

Учебная аналитика MOOK как инструмент прогнозирования успешности обучающихся

**Т. Ю. Быстрова, В. А. Ларионова, Е. В. Сеницын,
А. В. Толмачев**

Быстрова Татьяна Юрьевна

доктор философских наук, профессор
Уральского гуманитарного института.
E-mail: taby27@yandex.ru

Ларионова Виола Анатольевна

кандидат физико-математических
наук, доцент, заместитель проректора,
заведующий кафедрой Высшей школы
экономики и менеджмента. E-mail:
v. a. larioнова@urfu.ru

Сеницын Евгений Валентинович

доктор физико-математических наук,
профессор Высшей школы экономики
и менеджмента. E-mail: e. v. sinit-
syn@urfu.ru.

Толмачев Александр Владимирович

старший преподаватель Высшей школы
экономики и менеджмента. E-mail:
avtolmachev@urfu.ru

ФГАОУ ВО «Уральский федеральный
университет имени первого Прези-
дента России Б. Н. Ельцина». Адрес:
620002, Екатеринбург, ул. Мира, 19.

Аннотация. Авторы предлагают ис-
пользовать данные учебной аналити-
ки массовых открытых онлайн-курсов
в качестве основы для прогнозирова-
ния успешности обучающихся. Такое
прогнозирование актуально в усло-

виях запроса на адаптивное обучение,
к которому сегодня приходит высшая
школа. Междисциплинарная методо-
логия статьи позволяет интерпрети-
ровать эмпирические количествен-
ные данные о выполнении отдельных
видов заданий онлайн-курса для со-
ставления прогноза успеваемости
обучающихся, с одной стороны, и кор-
ректировки качества онлайн-курсов —
с другой. Результаты учебной анали-
тики дают возможность максималь-
но учитывать особенности работы
обучающихся с информацией и сте-
пень предварительной подготовлен-
ности к курсу. Приводятся результаты
применения предлагаемого алгорит-
ма учебной аналитики для анали-
за успешности обучающихся на кон-
кретных онлайн-курсах, созданных
в Уральском федеральном универси-
тете и размещенных на Национальной
платформе открытого образования.

Ключевые слова: массовые откры-
тые онлайн-курсы, учебная аналитика,
эмпирические данные, онлайн-обуче-
ние, контрольные мероприятия, мони-
торинг успеваемости.

DOI: 10.17323/1814-9545-2018-4-139-166

Статья поступила
в редакцию
в сентябре 2018 г.

Работа выполнена
при финансовой под-
держке, предостав-
ленной согласно По-
становлению № 211
Правительства Рос-
сийской Федера-
ции, контракт № 02.
А03.21.0006.

Появление и быстрый рост числа массовых открытых онлайн-
курсов (MOOK) на мировом рынке образования [Семенова,
Вилкова, Щеглова, 2018] привели к тому, что технологии он-

лайн-обучения за последнее десятилетие получили широкое распространение не только в секторе неформального образования, но в высшем и дополнительном профессиональном образовании [European Association of Distance Teaching Universities, 2018; Нетология-групп, 2017]. Использование МООК для реализации образовательных программ [Рощина, Рощин, Рудаков, 2018] дало вузам и учреждениям среднего профессионального образования возможность расширить образовательный выбор, предоставляемый студентам, и создало условия для виртуальной академической мобильности [Sancho, de Vries, 2013], для повышения доступности образования и снижения стоимости образовательных услуг [Ларионова, Третьяков, 2016]. Прибегая к использованию МООК, университеты сталкиваются с проблемой выбора качественных онлайн-курсов и с необходимостью оценивать эффективность онлайн-обучения. Стратегии выбора онлайн-курсов и методы оценивания результатов их применения нуждаются в системном изучении, результатом которого должно стать появление четких критериев для принятия решений. Аналитика учебной деятельности обучающихся на платформах открытого образования при освоении ими онлайн-курсов является одним из инструментов повышения качества обучения [O'Farrell, 2017]. Данные учебной аналитики не только позволяют осуществлять мониторинг успеваемости обучающихся, анализировать их активность и вовлеченность в процесс обучения, но и дают объективную информацию об эффективности применяемых методик и технологий онлайн-обучения.

На платформах открытого образования представлены многообразные онлайн-курсы [Hollands, Tirthali, 2014]. При этом качество МООК при включении в образовательную программу университета определяется эффективностью их применения для достижения целей обучения. Эффективность обучения вслед за специалистами [Загвязинский, Закирова, 2008; Самохин и др., 2018] мы понимаем как «степень соответствия результатов образовательной деятельности поставленным целям», а не просто как эквивалент экономической эффективности, определяемой соотношением полученных конкретных результатов обучения и затраченных ресурсов [Вишнякова, 1999]. Достоверность оценки эффективности онлайн-образования зависит от адекватности применяемых при оценивании результатов обучения контрольно-измерительных средств и точности их соотнесения с требуемыми уровнями освоения курса. В отличие от традиционного обучения, где преподаватель в очном взаимодействии с обучающимся дает субъективную оценку его знаний и навыков, в исключительно электронном обучении с применением МООК оценка образовательных результатов происходит, как правило, в автоматическом режиме с помощью тестов или в режиме взаимного оценивания учебных и контрольных

заданий. Для достижения объективности оценки необходимо обеспечить выполнение следующих условий, которые являются базовыми принципами классической и современной теорий тестирования (IRT) [Крокер, Алгина, 2010]:

- цели онлайн-курса должны быть сформулированы в связи с конкретными образовательными результатами [Нехаев, 2016];
- результаты обучения должны поддаваться измерению;
- необходимо выбрать валидные, надежные и чувствительные к уровню достижения результатов обучения контрольно-измерительные средства;
- результаты оценивания должны быть достоверны и репрезентативны [Шмелев, 2013].

Существующие психометрические методы позволяют оценить качество тестовых заданий на основе математических моделей и аналитических процедур, применяемых для анализа ответов обучающихся на конкретные вопросы тестов [Майоров, 2002; Звонников, Челышкова, 2012]. Предлагаемая в настоящей статье методика оценки информативности и качества контрольно-измерительных материалов в MOOK, основанная на теории информации, расширяет инструментарий психометрики и может быть использована в дополнение к известным средствам оценки валидности тестов.

Социальная необходимость изучения эффективности использования онлайн-технологий в образовании обусловлена актуальностью организации обучения в информационном обществе с его высокими темпами сменяемости технологий и накопления информации по модели обучения на протяжении всей жизни. Причинами пока еще недостаточно активного развития непрерывного обучения могут быть, в частности, несовершенство существующих онлайн-курсов и низкий уровень мотивации студентов, которые в основном относятся к так называемому поколению Z и характеризуются такими особенностями восприятия информации, как зависимость от технологий, нетерпение, стремление к участию [Freitas, Morgan, Gibson, 2015], привычка получать необходимую информацию посредством поиска в интернете [Грязнова, Муковозов, 2016; Guo, Kim, Rubin, 2014; Tyler-Smith, 2006]. Эффективность традиционных технологий обучения применительно к данному контингенту обучающихся оказывается невысокой, и на повестку дня выдвигается необходимость модернизации учебного процесса.

Наряду с социальной необходимостью изучения эффективности использования онлайн-технологий в образовании имеет и другой аспект — педагогический. В основу онлайн-обучения по-прежнему закладывается содержание консервативных

программ массового образования, и не учитываются требования новой образовательной парадигмы [Jansen, Schuwer, 2015; Kop, Fournier, Mak, 2011]. Представители традиционного подхода рассматривают контент MOOK как последовательность видеороликов и стандартных текстовых блоков для прочтения, притом что на протяжении последних 20 лет обучение понимается не просто как доступ к информации, но как приобретение конкретных практических навыков [Lundvall, Borrás, 1997; Nonaka, Takeuchi, 2011]. В результате статистика обучения с использованием открытых онлайн-курсов, как правило, демонстрирует резкое снижение интереса студентов к учебному процессу, а также разрыв между их ожиданиями и предложениями учебных заведений [Brown, Lally, 2017; Castano Muñoz et al., 2016]. Данные о сравнительной эффективности различных онлайн-технологий позволяют сократить этот разрыв.

Обеспечение эффективности онлайн-обучения личносно значимо для современного обучающегося. Человек цифровой эпохи испытывает потребность в индивидуализации траектории обучения, в ее адаптации на основе технологий искусственного интеллекта к собственным запросам и способностям. Массовые открытые онлайн-курсы предоставляют большие возможности для построения индивидуальных образовательных траекторий и непрерывного обучения на протяжении всей жизни человека [Deev, Glotova, Krevskiy, 2015], в том числе и в силу их способности адаптироваться под индивидуальные особенности и запросы обучающегося.

Технологические предпосылки настоящего исследования заложены в самом формате исключительно электронного обучения, которое сопровождается фиксацией результатов обучения слушателей в виде цифрового следа в электронной информационно-образовательной среде. Такая фиксация позволяет проследить образовательный путь обучающегося, выявить причинно-следственные связи между его активностями в онлайн-курсе и результатами обучения, изучить возможные причины его неуспеха и на основе текущей успеваемости прогнозировать итоговый прогресс. Кроме того, учебная аналитика является одним из немногих объективных показателей качества MOOK и служит инструментом для его улучшения.

Основная гипотеза исследования заключается в том, что на основе данных учебной аналитики можно получать объективную информацию об эффективности онлайн-обучения, а также прогнозировать успеваемость различных категорий слушателей. Целью настоящего исследования является разработка алгоритмов учебной аналитики для оценки качества контрольно-измерительных материалов онлайн-курсов, изучения профилей успеваемости слушателей и определения вероятности их успешности/неуспешности на примере массовых открытых

онлайн-курсов Уральского федерального университета, размещенных на Национальной платформе открытого образования. В рамках поставленной цели решаются следующие задачи: 1) проанализировать качество контрольно-измерительных материалов онлайн-курсов на основе эмпирических данных; 2) рассчитать и сравнить функции распределения успеваемости слушателей курсов для всех промежуточных контрольных испытаний и итогового тестирования; 3) провести кластеризацию слушателей по успеваемости и проанализировать их прогресс в динамике; 4) построить вероятностную модель изменения успеваемости разных категорий слушателей в процессе обучения. Исследование направлено на выявление факторов, оказывающих негативное воздействие на успеваемость обучающихся в процессе освоения ими онлайн-курсов. Результаты исследования помогут сформулировать рекомендации для авторов курсов по совершенствованию педагогических методик онлайн-обучения и повышению качества контрольно-измерительных инструментов, а также предложения для тьюторов и технических специалистов по сопровождению образовательного процесса, реализуемого с применением онлайн-курсов.

Под массовым открытым онлайн-курсом будем понимать размещенную в открытом доступе структурированную, методически обоснованную, целенаправленную последовательность учебно-методических, контрольно-измерительных материалов и других средств для реализации учебной деятельности в формате исключительно электронного обучения. Онлайн-курс задает методику обучения, ход освоения материалов курса, контрольные точки и процедуры проверки знаний и навыков обучающихся. Коммуникация слушателей с преподавателями и между собой осуществляется посредством сервисов электронной информационно-образовательной среды. Хорошо продуманный педагогический дизайн онлайн-курса обеспечивает достижение результатов обучения при условии соблюдения требований к входному уровню знаний и навыков обучающегося и достаточной мотивации к обучению.

К освоению онлайн-курса допускаются все желающие независимо от возраста, местоположения, полученного ранее образования и финансовых возможностей. Процесс обучения с использованием MOOK проходит, как правило, в асинхронном режиме, т. е. передача знаний от преподавателя к студенту происходит опосредованно. Такой режим обучения позволяет обучающимся самостоятельно планировать график прохождения курса с учетом индивидуальных предпочтений и возможностей, выбирать темп освоения материалов курса, а также время выполнения контрольных заданий. Такие курсы (*self-*

1. Теоретические аспекты эффективности онлайн-обучения

1.1. Особенности обучения с использованием MOOK

paced) не привязаны к определенным датам и доступны в режиме *on-demand*, т. е. в любое удобное для слушателя время. Для сохранения постоянного темпа освоения курса и повышения организованности слушателей в большинстве курсов вводятся ограничения по времени старта курса, по датам проведения онлайн-консультаций в формате вебинаров и по срокам прохождения контрольных мероприятий, в том числе итоговой аттестации.

Для получения сертификата, подтверждающего факт успешного освоения онлайн-курса, слушателю необходимо пройти итоговое контрольное мероприятие с идентификацией личности, которая осуществляется в режиме онлайн с использованием электронных сервисов прокторинга. Сертификат выдается слушателю, набравшему определенное количество баллов в течение всего процесса обучения (минимальный порог указывается в описании курса) и успешно сдавшему итоговый тест. Итоговый экзамен с прокторингом, необходимый для получения сертификата, как правило, бывает платным. Студенты вузов могут предъявить сертификат, подтверждающий факт успешного освоения онлайн-курса, для зачета результатов онлайн-обучения в основных или дополнительных профессиональных образовательных программах вузов. Процедура перезачета регулируется локальными нормативно-правовыми актами образовательных организаций.

1.2. Факторы, определяющие эффективность онлайн-обучения

В отличие от электронного учебно-методического комплекса как пакета образовательных и контролирующих ресурсов в соответствии с рабочей программой дисциплины (модуля), важным элементом онлайн-курса является организация процесса обучения, а также регулярный контроль результатов слушателей. В этом смысле электронный курс является авторской технологией обучения. Его эффективность определяется не столько качеством контента, сколько методикой преподавания, воплощенной в электронной образовательной среде, и качеством контрольно-измерительных материалов, позволяющих адекватно оценить уровень освоения материалов онлайн-курса обучающимися.

Компонентами онлайн-обучения, определяющими его эффективность, являются:

- методически обоснованное представление электронного контента согласно циклу усвоения человеком новой информации [Kolb, 1985];
- использование интерактивных элементов обучения;
- контроль результатов обучения и диагностика ошибок на всем протяжении курса;

- организация взаимодействия обучающихся в процессе освоения курса;
- поддержка обучающихся в процессе обучения и мотивирование их к достижению высоких результатов;
- применение активных методов обучения в режиме онлайн;
- получение обратной связи от обучающихся и ее статистическая обработка;
- оперативное внесение изменений и дополнений в электронный курс в случае необходимости [Jasnani, 2013].

Разработка такого электронного курса представляет собой сложную педагогическую задачу, требующую высокого уровня профессионального мастерства, методической подготовки, большого педагогического опыта и навыков владения информационными технологиями. Главное в процессе разработки эффективного электронного курса — это реализация интерактивных технологий, основанных на активных методах обучения, в электронной форме [Lisitsyna, Lyamin, 2014].

Таким образом, онлайн-курс за счет использования сервисов электронной информационно-обучающей среды обеспечивает функцию управления процессом обучения и может применяться без прямого участия преподавателя в работе с каждым из обучающихся. Сопровождение курса в этом случае состоит в обновлении его содержания во время и после завершения обучения и консультационной поддержке обучающихся. Поскольку поддержка выполняется в рамках представленного содержательного контента и методики обучения, она может осуществляться без участия автора курса и в большинстве случаев не требует от консультанта глубоких компетенций в изучаемой области. Таким образом, основная функция преподавателя заключается в создании электронного курса, а процесс его освоения слушателями могут сопровождать тьюторы, которые осуществляют методическую и организационную поддержку обучающихся, консультируют по вопросам выбора онлайн-курсов и зачета результатов обучения в основных образовательных программах, помогают с построением индивидуальных траекторий, обеспечивая тем самым условия для успешного прохождения обучающимися промежуточных и итоговых контрольных мероприятий.

В отличие от традиционного обучения, при котором преподаватель получает обратную связь от обучающегося только при непосредственном контакте с ним, обучение в онлайн-формате оставляет за собой «цифровой след». Все достижения слушателя и его активности в ходе освоения материалов MOOK фиксируются в электронной информационно-образовательной среде. Анализ таких данных — учебная аналитика — позволяет контро-

**1.3. Использование
данных учебной
аналитики для
поддержки
обучающихся**

лизовать регулярность занятий слушателя, осуществлять мониторинг его успеваемости, следить за ходом выполнения контрольных заданий.

В основе учебной аналитики лежит методология анализа больших данных о поведении обучающихся в процессе освоения курса [Usha Keshavamurthy, Guruprasad, 2014]. Данные учебной аналитики могут много рассказать о причинах неудачи или успеха ученика и позволяют спрогнозировать его будущее поведение. Анализ этих данных дает основания для корректировки контекста обучения, оказания поддержки обучающимся или адаптации их к новым условиям [O'Farrell, 2017]. Основными целями учебной аналитики являются:

- измерение, сбор и представление данных о пользовательском поведении;
- анализ успеваемости обучающихся на протяжении освоения курса;
- анализ поведенческих паттернов на основе больших данных;
- установление причинно-следственных связей между показателями успешности ученика и его учебными активностями;
- выявление ошибок и методических проблем в курсе;
- выработка рекомендаций по корректировке содержания курса;
- прогнозирование успешности обучающихся.

Методы учебной аналитики различны: от дескриптивной статистики до инструментов интеллектуальной обработки данных. Дополнительными источниками информации, кроме потоковых данных о пользовательском поведении с платформ открытого образования, могут служить административные базы образовательных учреждений, опросы слушателей, преподавателей, результаты претестов и др.

Мировыми лидерами в сфере учебной аналитики являются следующие исследовательские центры: ирландский Национальный форум совершенствования преподавания и обучения в высшем образовании (*National Forum for the Enhancement of Teaching and Learning in Higher Education*), Национальный исследовательский центр дистанционного образования и технологических достижений при Университете Висконсин — Милуоки (*National Research Center for Distance Education and Technological Advancements, University of Wisconsin-Milwaukee*), EdPlus Университета штата Аризона.

В настоящее время внимание исследователей Аризонского университета сосредоточено на поиске эффективных инструментов адаптивного обучения на основе больших данных о поведении обучающихся в процессе освоения электронных курсов. Выявляя на ранних стадиях обучения поведенческие паттерны

и группируя обучающихся на основании особенностей их учебных активностей, ученые исследуют факторы, оказывающие положительное влияние на успеваемость, и прогнозируют итоговый прогресс по курсу (например, [Sharkey, Ansari, 2014]).

Одной из задач применения описываемых ниже алгоритмов является анализ информативности и качества используемых в процессе обучения контрольных заданий. Эти задания должны, во-первых, обеспечивать дифференциацию слушателей по степени их успешности, а во-вторых, приводить к нарастающему прогрессу в освоении изучаемого материала и отражать этот прогресс. Не менее важной задачей является прогнозирование результатов прохождения контрольных точек курса слушателями с разным уровнем усвоения материала, отражаемым в результатах текущего контроля. Такое прогнозирование позволит адаптировать слушателей, различающихся по уровню успеваемости, к требованиям курса посредством дополнительных консультаций, индивидуальных заданий и т. д.

С позиций первой задачи одинаково неудачными следует признать и контрольные задания, которые не удастся выполнить подавляющему большинству слушателей, и задания, которые успешно выполняет подавляющее большинство обучающихся. Применение таких контрольных заданий не дает преподавателю, ведущему курс, никакой информации о прогрессе в обучении и успехах отдельных слушателей.

Для оценки информативности контрольных заданий с точки зрения дифференциации слушателей по степени их успешности использованы стандартные методы теории информации. Если обозначить через $\varphi(x)$ распределение оценок (в баллах), полученных слушателем при прохождении контрольной точки, то факт, что конкретный слушатель получил тот или иной балл, несет следующее количество двоичных единиц информации [Корн, Корн, 1973]:

$$I = - \int_0^{100} \varphi(x) \cdot \log_2(\varphi(x)) \cdot dx, \text{ бит.} \quad (1)$$

В практических расчетах область изменения оценок разбивалась на дискретные интервалы с шагом в 10 баллов, а интеграл, соответственно, трансформировался к сумме подынтегрального выражения по этим интервалам. Для удобства будем сравнивать эту величину с максимальным количеством информации, которому соответствует равномерное распределение $\varphi_1(x) = 1/n$, где n — количество интервалов:

$$I_{\max} = \log_2(n) = 3,22.$$

2. Методология исследования

В этом случае для характеристики информативности контрольной точки будем использовать величину:

$$(2) \quad inf = 100 \frac{I}{I_{max}},$$

округленную до целых баллов.

Для отражения прогресса в освоении материала и прогнозирования результатов изучаемого курса необходимо анализировать статистические характеристики последовательного прохождения контрольных точек каждым отдельным слушателем. В предыдущей работе авторов [Larionova et al., 2018] уже рассматривались изменения в ходе освоения дисциплины статистических распределений оценок для групп слушателей, объединенных по средней успеваемости предыдущих периодов обучения, — для отличников, хорошистов и т. д.

Для решения задачи отразить контрольно-измерительными инструментами прогресс обучающихся в изучении курса введем три группы слушателей, различающихся по уровню освоения курса:

- неуспевающие: слушатели этой группы не прошли контрольное мероприятие, т. е. получили менее 40 баллов — «неуд.»;
- успевающие удовлетворительно: слушатели, чьи оценки лежат в интервале от 40 до 60 баллов — «удовл.»;
- успевающие: слушатели, получившие 60 и более баллов — «успех».

Количество групп может быть увеличено, однако и трех достаточно, чтобы полно отразить уровень успешности слушателя в изучении материала и вместе с тем обеспечить наглядность результатов.

В процессе изучения курса и прохождения контрольных точек слушатели совершают переходы из одной из вышеперечисленных групп в другую. Отслеживая такие переходы по каждому слушателю, можно рассчитать вероятности переходов между группами по успеваемости для каждой контрольной точки. Точность расчета зависит от числа слушателей в выборке: чем их больше, тем достовернее могут быть оценены вероятности переходов. Эти расчеты позволят сделать определенные заключения о том, как контрольная точка отражает прогресс слушателя в изучении курса, а также спрогнозировать результаты прохождения контрольных точек разными категориями слушателей. Для такого прогноза необходимо накапливать информацию о вероятностях переходов слушателей между группами, различающимися по успеваемости, и обрабатывать большие объемы данных о прохождении обучающимися контрольных точек.

Обозначим через $|i\rangle$ и $|j\rangle$ состояния слушателя перед контрольной точкой и после ее прохождения соответственно (под состоянием понимаем принадлежность его к группе по успеваемости с номером i до начала испытания и с номером j после его прохождения, $i, j = 1, 2, 3$). Поставим в соответствие каждому переходу между группами по успеваемости оператор T_{ij} , действующий по правилу:

$$T_{ij} \cdot |i\rangle = |j\rangle. \quad (3)$$

Оператор T_{ij} является оператором перестановки $i \rightarrow j$, причем вероятности переходов определяются матрицей

$$\hat{P} = \begin{pmatrix} P_{11} & P_{12} & P_{13} \\ P_{21} & P_{22} & P_{23} \\ P_{31} & P_{32} & P_{33} \end{pmatrix}. \quad (4)$$

Матрица $\hat{P}$ не является симметричной, кроме того, ее элементы удовлетворяют соотношению:

$$\sum_{j=1}^3 P_{ij} = 1. \quad (5)$$

Для определения численности слушателей в каждой из групп при вероятностях (4) можно использовать модель, предложенную в [Astratova et al., 2017], которая позволяет определить вероятность того, что в группах 1, 2, 3 будет соответственно X_1 , X_2 , X_3 членов в момент времени $t - P(X_1, X_2, X_3 | t)$. Уравнение для $P(X_1, X_2, X_3 | t)$ имеет вид:

$$\begin{aligned} \frac{\partial P(X_1, X_2, X_3 | t)}{\partial t} = & P(X_1, X_2, X_3 | t) \cdot \left\{ (1-z) \cdot \sum_{j=1}^3 P_{jj} - \sum_{j=1}^3 X_j \right\} + \\ & + z \cdot \sum_{i=1}^3 (X_i + 1) \cdot P(\dots, X_{i+1}, \dots | t) + (1-z) \cdot \sum_{j=1, i \neq j}^3 P_{ij} \cdot (X_i + 1) \times \\ & \times P(\dots, X_{i+1}, \dots | t). \end{aligned} \quad (6)$$

Здесь z — вероятность прекращения обучения слушателем в единицу времени. В дальнейшем будем полагать, что $z = 0$ (для этого на подготовительном этапе исключим соответствующих лиц из анализируемых данных).

Уравнение (6) может быть решено в общем виде, однако для большинства задач можно ограничиться рассмотрением средних и ковариаций:

$$\begin{aligned} \bar{X}_i = \langle X_i \rangle &= \int_0^{100} X_i \cdot P(X_1, X_2, X_3 | t) \cdot dX_i, \\ \sigma_{ij} = (X_i - \langle X_i \rangle) \cdot (X_j - \langle X_j \rangle) &= \int_0^{100} X_i \cdot X_j \cdot P(X_1, X_2, X_3 | t) \cdot dX_i \cdot dX_j - \langle X_i \rangle \cdot \langle X_j \rangle. \end{aligned} \quad (7)$$

Можно показать, что выполняются следующие соотношения:

$$(8) \quad \bar{X}_i \sim N, \\ \sigma_{ij} \sim \sqrt{N}.$$

Здесь N — общее число слушателей. Таким образом, при $N \rightarrow \infty$ коэффициенты вариации стремятся к нулю:

$$C_v(ij) \sim \frac{1}{\sqrt{N}} \rightarrow 0,$$

что отражает закон больших чисел. Таким образом, при большой численности слушателей N их распределение по группам практически неслучайно и число участников группы близко к $\langle X_i \rangle$, причем:

$$\langle X_1 \rangle + \langle X_2 \rangle + \langle X_3 \rangle = N.$$

Уравнение для X_i имеет вид:

$$(9) \quad \frac{\partial X_i}{\partial t} = \sum_{k=1}^3 [\tilde{P}_{ki} \cdot X_k - \tilde{P}_{ik} \cdot X_i],$$

где

$$(10) \quad \tilde{P}_{ki} = 0 \text{ для } k = i; \\ \tilde{P}_{ki} = P_{ki}, \text{ определяемое (4) для } k \neq i.$$

Матрицу переходов (4) можно связать с задачей о случайных блужданиях по ориентированному графу, вершины которого соответствуют группам $i = 1, 2, 3$, а вероятности переходов между вершинами определяются из (4) [Лесковец, Раджараман, Ульман, 2016].

Вероятности переходов (4) однозначно определяют влияние контрольной точки на распределение слушателей по группам успеваемости и могут косвенно служить индикаторами качества контрольного задания, однако непосредственное использование (4) неудобно — прежде всего из-за большого количества параметров (девять вероятностей переходов) и их сложной, хотя и однозначной, связи с традиционными и понятными характеристиками успехов в обучении. Поэтому в качестве наглядной характеристики будем использовать вектор $\vec{\alpha}_x = \{X_1, X_2, X_3\}$, определяющий установившееся распределение слушателей по группам успеваемости $j = 1, 2, 3$. Его можно рассматривать как установившееся решение уравнения (9), соответствующее стационарному случаю $\left(\frac{\partial X_i}{\partial t} = 0\right)$, или как предельное распределение, возникающее после многократных переходов вида

$\vec{\alpha}_x(n) = \hat{P} \cdot \vec{\alpha}_x(n-1)$ на графе, соответствующем матрице (4) [Astratova et al., 2017] при $n \rightarrow \infty$. Этот предельный случай соответствует гипотетической ситуации многократного прохождения рассматриваемой контрольной точки группами слушателей со статистически эквивалентными характеристиками успехов в обучении. Нетрудно убедиться, что $\vec{\alpha}_x(n \rightarrow \infty) \equiv \vec{\alpha}_x$ удовлетворяет уравнению [Ibid.]:

$$\vec{\alpha}_x = \hat{P} \cdot \vec{\alpha}_x. \quad (11)$$

Таким образом, $\vec{\alpha}_x$ — это собственный вектор $\hat{P}$ (4), соответствующий собственному значению 1. Используя (5), (10), можно показать, что $\vec{\alpha}_x$ в (11) соответствует установившемуся решению (9) для $\frac{\partial X_i}{\partial t} = 0$.

Соотношение, аналогичное (11), можно использовать при известной матрице $\hat{P}$ (4) и для прогнозирования результатов прохождения контрольной точки. Пусть $\vec{\alpha}_x(0)$ — вектор, характеризующий распределение слушателей по группам успеваемости до прохождения контрольной точки, а $\vec{\alpha}_x(1)$ — после ее прохождения, тогда, как известно из теории марковских процессов [Максимов, 2001], эти два вектора связаны соотношением:

$$\vec{\alpha}_x(1) = \hat{P} \cdot \vec{\alpha}_x(0). \quad (12)$$

Здесь $\hat{P}$ — матрица вида (4), соответствующая анализируемой контрольной точке.

В качестве примера применения алгоритма, описанного в предыдущем разделе, проанализируем данные по онлайн-курсу «Инженерная механика» Уральского федерального университета, размещенному на Национальной платформе открытого образования¹. Данный курс включает следующие контрольно-измерительные инструменты (контрольные точки):

- 16 тестов по теоретическому материалу (Т);
- 18 домашних заданий (ДЗ);
- 5 базовых проектов по курсу (БП);
- итоговый тест (ИТ).

В исходной базе данных прохождение каждой контрольной точки оценивалось по 100-балльной системе, и каждой из них присваивался свой весовой коэффициент k_p , $p = 1, \dots, 4$. С использованием весовых коэффициентов $0 \leq k_p \leq 1$ и оценок $0 \leq B_j(C) \leq 100$, полученных слушателем на каждой контрольной

3. Практическое применение алгоритма учебной аналитики и обсуждение результатов

¹ <https://openedu.ru/course/urfu/ENGM/>

точке, где $C = T, ДЗ, БП, ИТ$, рассчитывались следующие показатели:

- средняя оценка по текущей успеваемости

$$(13) \quad Avg = k_1 \frac{1}{16} \sum_{j=1}^{16} B(T)_j + k_2 \frac{1}{18} \sum_{j=1}^{18} B(ДЗ)_j + k_3 \frac{1}{5} \sum_{j=1}^5 B(БП)_j;$$

- итоговая оценка курса

$$(14) \quad Grade = Avg + k_4 B(ИТ).$$

- В соответствии с установками авторов курса коэффициенты k_p принимали значения: $k_1 = 0,16$; $k_2 = 0,34$; $k_3 = 0,1$; $k_4 = 0,4$. Таким образом, максимальное значение Avg равно 60. Для удобства сравнения результатов разных контрольных точек с этой величиной для каждого слушателя с идентификатором i она также пересчитывалась в 100-балльный эквивалент по формуле

$$(15) \quad Avg_j(100) = 100 \cdot \frac{Avg_i}{\text{Max}\{Avg_j | j = 1, \dots, N\}},$$

где N — общее число слушателей анализируемого курса.

Каждой контрольной точке могут быть поставлены в соответствие следующие характеристики:

- средняя оценка, полученная слушателями курса;

- коэффициент решаемости задания: $k_i = \frac{C_i}{N}$,

где C_i — число слушателей, выполнивших задание, N — общее число слушателей курса;

- функция плотности распределения вероятностей оценок, полученных слушателями при прохождении контрольной точки (более сложная характеристика).

С помощью последней характеристики может быть рассчитана информативность контрольной точки (2).

Рассмотрим в качестве примера тест 1. Функция распределения баллов, полученных по данному тесту, представлена на рис. 1. Для сравнения на этом же рисунке представлено нормальное распределение. Даже без использования специальных критериев видно, что отклонения фактического распределения баллов по тесту 1 от нормального существенны и не могут быть объяснены случайными причинами. Расчет информации, которая содержится в факте «слушатель получил по тесту 1 определенное количество баллов», по формуле (1) дает: $I = 1,47$ бит.

Для расчета информативности всех контрольных точек типа «тест» используем формулу (2). Полученные результаты представлены на рис. 2.

Из рис. 2, в частности, видно, что наиболее информативны тесты 2, 14 и 15. Именно они эффективно дифференцируют слу-

Рис. 1. Плотность распределения вероятности баллов, полученных по тесту 1 $\varphi_1(x)$

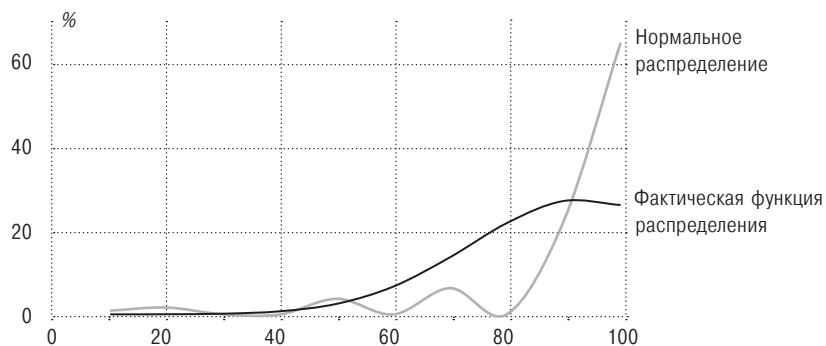


Рис. 2. Характеристика информативности тестов, рассчитанная по формуле (2)

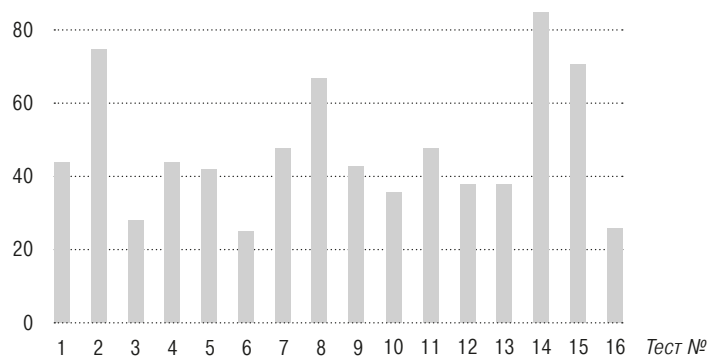


Таблица 1. Характеристики наиболее и наименее информативных тестов

Тест №	Информативность $inf(2)$	Средняя оценка	Коэффициент решаемости
6	25	96,3	0,993
16	26	95,8	0,985
2	75	80,2	0,898
14	85	68,5	0,797
15	71	72,4	0,869
Относительная разница между максимальным и минимальным значением $(Max - Min)/Min, \%$	240	40,5	24,6

Рис. 3. Плотность распределения вероятности получения заданных оценок по текущей успеваемости, итоговому тесту и *Grade*

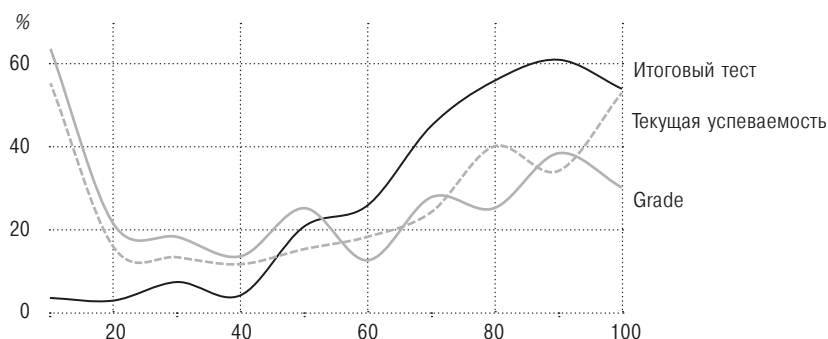


Таблица 2. Информативность контрольных мероприятий

№	Контрольное мероприятие	<i>Inf</i> , формула (2)
1	Текущая успеваемость (<i>Avg</i>)	93
2	Итоговый тест	84
3	<i>Grade</i>	94

шателей по уровню усвоения материала, тогда как, например, тесты 3, 6 и 16 наименее информативны: подавляющее большинство слушателей сдают их на отлично — возможно, данные задания слишком легки. В табл. 1 информативность наиболее и наименее информативных тестов сопоставлена с другими характеристиками контрольной точки.

Разница между наиболее и наименее информативными тестами в информативности, рассчитанной по формуле (2), существенно больше разницы по таким характеристикам, как средняя оценка и коэффициент решаемости. Таким образом, информативность наиболее удобна для сравнения контрольных точек и оценки их качества.

Среди всех типов контрольных точек наибольший интерес представляют:

- текущая успеваемость (*Avg*);
- итоговый тест (ИТ);
- *Grade* — интегральная оценка прохождения курса слушателем с учетом текущей успеваемости и итогового теста.

Плотности распределения вероятностей оценок по баллам для этих типов контрольных мероприятий представлены на рис. 3.

Таблица 3. Установившиеся распределения слушателей по группам успеваемости и информативность (i) для разных контрольных точек

Доля слушателей в группе	Тест	ДЗ	БП	ИТ
x_1 (неуспевающие)	0,336	0,436	0,658	0,296
x_2 (успевающие удовлетворительно)	0,002	0,087	0	0,129
x_3 (успевающие)	0,662	0,477	0,342	0,575
inf (информативность)	59	84	58	86

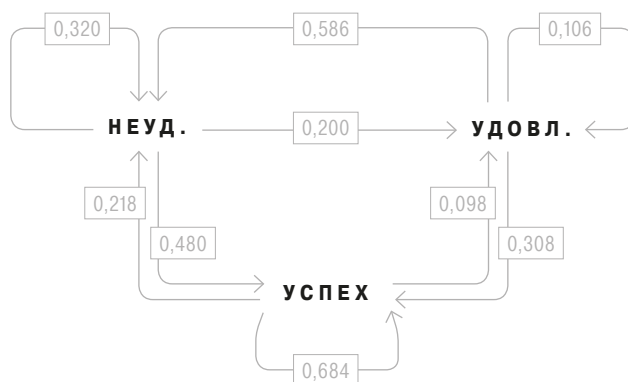
По каждой контрольной точке, представленной на рис. 3, спектр оценок достаточно широк, т.е. каждая из них обеспечивает хорошую дифференциацию слушателей. Данные по информативности соответствующих контрольных мероприятий inf , рассчитываемой по формуле (2), представлены в табл. 2.

Таким образом, такие интегральные контрольные мероприятия, как средняя текущая успеваемость и $Grade$, отражающие усвоение слушателем материала в процессе изучения всего курса, обладают высокой информативностью, что не всегда можно сказать об отдельных контрольных точках (см. табл. 1). Информативность итогового тестирования несколько меньше, но также достаточно высока.

Далее мы анализируем не отдельные контрольные точки, а их серии, соответствующие различным типам заданий (Т, ДЗ, БП, ИТ). При этом в качестве входного состояния выбирается состояние после прохождения первой контрольной точки серии. Таким образом мы получаем возможность анализировать все серии контрольных точек независимо друг от друга и снимаем проблему отсутствия входного тестирования на большинстве онлайн-курсов (информация о входном уровне подготовки обучающихся слушателей, как правило, недоступна). Полученные результаты представлены в табл. 3. Результаты анализа можно представить в еще более компактном виде, если использовать фактор inf (2). В данном случае он отражает информативность распределения слушателей по группам после прохождения контрольной точки.

Группы контрольно-измерительных инструментов типа «тест», а также «базовый проект», по сути, приводят к расслоению слушателей на две группы: неуспевающие и успевающие. При этом промежуточная группа — успевающие удовлетворительно — практически отсутствует. Эти данные свидетельствуют, в частности, о невысокой информативности соответствующих групп контрольных точек, что подтверждает последняя строка в табл. 2. Действительно, после прохождения этих контрольных

Рис. 4. **Ориентированный граф для переходов между группами по успеваемости, порождаемых контрольной точкой «итоговый тест».** Числа соответствуют вероятностям переходов между группами (16)



точек слушатели получают либо неудовлетворительный результат, либо высокие оценки. Вполне возможно, что предлагаемые задания излишне легки или оценивание результатов производится формально, по принципу «сдал — не сдал», особенно при приеме проекта. Разумеется, могут быть и другие причины отмеченного расслоения. В любом случае очевидно, что проведение описанного анализа дает авторам курса полезную информацию для оценки качества контрольно-измерительных инструментов.

При рассмотрении итогового теста в качестве входного состояния $|i\rangle$ можно использовать данные по текущей успеваемости (Avg). Таким образом, мы будем подсчитывать переходы между различающимися по успеваемости группами в процессе итогового тестирования: $Avg(i) \rightarrow ИТ(j)$ (здесь i и j — группы по успеваемости). Подсчет соответствующих пар $\{ij\}$ для переходов между группами по успеваемости после итогового теста дает следующую матрицу:

$$(16) \quad \hat{P} = \begin{pmatrix} 0,320 & 0,586 & 0,218 \\ 0,200 & 0,106 & 0,098 \\ 0,480 & 0,308 & 0,684 \end{pmatrix}.$$

Вероятности переходов можно изобразить в виде ориентированного графа, представленного на рис. 4.

На рис. 4 видно, что наиболее вероятны переходы из группы «успех» в нее же, а также из группы «удовл.» в группу «неуд.». Неожиданно высокой оказалась также вероятность перехода из группы «неуд.» в группу «успех». Впрочем, исследователи

из Университета штата Аризона тоже выделяют такой психотип обучающихся, называя их «кенгуру» ([Johnson, 2018]).

Предположим, что после завершения курса перед итоговым тестированием слушатели распределены по группам успеваемости равномерно:

$$x_1 = x_2 = x_3 = \frac{1}{3}.$$

При найденных вероятностях переходов (16) после прохождения итогового тестирования прогнозируемое распределение слушателей по группам успеваемости в соответствии с (12) будет иметь вид: $x_1 = 0,375$; $x_2 = 0,135$; $x_3 = 0,491$. Если прогнозируемое распределение слушателей по группам не устраивает преподавателей (например, в рассмотренном примере прогнозируется рост группы неуспевающих по сравнению с текущей успеваемостью), они могут предпринять дополнительные меры для поддержки обучающихся, чтобы повысить успеваемость.

Сравним результаты предлагаемого алгоритма учебной аналитики для разных онлайн-курсов. Ввиду того что структура контрольных точек для каждого из анализируемых курсов индивидуальна, целесообразно сравнивать переходы $Avg(i) \rightarrow ИТ(j)$ (i, j — группы «неуд.», «удовл.», «успех»), так как данные по средней текущей успеваемости и итоговое тестирование имеются в каждом курсе. Результаты представлены в табл. 4.

Прогнозируемая доля группы «успех» для курса «Начертательная геометрия и инженерная графика» минимальна, тогда как доля группы «неуд.», наоборот, максимальна среди рассмотренных курсов. Итоговое тестирование по этому курсу наиболее информативно.

Вероятности переходов $Avg(i) \rightarrow ИТ(j)$ между группами по успеваемости для курса «Начертательная геометрия и инженерная графика» представлены в таблице 5².

Вероятность перехода из группы «удовл.» в какую-либо группу, кроме группы «неуд.», крайне мала, тогда как вероятность перехода из группы «неуд.» в группу «успех» (психотип «кенгуру») достаточно высока (0,35). Группа «успех» на итоговом тестировании имеет тенденцию к расслоению: обучающиеся, отнесенные по текущей успеваемости к этой группе, либо переходят в группу «неуд.» (с вероятностью 0,38), либо — с большей вероятностью (0,54) — остаются в группе «успех».

При тестировании по курсу «Технология конструкционных материалов» психотип «кенгуру» проявляет себя сильнее, чем при тестировании по курсу «Начертательная геометрия и ин-

² В матрице представлены вероятности переходов из групп, соответствующих столбцам, в группы, соответствующие строкам (таким образом, сумма элементов по каждому столбцу равна 1).

Таблица 4. Прогнозируемое установившееся распределение слушателей по группам успеваемости для переходов Avg → ИТ в разных онлайн-курсах

	x_1 (неуд.)	x_2 (удовл.)	x_3 (успех)	inf (информативность)
Инженерная механика	0,296	0,129	0,575	86
Технология конструкционных материалов	0,197	0,105	0,698	73
Начертательная геометрия и инженерная графика	0,48	0,149	0,371	91

Таблица 5. Вероятности переходов между группами успеваемости Avg → ИТ для онлайн-курса «Начертательная геометрия и инженерная графика»

	Неуд.	Удовл.	Успех
Неуд.	0,40	0,98	0,38
Удовл.	0,25	0,00	0,08
Успех	0,35	0,02	0,54

Таблица 6. Вероятности переходов между группами успеваемости Avg → ИТ для онлайн-курса «Технология конструкционных материалов»

	Неуд.	Удовл.	Успех
Неуд.	0,13	0,64	0,15
Удовл.	0,40	0,07	0,03
Успех	0,47	0,29	0,83

женерная графика» (вероятность соответствующих переходов 0,47), тогда как вероятность перехода из группы «успех» в группу «неуд.» по результатам итогового тестирования здесь невысока (0,15). Большая часть группы «успех» с вероятностью 0,83 сохраняет высокие результаты. Наименее вероятен для этого курса переход из группы «успех» в группу «удовл.» (табл. 6).

Курс «Инженерная механика» был рассмотрен выше (см. рис. 4). По вероятностям переходов $Avg(i) \rightarrow ИТ(j)$ он существенно отличается от остальных курсов и дает, на наш взгляд, наиболее адекватное распределение итоговых баллов успеваемости обучающихся, что свидетельствует о достаточной надежности системы оценивания в данном онлайн-курсе, высоком уровне тьюторской поддержки и методически обоснованном содержании курса, обеспечивающем учебные достижения слушателей.

Онлайн-обучение представляет собой новую образовательную парадигму, возникновение которой обусловлено социокультурными процессами современности, прежде всего коммуникационными. Оно предполагает более активную обратную связь для обучающегося, которая определяет конкретную траекторию обучения, а в конечном счете вовлекает его в обучение на протяжении всей жизни. Образование перестает быть монологичным, студент становится активным субъектом процесса обучения. Предложенный в статье способ прогнозирования успешности прохождения онлайн-курса способствует обеспечению обратной связи для обучающегося, формированию его индивидуальной траектории обучения и со временем может стать неотъемлемой частью онлайн-обучения.

Проведенные исследования на основе данных учебной аналитики показывают, что:

- анализ информативности контрольно-измерительных инструментов в соответствии с описанной методикой может дать авторам курса полезную информацию о качестве контрольных мероприятий дополнительно к традиционно используемому психометрическому анализу;
- прослеживание индивидуальных траекторий прохождения слушателями курса контрольных точек и определение на его основе вероятностей переходов между группами слушателей по успеваемости может использоваться для расчета распределения слушателей по группам успеваемости после прохождения соответствующей контрольной точки. Такой расчет позволяет получить дополнительные данные для оценки качества контрольно-измерительных инструментов;
- знание вероятностей переходов между группами слушателей по успеваемости позволяет прогнозировать итоговое распределение слушателей по таким группам и предпринимать необходимые меры для корректировки усилий.

4. Выводы

1. Вишнякова С. М. (1999) Профессиональное образование: Словарь. Ключевые понятия, термины, актуальная лексика. Москва: НМЦ СПО.
2. Грязнова Ю., Муковозов О. (2016) Пилотное исследование РАСО «Как поколение Z воспринимает информацию» / IV Международная научно-практическая конференция «Коммуникация в социально-гуманитарном знании, экономике, образовании» (Минск, 7–9 апреля 2016 г.). Минск: БГУ.
3. Загвязинский В. И., Закирова А. Ф., Строкова Т. А. и др. (2008) Педагогический словарь. М.: Академия.
4. Звонников А. И., Челышкова М. Б. (2012) Контроль качества результатов обучения при аттестации (компетентностный подход): учеб. пособие. М.: Логос.

Литература

5. Корн Г., Корн Т. (1973) Справочник по математике для научных работников и инженеров. М.: Главная редакция физико-математической литературы.
6. Крокер Л., Алгина Дж. (2010) Введение в классическую и современную теорию тестов. М.: Логос.
7. Ларионова В. А., Третьяков В. С. (2016) Открытые онлайн-курсы как инструмент модернизации образовательной деятельности в вузе // Высшее образование в России. № 7. С. 55–66.
8. Лесковец Ю., Раджараман А., Ульман Дж. Д. (2016) Анализ больших наборов данных. М.: ДМК Пресс.
9. Майоров А. Н. (2002) Теория и практика создания тестов для системы образования. М.: Интеллект-Центр.
10. Максимов Ю. Д. (ред.) (2001) Вероятностные разделы математики. СПб.: Иван Федоров.
11. Нетология-групп (2017) Исследование российского рынка онлайн-образования и образовательных технологий. <http://edumarket.digital>
12. Нехаев И. Н. (2016) Анализ качества процесса обучения с использованием онлайн-курсов // Лучшие практики электронного обучения: материалы II методической конференции. Томск: Изд-во Томского ун-та. С. 8–14.
13. Рощина Я. М., Рощин С. Ю., Рудаков В. Н. (2018) Спрос на массовые открытые онлайн-курсы (МООС): опыт российского образования // Вопросы образования / Educational Studies Moscow. № 1. С. 174–199. doi: 10.17323/1814-9545-2018-1-174-199.
14. Самохин И. С., Сергеева М. Г., Соколова Н. Л., Марченко Е. А. (2018) Содержание понятия «эффективность образования» в контексте инклюзивных тенденций современной школы // Научный диалог. № 1. С. 278–288.
15. Семенова Т. В., Вилкова К. А., Щеглова И. А. (2018) Рынок массовых открытых онлайн-курсов: перспективы для России // Вопросы образования / Educational Studies Moscow. № 2. С. 173–197. doi: 10.17323/1814-9545-2018-2-173-197.
16. Шмелев А. Г. (2013) Практическая тестология: тестирование в образовании, прикладной психологии и управлении персоналом. М.: Маска.
17. Astratova G., Sinicin E., Toporkova E., Frishberg L., Karabanova I. (2017) Mechanism of Information Model Development for Company Brand Assessment within Marketing Strategy // Karpov A., Martyushev N. (eds) Proceedings of the International Conference on Trends of Technologies and Innovations in Economic and Social Studies (Tomsk, 28–30 June 2017). Tomsk: Atlantis. P. 20–25.
18. Brown K., Lally V. (2017) It Isn't Adding Up: The Gap between the Perceptions of Engineering Mathematics Students and Those Held by Lecturers in the First Year of Study of Engineering // Proceedings of the 10th Annual International Conference of Education, Research and Innovation — ICE-RI2017 (Seville, Spain, 16–18 November 2017). Valencia: IATED. P. 317–321.
19. Castaño Muñoz J., Punie Y., Inamorato dos Santos A., Mitic M., Morais R. (2016) How Are Higher Education Institutions Dealing with Openness? A Survey of Practices, Beliefs and Strategies in Five European Countries. Institute for Prospective Technological Studies. JRC Science for Policy Report, EUR27750EN. <http://publications.jrc.ec.europa.eu/repository/bitstream/JRC99959/Ifna27750enn.pdf>
20. Deev M. V., Glotova T. V., Krevskiy I. G. (2015) Individualized Learning Trajectories Using Distance Education Technologies // Creativity in Intelligent, Technologies and Data Science. Ser. Communications in Computer and Information Science. Part XI. Vol. 535. P. 778–792.

21. European Association of Distance Teaching Universities (2018) The 2018 OpenupEd Trend Report on MOOCs. https://www.openuped.eu/images/Publications/The_2018_OpenupEd_trend_report_on_MOOCs.pdf
22. Freitas S. I., Morgan J., Gibson D. (2015) Will MOOCs Transform Learning and Teaching in Higher Education? Engagement and Course Retention in Online Learning Provision // *British Journal of Educational Technology*. Vol. 46. No 3. P. 455–471.
23. Guo P. J., Kim J., Rubin R. (2014) How Video Production Affects Student Engagement: An Empirical Study of MOOC Videos // *Proceedings of the First ACM Conference on Learning and Scale* (Atlanta, GA, 4–5 March 2014). New York: ACM. P. 41–50.
24. Hollands F. M., Tirthali D. (2014) MOOCs: Expectations and Reality. Full Report. Center for Benefit-Cost Studies of Education, Teachers College, Columbia University, NY. http://cbcse.org/wordpress/wp-content/uploads/2014/05/MOOCs_Expectations_and_Reality.pdf
25. Jansen D., Schuwer R. (2015) Institutional MOOC Strategies in Europe. Status Report Based on a Mapping Survey Conducted in October—December 2014. http://www.eadtu.eu/documents/Publications/OEenM/Institutional_MOOC_strategies_in_Europe.pdf
26. Jasnani P. (2013) Designing MOOCs. A White Paper on Instructional Design for MOOCs. http://www.tatainteractive.com/pdf/Designing_MOOCs-A_White_Paper_on_ID_for_MOOCs.pdf
27. Johnson D. (2018) Driving Performance & Persistence: How ASU is Improving Learner Outcomes with an Active Adaptive Approach. <https://youtu.be/ASekB3jEIBs>
28. Kolb D. A. (1985) *Learning Style Inventory*. Technical Manual. Boston: McBer and Company.
29. Kop R., Fournier H., Mak S. F. J. (2011) A Pedagogy of Abundance or a Pedagogy to Support Human Beings? Participant Support on Massive Open Online Courses // *The International Review of Research in Open and Distance Learning*. Vol. 12. No 7. P. 74–93.
30. Larionova V., Brown K., Bystrova T., Sinitsyn E. (2018). Russian Perspectives of Online Learning Technologies in Higher Education: An Empirical Study of a MOOC // *Research in Comparative and International Education*. Vol. 13. No 1. P. 70–91.
31. Lisitsyna L., Lyamin A. (2014) Approach to Development of Effective E-Learning Courses // *Smart Digital Futures*. Vol. 262. P. 732–738.
32. Lundvall B. A., Borrás S. (1997) *The Globalising Learning Economy: Implications for Innovation Policy*. Report Based on Contributions from Seven Projects under the TSER Programme DG XII. www.globelicsacademy.org/2011_pdf/Lundvall%20Borrass%201997.pdf
33. Nonaka I., Takeuchi H. (2011) *The Company Is the Creator of Knowledge. Origin and Development of Innovations in Japanese Firms*. M.: Olimp-Business.
34. O'Farrell L. (2017) Using Learning Analytics to Support the Enhancement of Teaching and Learning in Higher Education // *National Forum for the Enhancement of Teaching and Learning in Higher Education*. https://www.teachingandlearning.ie/wp-content/uploads/2018/01/Final_LA-Briefing-Paper_Web-with-doi.pdf
35. Sancho T., de Vries F. (2013) Virtual Learning Environments, Social Media and MOOCs: Key Elements in the Conceptualisation of New Scenarios in Higher Education // *Open Learning*. Vol. 28. No 3. P. 166–170.
36. Sharkey M., Ansari M. (2014). Deconstruct and Reconstruct: Using Topic Modeling on an Analytics Corpus // *LAK Data Challenge*. http://ceur-ws.org/Vol-1137/lakdatachallenge2014_submission_1.pdf

37. Tyler-Smith K. (2006) Early Attrition among First Time E-Learners: A Review of Factors that Contribute to Drop-Out, Withdrawal and Non-Completion Rates of Adult Learners Undertaking E-Learning Programmes // Journal of Online Learning Teaching. Vol. 2. No 2. P. 73–85.
38. Usha Keshavamurthy, Guruprasad H. S. (2014) Learning Analytics: A Survey // International Journal of Computer Trends and Technology (IJCTT). Vol. 18. No 6. P. 260–264.

Learning Analytics in Massive Open Online Courses as a Tool for Predicting Learner Performance

Tatiana Bystrova

Doctor of Sciences in Philosophy, Professor at Ural Institute for the Humanities.
E-mail: tatiana.bystrova@urfu.ru

Authors

Viola Larionova

Candidate of Sciences in Mathematical Physics, Associate Professor, Deputy Provost, Head of an academic department, Graduate School of Economics and Management. E-mail: v.a.larionova@urfu.ru

Evgueny Sinitsyn

Doctor of Sciences in Mathematical Physics, Professor, Graduate School of Economics and Management. E-mail: e.v.sinitcyn@urfu.ru.

Alexander Tolmachev

Senior Lecturer, Graduate School of Economics and Management. E-mail: avtolmachev@urfu.ru

Ural Federal University named after the first President of Russia B. N. Yeltsin & Address: 19 Mira Str., 620002 Ekaterinburg, Russian Federation.

Learning analytics in MOOCs can be used to predict learner performance, which is critical as higher education is moving towards adaptive learning. Interdisciplinary methods used in the article allow for interpreting empirical qualitative data on performance in specific types of course assignments to predict learner performance and improve the quality of MOOCs. Learning analytics results make it possible to take the most from the data regarding the ways learners engage with information and their level of skills at entry. The article presents the results of applying the proposed learning analytics algorithm to analyze learner performance in specific MOOCs developed by Ural Federal University and offered through the National Open Education Platform.

Abstract

massive open online courses, learning analytics, empirical evidence, online learning, assessment tools, checkpoint assignments, academic performance monitoring.

Keywords

- Astratova G., Sinicin E., Toporkova E., Frishberg L., Karabanova I. (2017) Mechanism of Information Model Development for Company Brand Assessment within Marketing Strategy. Proceedings of the *International Conference on Trends of Technologies and Innovations in Economic and Social Studies (Tomsk, 28–30 June 2017)*, Tomsk: Atlantis, pp. 20–25.
- Brown K., Lally V. (2017) It Isn't Adding Up: The Gap between the Perceptions of Engineering Mathematics Students and Those Held by Lecturers in the First Year of Study of Engineering. Proceedings of the *10th Annual International Conference of Education, Research and Innovation—ICERI2017 (Seville, Spain, 16–18 November 2017)*, Valencia: IATED, pp. 317–321.
- Castañó Muñoz J., Punie Y., Inamorato dos Santos A., Mitic M., Morais R. (2016) *How Are Higher Education Institutions Dealing with Openness? A Survey of Practices, Beliefs and Strategies in Five European Countries*. Institute for Prospective Technological Studies. JRC Science for Policy Report, EUR27750EN. Available at: <http://publications.jrc.ec.europa.eu/repository/bitstream/JRC99959/lfn27750enn.pdf> (accessed 2 October 2018).

References

- Crocker L., Algina J. (2010) *Vvedenie v klassicheskuyu i sovremennuyu teoriyu testov* [Introduction to Classical and Modern Test Theory]. Moscow: Logos.
- Deev M. V., Glotova T. V., Krevskiy I. G. (2015) Individualized Learning Trajectories Using Distance Education Technologies. *Creativity in Intelligent, Technologies and Data Science. Ser. Communications in Computer and Information Science*, part XI, vol. 535, pp. 778–792.
- European Association of Distance Teaching Universities (2018) *The 2018 OpenupEd Trend Report on MOOCs*. Available at: https://www.openuped.eu/images/Publications/The_2018_OpenupEd_trend_report_on_MOOCs.pdf (accessed 2 October 2018).
- Freitas S. I., Morgan J., Gibson D. (2015) Will MOOCs Transform Learning and Teaching in Higher Education? Engagement and Course Retention in On-line Learning Provision. *British Journal of Educational Technology*, vol. 46, no 3, pp. 455–471.
- Gryaznova Y., Mukovozov O. (2016) *Pilotnoe issledovanie RASO “Kak pokolenie Z vosprinimaet informatsiyu”* [Pilot Study by the Russian Public Relations Association: How Generation Z Perceives Information]. Paper presented at 4th International Applied Research Conference on Communication in Sociology, Humanities, Economics and Education (Minsk, Belarus, 7–9 April, 2016), Minsk: Belarusian State University.
- Guo P. J., Kim J., Rubin R. (2014) How Video Production Affects Student Engagement: An Empirical Study of MOOC Videos. *Proceedings of the First Conference on Learning and Scale (Atlanta, GA, 4–5 March 2014)*, New York: ACM, pp. 41–50.
- Hollands F. M., Tirthali D. (2014) *MOOCs: Expectations and Reality. Full Report*. Available at: http://cbcse.org/wordpress/wp-content/uploads/2014/05/MOOCs_Expectations_and_Reality.pdf (accessed 2 October 2018).
- Jansen D., Schuur R. (2015) *Institutional MOOC Strategies in Europe. Status Report Based on a Mapping Survey Conducted in October-December 2014*. Available at: http://www.eadtu.eu/documents/Publications/OEenM/Institutional_MOOC_strategies_in_Europe.pdf (accessed 2 October 2018).
- Jasnani P. (2013) *Designing MOOCs. A White Paper on Instructional Design for MOOCs*. Available at: http://www.tatainteractive.com/pdf/Designing_MOOCs-A_White_Paper_on_ID_for_MOOCs.pdf (accessed 2 October 2018).
- Johnson D. (2018) *Driving Performance & Persistence: How ASU is Improving Learner Outcomes with an Active Adaptive Approach*. Available at: <https://youtu.be/ASekB3jEIBs> (accessed 2 October 2018).
- Kolb D. A. (1985) *Learning Style Inventory. Technical Manual*. Boston: McBer and Company.
- Kop R., Fournier H., Mak S. F. J. (2011) A Pedagogy of Abundance or a Pedagogy to Support Human Beings? Participant Support on Massive Open Online Courses. *The International Review of Research in Open and Distance Learning*, vol. 12, no 7, pp. 74–93.
- Korn G., Korn T. (1973) *Spravochnik po matematike dlya nauchnykh rabotnikov i inzhenerov* [Mathematical Handbook for Scientists and Engineers]. Moscow: Chief Editorial Board for Physics and Mathematics Literature.
- Larionova V., Brown K., Bystrova T., Sinitsyn E. (2018) Russian Perspectives of Online Learning Technologies in Higher Education: An Empirical Study of a MOOC. *Research in Comparative and International Education*, vol. 13, no 1, pp. 70–91.
- Larionova V., Tretyakov V. (2016) Otkrytye onlayn-kursy kak instrument modernizatsii obrazovatelnoy deyatel'nosti v vuse [Open Online Courses as a Tool

- for Modernization of Educational Process in Universities. *Higher Education in Russia*, no 7, pp. 55–66.
- Leskovec J., Rajaraman A., Ullman J. (2016) *Analiz bolshikh naborov dannykh* [Mining of Massive Datasets]. Moscow: DMK Press.
- Lisitsyna L., Lyamin A. (2014) Approach to Development of Effective E-Learning Courses. *Smart Digital Futures*, vol. 262, pp. 732–738.
- Lundvall B. A., Borrás S. (1997) *The Globalising Learning Economy: Implications for Innovation Policy. Report Based on Contributions from Seven Projects under the TSER Programme DG XII*. Available at: www.globelicsacademy.org/2011_pdf/Lundvall%20Borras%201997.pdf (accessed 2 October 2018).
- Maksimov Y. (ed.) (2001) *Veroyatnostnye razdely matematiki* [Probability-Related Areas of Mathematics]. Saint Petersburg: Ivan Fedorov.
- Mayorov A. (2002) *Teoriya i praktika sozdaniya testov dlya sistemy obrazovaniya* [Theory and Practice of Developing Tests for the Education System]. Moscow: Intellekt-Tsentr.
- Nekhaev I. (2016) Analiz kachestva protsessa obucheniya s ispolzovaniem on-layn-kursov [Analyzing the Quality of Learning with MOOCs]. Proceedings of the *Best Practices in Digital Learning: 2nd Methodological Conference (Tomsk, 26–27 May 2016)*, Tomsk: Tomsk State University, pp. 8–14.
- Netology Group (2017) *Issledovanie rossiyskogo rynka onlayn-obrazovaniya i obrazovatelnykh tekhnologiy* [Research in Russia's Online Learning Market and Education Technology]. Available at: <http://edumarket.digital> (accessed 2 October 2018).
- Nonaka I., Takeuchi H. (2011) *The Company Is the Creator of Knowledge. Origin and Development of Innovations in Japanese Firms*. Moscow: Olimp-Business.
- O'Farrell L. (2017) *Using Learning Analytics to Support the Enhancement of Teaching and Learning in Higher Education*. Paper presented at National Forum for the Enhancement of Teaching and Learning in Higher Education. Available at: https://www.teachingandlearning.ie/wp-content/uploads/2018/01/Final_LA-Briefing-Paper_Web-with-doi.pdf (accessed 2 October 2018).
- Roshchina J., Roshchin S., Rudakov V. (2018) Spros na massovye otkrytye on-layn-kursy (MOOC): opyt rossiyskogo obrazovaniya [The Demand for Massive Open Online Courses (MOOC): Evidence from Russian Education]. *Voprosy obrazovaniya / Educational Studies Moscow*, no 1, pp. 174–199. doi: 10.17323/1814-9545-2018-1-174-199.
- Samokhin I., Sergeyeva M., Sokolova N., Marchenko E. (2018) Soderzhanie ponyatiya “effektivnost obrazovaniya” v kontekste inkluzivnykh tendentsiy sovremenniy shkoly [Concept of “Educational Efficiency” in Context of Inclusive Trends of Modern School]. *Nauchnyy dialog*, no 1, pp. 278–288.
- Sancho T., de Vries F. (2013) Virtual Learning Environments, Social Media and MOOCs: Key Elements in the Conceptualisation of New Scenarios in Higher Education. *Open Learning*, vol. 28, no 3, pp. 166–170.
- Semenova T., Vilkova K., Shcheglova I. (2018) Rynok MOOK: perspektivy dlya Rossii [The MOOC Market: Prospects for Russia]. *Voprosy obrazovaniya / Educational Studies Moscow*, no 2, pp. 173–197. doi: 10.17323/1814-9545-2018-2-173-197.
- Sharkey M., Ansari M. (2014). Deconstruct and Reconstruct: Using Topic Modeling on an Analytics Corpus. *LAK Data Challenge*. Available at: http://ceur-ws.org/Vol-1137/lakdatachallenge2014_submission_1.pdf (accessed 2 October 2018).
- Shmelev A. (2013) *Prakticheskaya testologiya: testirovanie v obrazovanii, prikladnoy psikhologii i upravlenii personalom* [Testing in Education, Applied Psychology and Human Resource Management]. Moscow: Maska.

- Tyler-Smith K. (2006) Early Attrition among First Time E-Learners: A Review of Factors that Contribute to Drop-Out, Withdrawal and Non-Completion Rates of Adult Learners Undertaking E-Learning Programmes. *Journal of Online Learning Teaching*, vol. 2, no 2, pp. 73–85.
- Usha Keshavamurthy, Guruprasad H. S. (2014) Learning Analytics: A Survey. *International Journal of Computer Trends and Technology (IJCTT)*, vol. 18, no 6, pp. 260–264.
- Vishnyakova S. (1999) *Professionalnoe obrazovanie: Slovar. Klyuchevye ponyatiya, terminy, aktualnaya leksika* [Vocational Education: Vocabulary. Keywords, Terminology, Language]. Moscow: Research and Methodology Center for Secondary Vocational Education.
- Zagvyazinsky V., Zakirova A., Strokova T. et al. (2008) *Pedagogicheskiy slovar* [Pedagogical Vocabulary]. Moscow: Akademiya.
- Zvonnikov A., Chelyshkova M. (2012) *Kontrol kachestva rezultatov obucheniya pri attestatsii (kompetentnostny podkhod). Ucheb. Posobie* [Monitoring Educational Quality in Teacher Appraisal (Competency-Based Approach). Study Guide]. Moscow: Logos.