	0.767	3.000	4.000	5.000	5.000	5.000

и общее время, проведенное на курсе. Существует заметная корреляция между итоговой оценкой и кликами по заданию 4 и по теории темы 5. Умеренная корреляция существует между итоговой оценкой и следующими показателями: количество кликов по форуму и чату, количество кликов по заданию 1, количество кликов по теории темы 2, количество кликов по заданию 3, количество кликов по теории 4 и заданию 5. Между количеством кликов по заданию 2 и итоговой оценкой корреляция слабая.

Существует также сильная корреляция между временем, проведенным в системе, и оценками за каждое задание. Наиболее сильно коррелирующие переменные были вынесены в отдельную матрицу (рис. 3).

Существует заметная корреляция (0.65) между количеством кликов по заданию темы 4 и соответствующей оценкой, а также между количеством кликов по теории темы 5 и оценкой по данной теме (0.54). Между остальными показателями количества кликов по теории и заданиям по определенным темам и полученными оценками также существует умеренная или слабая связь. Причем в большинстве случаев многократный доступ к теории имеет боль-

шую связь с итоговой оценкой, чем изучение текста задания.

Оценка наиболее востребованных элементов курса на основе анализа количества кликов показала их несущественное распределение: меньше всего пользователи обращались к заданию по теме 3 и к теории по теме 2, а больше всего — к теории по теме 1 и к теории и заданию по теме 5. Как видно из матрицы корреляции (рис. 2), итоговая оценка также имеет наибольшую зависимость от оценки по теме 5. Можно судить о том, что данная тема показала студентам наиболее сложной, или о том, что они наиболее ответственно подошли к выполнению последнего задания, чтобы увеличить свой итоговый балл. Количество кликов по всем заданиям и всей теории примерно одинаковое, однако количество кликов по теории незначительно превышает количество кликов по заданиям.

Рассмотрим среднее время в минутах, которое проводили студенты, получившие оценки «удовлетворительно», «хорошо» и «отлично» (рис. 4).

Из гистограммы видно, что те, кто получил оценку «отлично», почти в два раза больше находились на курсе, чем те, кто получил оценку «хорошо»,

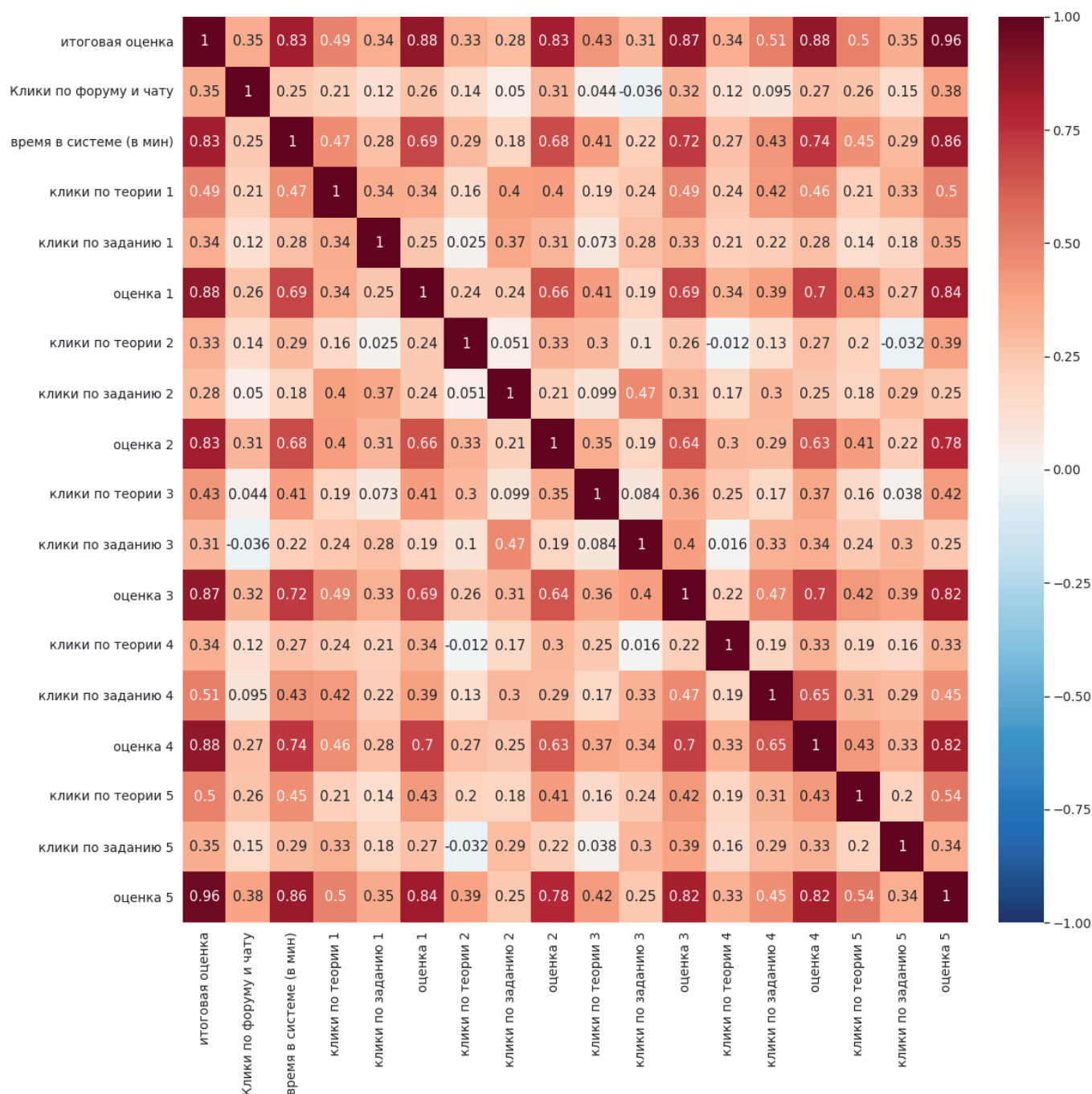


Рис. 2. Матрица корреляции признаков

Fig. 2. Feature correlation matrix

и в 3,4 раза больше, чем студенты с оценкой «удовлетворительно». Студенты с оценками «хорошо» в 1,75 раза больше времени провели в системе, чем студенты, получившие оценку «удовлетворительно». Таким образом, становится ясно, что чем дольше студент изучает учебный материал и выполняет задание, тем выше его балл.

В исследовании рассматривалась и активность посещения форума и чата (рис. 5).

Доля тех, кто заходил на форум и чат курса один раз или менее, составила 9,73 %. Доля тех, кто посещал форум или чат более одного раза, — 90,27 %.

Из гистограммы видно, что среди отличников больше тех, кто посещал форум более одного раза. Чем выше оценка, тем чаще студент открывал сообщения форума и чата. Большинство из тех, кто зашел на форум всего один раз или не посетил его совсем, получили оценку «удовлетворительно», и меньше всего среди них получивших «хорошо». Это объясняет тот факт, что между данным показателем и итоговой оценкой корреляция средняя (0.35).

Для выявления закономерностей данных и для изучения взаимосвязи между итоговой оценкой студента и его пользовательским поведением на курсе

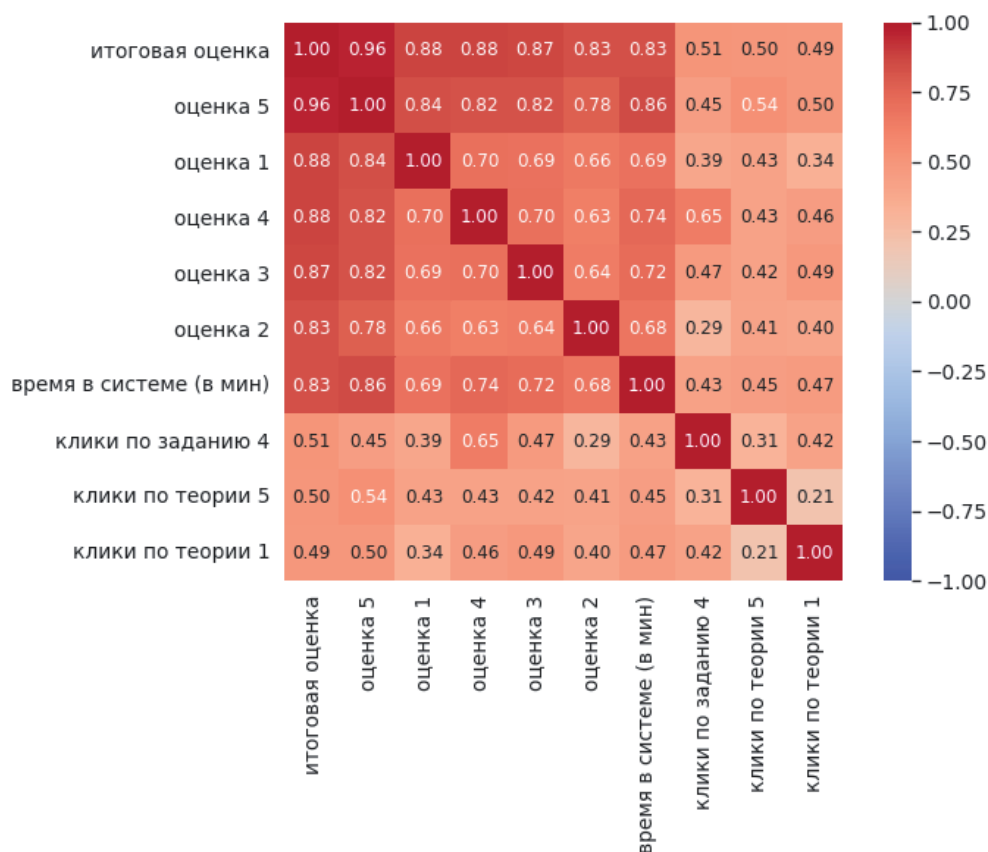


Рис. 3. Матрица корреляции наиболее сильно зависимых переменных

Fig. 3. Correlation matrix of the most strongly dependent variables

в цифровой образовательной среде была проведена множественная регрессия. Оценки по каждой из тем при этом не учитывались. Были использованы следующие факторные признаки:

x_1 — общее время, проведенное пользователем в системе;

x_2 — количество кликов по форуму и чату;

x_3 — количество кликов по теории темы 1;

x_4 — количество кликов по заданию темы 1;

x_5 — количество кликов по теории темы 2;

x_6 — количество кликов по заданию темы 2;

x_7 — количество кликов по теории темы 3;

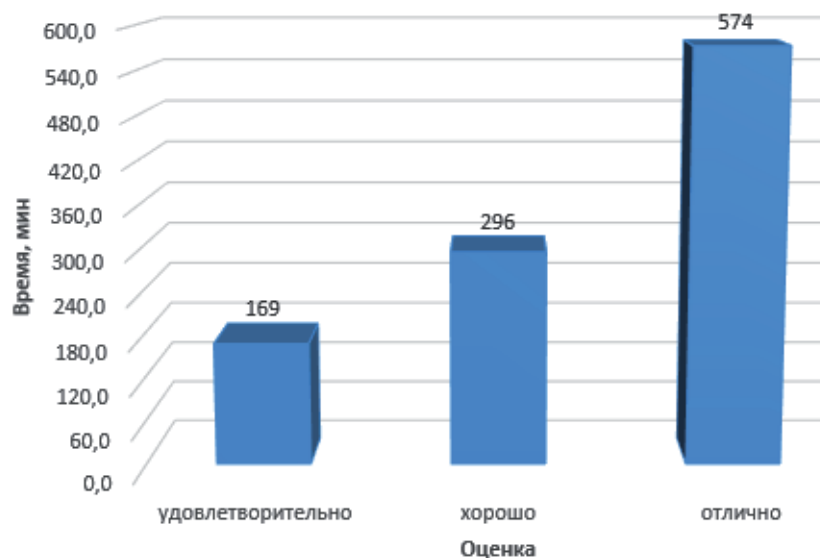


Рис. 4. Итоговая оценка и среднее время, проведенное в электронной информационно-образовательной среде

Fig. 4. Final grade and average time spent in the electronic information educational environment

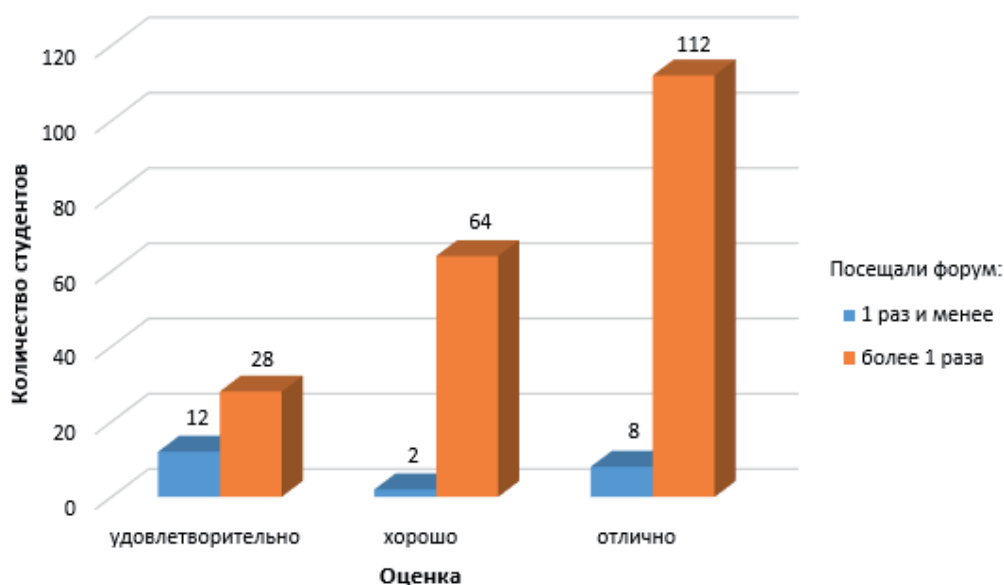


Рис. 5. Зависимость оценки от посещения форума и чата

Fig. 5. Dependence of the grade on the forum and chat attendance

x_8 — количество кликов по заданию темы 3;

x_9 — количество кликов по теории темы 4;

x_{10} — количество кликов по заданию темы 4;

x_{11} — количество кликов по теории темы 5;

x_{12} — количество кликов по заданию темы 5.

Данные были разделены на обучающую и тестовую выборки по 80 % и 20 % соответственно.

Прежде чем приступить к построению регрессионной модели, данные были стандартизованы, а также исключено наличие межфакторной корреляции. Было выявлено, что между факторами существует коллинеарность. Для избавления от

мультиколлинеарности был использован метод исключения, в основе которого лежит пошаговое исключение мультиколлинеарных признаков. После избавления от мультиколлинеарности было оставлено пять факторных признаков (рис. 6).

Значимой является t-статистика каждого из факторных признаков, входящих в модель. Это факторы, обозначающие время, проведенное в системе; количество кликов по форуму и чату; количество кликов по теории тем 2 и 4; клики по заданиям тем 4 и 5. $F_{\text{набл}} > F_{\text{табл}}$, следовательно, уравнение в целом является статистически значимым и надежным.

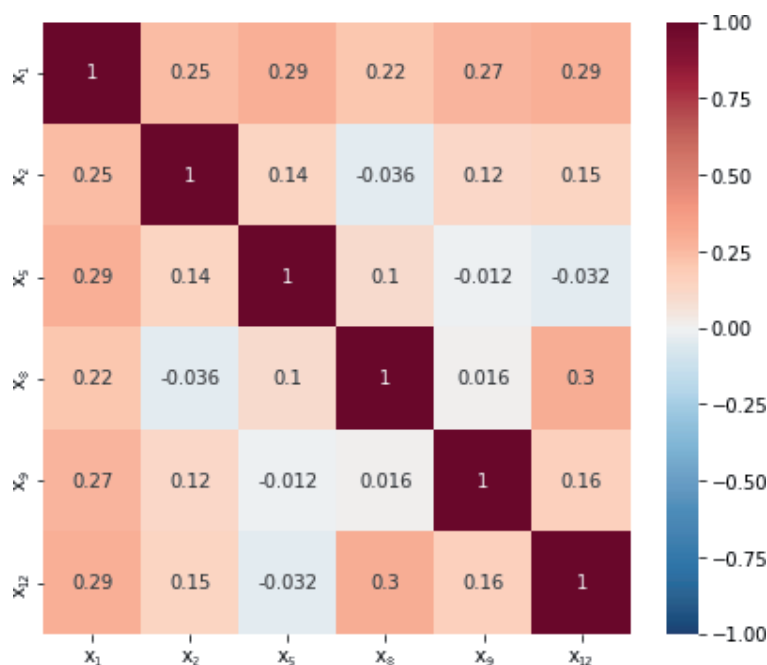


Рис. 6. Переменные, входящие в модель

Fig. 6. Variables included in the model

Уравнение множественной регрессии имеет следующий вид:

$$y = 2,6020 + 0,0025x_1 + 0,0236x_2 + 0,0107x_5 + 0,0163x_8 + 0,0274x_9 + 0,0128x_{12}.$$

На обучающих данных коэффициент детерминации R^2 составил 0,760. Данное значение больше 0,7, поэтому модель можно считать качественной. Между факторными переменными и результативной переменной существует линейная связь. Доля дисперсии зависимой переменной (итоговая оценка), объясняемая рассматриваемой моделью, составляет 76 %.

Для проверки обнаруженных закономерностей модель была применена к тестовым данным. Коэффициент детерминации R^2 на тестовых данных незначительно ниже коэффициента на обучающих данных и составляет 0,749.

Средняя квадратическая ошибка MSE на обучающей выборке составила 0,109, на тестовой — 0,123, что подтверждает качество модели.

4. Обсуждение и выводы

Исследование показало, что между поведением пользователя в рамках обучения на онлайн-курсе в цифровой образовательной среде и его итоговой оценкой существует линейная зависимость. Оценки за задания по каждой из тем оказывают практически одинаковое влияние на итоговую оценку.

В соответствии с анализом количество времени, посвященное изучению учебного материала, а также количество обращений к форуму и чату курса, на которых студент может получить ответы на возникающие в ходе выполнения заданий вопросы или рекомендации по их выполнению, имеют высокую значимость и влияют на успешность прохождения студентом электронного курса.

Исходя из этих выводов, студентам могут быть даны рекомендации о том, что наибольшее внимание стоит уделить изучению темы 2, внимательнее выполнить задания тем 3–5, так как именно от данных элементов курса в наибольшей степени зависит итоговая оценка. Стоит также уделить особое внимание активности на форуме и в чате и больше времени проводить с учебным материалом для его качественного усвоения.

Использование анализа образовательных данных может дать рекомендации студентам по изучению тех или иных тем в ходе прохождения курса. Полученная модель может применяться для дальнейшего прогнозирования эффективности обучения студентов на курсе «Программирование», а также на курсах со схожей структурой после их анализа.

К примеру, студент с показателями $x_1 = 190$, $x_2 = 2$, $x_5 = 6$, $x_8 = 5$, $x_9 = 5$, $x_{12} = 6$ согласно модели будет иметь итоговую оценку 3,8. Однако если время изучения курса составит не 190, а 300 мин, его предсказываемая оценка будет составлять 4,1. Стоит отметить, что итоговая оценка складывается

из промежуточных оценок и прогноз, построенный моделью, является только рекомендацией для студента по наиболее эффективному изучению курса.

Таким образом, современные инструменты анализа данных позволяют строить модели прогнозирования успешности обучения студентов в цифровой образовательной среде на основе их активности и образовательных результатов для наиболее эффективного управления процессом обучения в высших учебных заведениях.

Список источников / References

1. Кадры для цифровой экономики: Министерство цифрового развития, связи и массовых коммуникаций Российской Федерации. Режим доступа: <https://digital.gov.ru/ru/activity/directions/866/>
[Personnel for the digital economy: Ministry of Digital Development, Communications and Mass Media of the Russian Federation (In Russian.) Available at: <https://digital.gov.ru/ru/activity/directions/866/>]
2. Указ Президента Российской Федерации от 09.05.2017 № 203: «О Стратегии развития информационного общества в Российской Федерации на 2017–2030 годы». Режим доступа: http://www.consultant.ru/document/cons_doc_LAW_216363/
[Decree of the President of the Russian Federation No. 203 dated 09.05.2017: “On the Strategy for the Development of the Information Society in the Russian Federation for 2017–2030”. (In Russian.) Available at: http://www.consultant.ru/document/cons_doc_LAW_216363/]
3. Система управления обучением. Словарь-справочник. АНО ДПО «Корпоративный университет Сбербанка», 2012–2022. Режим доступа: <https://sberuniversity.ru/edutech-club/glossary/904/>
[Learning management system. ANO FVE “Corporate University of Sberbank”, 2012–2022. (In Russian.) Available at: <https://sberuniversity.ru/edutech-club/glossary/904/>]
4. Moodle — Open-source learning platform. Available at: <https://moodle.com/>
5. Мухаметшин Л. М., Салехова Л. Л., Мухаметшина М. М. Использование системы LMS Moodle в современном образовательном процессе. *Филология и культура*. 2019;(2(56)):274–279.
[Mukhametshin L. M., Salekhova L. L., Mukhametshina M. M. Using the LMS Moodle system in the modern educational process. *Philology and Culture*. 2019;(2(56)):274–279. (In Russian.)]
6. Фиофанова О. А. Smart Big Data в публичных докладах. *Образовательная политика*. 2020;(4(84)):70–77. DOI: 10.22394/2078-838X-2020-4-70-77
[Fiofanova O. A. Smart Big Data in public reports. *Educational Policy*. 2020;(4(84)):70–77. (In Russian.) DOI: 10.22394/2078-838X-2020-4-70-77]
7. Toktarova V. I. Pedagogical management of learning activities of students in the electronic educational environment of the university: A differentiated approach. *International Education Studies*. 2015;8(5):205–212. DOI: 10.5539/ies.v8n5p205
8. Есин Т. Е., Глухих И. Н. Использование методов интеллектуального анализа данных для получения знаний об образовательном процессе. *Математическое и информационное моделирование: сборник научных трудов*. Тюмень: Тюменский государственный университет; 2017;15(1):139–149.
[Esin T. E., Glukhikh I. N. Using data mining methods to obtain knowledge about the educational process. *Mathematical and Information Modeling: A Collection of Scientific Papers*. Tyumen: Tyumen State University; 2017;15(1):139–149. (In Russian.)]

9. Белоножко П. П., Карпенко А. П., Храмов Д. А. Анализ образовательных данных: направления и перспективы применения. *Интернет-журнал «Наукоедение»*. 2017;9(4):57.

[Belonozhko P. P., Karpenko A. P., Khramov D. A. Algorithm of rangovy optimization of access to data in information system. *Electronic Scientific Publication of "Naukovedenie"*. 2017;9(4):57. (In Russian.)]

10. Ma H. Y., Yao J., Liu L. W. Research on the correlation between learning effectiveness and online learning behavior based on online education scene. *Creative Education*. 2017;(8(13)):2187–2198. DOI: 10.4236/ce.2017.813149

11. de Barba P. G. et al. The importance and meaning of session behaviour in a MOOC. *Computers & Education*. 2020;(146).

12. Kim J. H., Park Y., Song J., Jo I.-H. Predicting students' learning performance by using online behavior patterns in blended learning environments: Comparison of two cases on linear and non-linear model. *The 10th International Conference on Educational Data Mining*; 2019;(168):69–78.

13. Фиофанова О. А. Анализ больших данных в сфере образования: методология и технологии: монография. М.: Издательский дом «Дело»; 2020. 200 с.

[Fiofanova O. A. Big data analysis in the field of education: Methodology and technologies: monograph. Moscow, Delo Publishing House; 2020, 200 p. (In Russian.)]

14. Сосницкий В. Н., Потанин Н. И., Шевелева Л. В. Проблемы статистического анализа средней успеваемости студентов. *Фундаментальные исследования*. 2013;(10-2):316–320.

[Sosnitskiy V. N., Potanin N. I., Sheveleva L. V. Problems statistical analysis middle evaluation of students. *Fundamental Research*. 2013;(10-2):316–320. (In Russian.)]

15. Baker R. Data mining for education. *International Encyclopedia of Education*. 2010;(7):112–118.

16. Луцицына Л. С., Першин А. А., Усков В. Л. К вопросу повышения результативности массового онлайн-курса. *Научно-технический вестник информационных технологий, механики и оптики*. 2014;(5(93)):164–171.

[Lisitsyna L. S., Pershin A. A., Uskov V. L. New approaches to efficiency of massive online course. *Scientific and Technical Journal of Information Technologies, Mechanics and Optics*. 2014;(5(93)):164–171. (In Russian.)]

17. Вилкова К. А., Захарова У. С. Учебная аналитика в традиционном образовании: ее роль и результаты. *Университетское управление: практика и анализ*. 2020;24(3):59–76. DOI: 10.15826/umpa.2020.03.026

[Vilkova K. A., Zakharova U. S. Learning analytics in conventional education: Its role and outcomes. *Journal University Management: Practice and Analysis*. 2020;24(3):59–76. (In Russian.) DOI: 10.15826/umpa.2020.03.026]

18. Шухман А. Е., Парфенов Д. И., Легашев Л. В., Гришина Л. С. Анализ и прогнозирование успеваемости обучающихся при использовании цифровой образовательной среды. *Высшее образование в России*. 2021;30(8-9):125–133. DOI: 10.31992/0869-3617-2021-30-8-9-125-133

[Shukhman A. E., Parfenov D. I., Legashev L. V., Grishina L. S. Analysis and forecasting students' academic performance using a digital educational environment. *Higher Education in Russia*. 2021;30(8-9):125–133. (In Russian.) DOI: 10.31992/0869-3617-2021-30-8-9-125-133]

19. Токтарова В. И., Федорова С. Н. Информационно-образовательная среда вуза: интерпретационный и содержательный анализ. *Вестник Марийского государственного университета*. 2018;12(4):77–87. DOI: 10.30914/2072-6783-2018-12-4-77-87]

[Toktarova V. I., Fedorova S. N. Electronic educational environment of higher education institution: Interpretative and content analysis. *Bulletin of the Mari State University*. 2018;12(4):77–87. (In Russian.) DOI: 10.30914/2072-6783-2018-12-4-77-87]

20. Чикилева Л. С. Оптимизация учебного процесса в условиях цифровой образовательной среды как предиктор успешности дистанционного обучения. *Российский гуманитарный журнал*. 2021;10(1):42–53. DOI: 10.15643/libartrus-2021.1.4

[Chikileva L. S. Optimization of the educational process in the digital educational environment as a predictor of successful distance learning. *Liberal Arts in Russia*. 2021;10(1):42–53. (In Russian.) DOI: 10.15643/libartrus-2021.1.4]

Информация об авторах

Токтарова Вера Ивановна, доктор пед. наук, доцент, профессор кафедры прикладной математики и информатики, Марийский государственный университет, г. Йошкар-Ола, Республика Марий Эл, Россия; ORCID: <https://orcid.org/0000-0002-3590-3053>; e-mail: toktarova@yandex.ru

Попова Олеся Геннадьевна, магистрант 2-го курса, Институт цифровых технологий, Марийский государственный университет, г. Йошкар-Ола, Республика Марий Эл, Россия; ORCID: <https://orcid.org/0000-0002-2666-1005>; e-mail: olesya_popova10@mail.ru

Information about the authors

Vera I. Toktarova, Doctor of Sciences (Education), Docent, Professor at the Department of Applied Mathematics and Informatics, Institute of Digital Technologies, Mari State University, Yoshkar-Ola, The Mari El Republic, Russia; ORCID: <https://orcid.org/0000-0002-3590-3053>; e-mail: toktarova@yandex.ru

Olesya G. Popova, 2nd year master's student at Institute of Digital Technologies, Mari State University, Yoshkar-Ola, The Mari El Republic, Russia; ORCID: <https://orcid.org/0000-0002-2666-1005>; e-mail: olesya_popova10@mail.ru

Поступила в редакцию / Received: 17.06.2022.

Поступила после рецензирования / Revised: 04.08.2022

Принята к печати / Accepted: 09.08.2022.