

Учебная аналитика MOOK как инструмент анализа математической тревожности

Ю. Ю. Дюличева

Статья поступила
в редакцию
в августе 2021 г.

Дюличева Юлия Юрьевна — кандидат физико-математических наук, доцент ФГАОУ ВО «Крымский федеральный университет имени В. И. Вернадского». Адрес: 295007, Симферополь, просп. Академика Вернадского, 4. E-mail: dyulichева_yu@mail.ru

Аннотация

Исследование посвящено извлечению описаний математической тревожности из отзывов на массовые открытые онлайн-курсы по математике (MOOK) с помощью методов анализа текстовых данных. Эмоциональные состояния обучающихся, связанные с математической фобией, являются серьезным препятствием в изучении математики и получении базовых математических знаний, необходимых для будущей профессиональной деятельности. На платформах MOOK накапливаются большие объемы данных, среди которых отзывы на онлайн-курсы представляют особый интерес. Эмпирическую основу исследования составили материалы 38 онлайн-курсов по математике на *Udemy* и 1898 отзывов обучающихся. Применение алгоритма анализа тональности VADER, кластерного анализа текстов отзывов с негативной тональностью на основе метода *kMeans* и векторного представления предложений с помощью модели представления языка BERT позволило выделить кластеры с описанием различных отрицательных эмоций, связанных с прошлым фрустрирующим опытом при изучении математики, кластер с описанием сожалений в связи с упущенными возможностями из-за негативного отношения к математике, а также кластер с описанием постепенного преодоления математической тревожности в процессе изучения онлайн-курсов по математике. Построенный граф знаний позволил визуализировать некоторые закономерности, связанные с различными отрицательными эмоциями, которые возникали у обучающихся при изучении математики.

Ключевые слова

массовые открытые онлайн-курсы (MOOK), математическая тревожность, учебная аналитика, VADER, BERT, анализ текстов.

Для цитирования

Дюличева Ю. Ю. (2021) Учебная аналитика MOOK как инструмент анализа математической тревожности // Вопросы образования / Educational Studies Moscow. № 4. С. 243–265. <https://doi.org/10.17323/1814-9545-2021-4-243-265>

Learning Analytics in MOOCs as an Instrument for Measuring Math Anxiety

Y. Y. Dyulicheva

Yulia Y. Dyulicheva, Candidate of Sciences in Mathematical Physics, Associate Professor, V. I. Vernadsky Crimean Federal University. Address: 4 Akademika Vernadskogo Ave, 295007 Simferopol. E-mail: dyulicheva_yu@mail.ru

Abstract In this paper, math anxiety descriptions are extracted from Massive Open Online Course (MOOC) reviews using text mining techniques. Learners' emotional states associated with math phobia represent substantial barriers to learning mathematics and acquiring basic mathematical knowledge required for future career success. MOOC platforms accumulate big sets of educational data, learners' feedback being of particular research interest. Thirty-eight math MOOCs on Udemy and 1,898 learners' reviews are investigated in this study. VADER sentiment analysis, *k*-means clustering of content with negative sentiment, and sentence embedding based on the Bidirectional Encoder Representations from Transformers (BERT) language model allow identifying a few clusters containing descriptions of various negative emotions related to bad math experiences in the past, a cluster with descriptions of regrets about missed opportunities due to negative attitudes towards math in the past, and a cluster describing gradual overcoming of math anxiety while progressing through a math MOOC. The constructed knowledge graph makes it possible to visualize some regularities pertaining to different negative emotions experienced by math MOOC learners.

Keywords BERT, learning analytics, Massive Open Online Courses (MOOCs), math anxiety, text mining, VADER.

For citing Dyulicheva Y. Y. (2021) Uchebnaya analitika MOOK kak instrument analiza matematicheskoy trevozhnosti [Learning Analytics in MOOCs as an Instrument for Measuring Math Anxiety]. *Voprosy obrazovaniya / Educational Studies Moscow*, no 4, pp. 243–265. <https://doi.org/10.17323/1814-9545-2021-4-243-265>

С быстрым ростом количества массовых открытых онлайн-курсов (MOOK) и их аудитории по всему миру произошло накопление больших объемов разнородных данных об обучающихся — от возраста, статуса и начального уровня знаний до динамики и успешности прохождения курсов. Дополнительные данные могут быть получены на начальном и заключительном этапах обучения с помощью опросников, промежуточного и итогового тестирования, а также в процессе обучения при сборе данных о просмотрах и последовательности изучения видеофрагментов, потоке кликов обучающихся, количестве скачиваний учебного материала и т. п.

Взаимодействие инструктора с обучающимися в процессе изучения MOOK ограничено, свое отношение к содержанию онлайн-курса и инструктору обучающиеся могут проявить в отзывах о курсах и преподавателях, а также в комментариях

ях на форумах, в мессенджерах и социальных сетях. На основе этих отзывов выросло новое направление учебной аналитики — анализ текстовых образовательных данных.

Массовые открытые онлайн-курсы по математике часто разрабатываются с целью повышения уровня базовых знаний по математике, например проект бесплатного онлайн-курса для взрослых с низким уровнем знаний по математике и статистике с изложением требований, которым должен удовлетворять универсальный онлайн-курс с доступным объяснением основных математических понятий [Griffiths et al., 2019]. MOOK по математике создаются в качестве дополнительной поддержки студентов с низким уровнем математических знаний и высоким уровнем стресса и дискомфорта, который возникает при изучении очных курсов по математике, а также в качестве инструмента для обновления знаний по математике после каникул [Lambert, 2015]. MOOK по математике могут применяться для совершенствования профессиональных навыков и как сеть для обмена эффективными методиками преподавания в педагогическом сообществе [Taranto, Robutti, Arzarello, 2020], а также могут подстраиваться под нужды определенных целевых групп: тех, кто стремится обновить знания по математике; тех, кому необходимы базовые знания; учителей и обучающихся, осваивающих различные методики преподавания математики [Soares, Lopes, 2016]. Особую группу пользователей MOOK по математике составляют обучающиеся с глубоким нарушением зрения и слепотой [Косова, Изетова, 2020].

В последние годы наблюдается особый интерес со стороны научного сообщества к задачам учебной аналитики MOOK, о чем свидетельствует появление датасетов с данными MOOK на платформе анализа данных и проведение соревнований по машинному обучению *Kaggle* (*kaggle.com*). Например, на платформе *Kaggle* представлен датасет с данными ЭЭГ головного мозга студентов, которые изучали видеоконтент MOOK, датасет с данными о видеолекциях MOOK и т. п.

Основные направления учебной аналитики MOOK

К основным направлениям учебной аналитики MOOK относятся анализ тональности текстовых данных MOOK, выявление целевых групп слушателей MOOK, аналитика на основе анализа потока кликов и ресурсов MOOK, оценивание качества MOOK, прогнозирование отсева/оттока слушателей MOOK, разработка систем рекомендаций по выбору MOOK и их контента.

Анализ тональности текстовых данных MOOK часто используется для изучения отношений, мнений или настроений обучающихся, для чего выявляются закономерности в текстовых

данных и оценивается их сентимент. Полученные с помощью анализа тональности отзывов обучающихся данные способствуют выявлению причин отсеивания и снижения интереса обучающихся при изучении MOOK, разработке стратегий улучшения учебного контента MOOK и методик обучения, а также пониманию поведения студентов. Изучение данных социальных сетей — профилей и комментариев — дает представление об эмоциях и настроении обучающихся при изучении MOOK. Для анализа комментариев на форумах и отзывов обучающихся на онлайн-курсы используются иерархические рекуррентные нейронные сети [Caruano et al., 2020]; отзывы на MOOK с учетом двух полярностей (положительные/отрицательные) анализируются на основе модели ALBERT-BiLSTM с тремя слоями: эмбединг-слой, семантический слой и выходной слой [Wang, Huang, Zhou, 2021]; тональность сообщений дискуссионных форумов MOOK и ее влияние на отчисление обучающихся оцениваются с применением методов анализа выживаемости [Wen, Yang, Rose, 2014]; извлечь скрытые закономерности, которые возникают в результате взаимодействия между участниками форума MOOK, и построить модель распространения знаний можно на основе анализа социальных сетей, когортного анализа и выделения групп обучающихся с учетом их активности на форумах [Sinha, 2014]; соответствие между сообщениями студентов на форумах и учебным контентом MOOK устанавливается на основе тематического моделирования [Wong, Wong, Hindle, 2019]. Еще одной важной задачей учебной аналитики текстовых образовательных данных является идентификация учебных ресурсов (тегирование) с помощью анализа сообщений дискуссионного форума, ее решение основано на применении архитектуры теггеров LSTM-CRF [An et al., 2018].

Выявление целевых групп слушателей MOOK может помочь инструктору MOOK разработать стратегии для более эффективного взаимодействия с аудиторией онлайн-курса. Для решения задач учебной аналитики такого типа часто привлекается кластерный анализ. Удовлетворенность онлайн-курсами зависит от настроений целевых групп обучающихся, которые выявляются с помощью алгоритма VADER. В частности, показано, что новички, как правило, удовлетворены содержанием MOOK, а более опытные слушатели ожидают углубленного рассмотрения тем [Lundqvist, Liyanagunawardena, Starkey, 2021]. Еще одним критерием для выделения целевых групп являются различия между обучающимися при просмотре последовательности видеофрагментов MOOK, на их основе осуществляется персонализация учебного контента [Zhang, Liu, Liu, 2020].

Аналитика на основе анализа потока кликов и ресурсов MOOK дает возможность выявить сложные или скучные фрагменты учебного контента и разработать стратегии персонализации учебного контента MOOK с целью повышения качества учебных материалов и ресурсов. К задачам учебной аналитики этого направления относится прогнозирование с помощью рекуррентных нейронных сетей последовательности, в которой обучающиеся будут просматривать ресурсы MOOK в следующий раз, на основе анализа предыдущих последовательностей просмотров [Tang, Peterson, Pardos, 2016]; построение модели поведения обучающихся на основе анализа последовательностей просмотра видеоконтента и потоков кликов, а также таких характеристик, как длительность просмотра учебных видеофрагментов и показатель успешности ответов на вопросы викторины [Brinton et al., 2015]. На основании полученных данных можно построить индивидуальные траектории для просмотра учебного видеоконтента и разработать персонифицированные стратегии для предложения учебных ресурсов MOOK.

Оценивание качества MOOK, разработка обоснованных критериев такого оценивания — сложная задача, которая решается, в частности, средствами учебной аналитики. Для изучения поведения обучающихся и их отношения к представленному на онлайн-курсах учебному контенту используется, например, исследование видеотрафика, сообщений на форумах и отслеживается количество полученных сертификатов об окончании MOOK [Luo et al., 2018].

Прогнозирование отсева/оттока слушателей MOOK. С целью разработки стратегий, направленных на удержание обучающихся, контролируются изменения заинтересованности слушателей в процессе изучения онлайн-курсов. Один из подходов к решению этой задачи — анализ активности обучающихся и своевременное реагирование на отсутствие обратной связи со стороны слушателей курса. Активность обучающихся прогнозируют, в частности, на основе анализа еженедельной активности с помощью нейронных сетей с архитектурой RNN-LSTM [Liu et al., 2018]. Перспективы оттока обучающихся оценивают на основе анализа потока кликов при просмотре учебного видеоконтента [Jeon, Park, 2020]. Учебную аналитику используют для прогнозирования успешности обучающихся [Быстрова и др., 2018].

Разработка систем рекомендаций по выбору MOOK и их контента. Количество предлагаемых MOOK постоянно растет, выбрать курс, соответствующий личным потребностям и доступный по цене, становится все труднее. Одним из подходов к разра-

ботке рекомендаций является построение на основе больших данных распределенного дерева курсов, позволяющего выделять предпочтительные области курсов для поиска оптимального решения [Hou et al., 2016].

Разработка алгоритмов для анализа больших данных в сфере образования привела к появлению аналитических инструментов, в том числе для оценки MOOK. Например, MOOCviz—это платформа для визуализации данных edX и аналитики данных Coursera. MOOCdb содержит базы данных, в которых представлены сведения об обучающихся, их активности и полученной от них обратной связи. Авторы MOOCviz реализовали когортный и статистический анализ, а также использовали различные эвристические подходы для выявления когорт слушателей MOOK с учетом используемых учебных ресурсов, стран проживания обучающихся и т. п. [Dernoncourt et al., 2013]. Примерами аналитических инструментов для анализа MOOK являются *PerspectivesX*, предназначенный для управления контентом и анализа учебной деятельности с учетом данных о взаимодействии с обучающимися на основе методов тематического моделирования и глубокого обучения, и *MessageLens* для визуальной аналитики тем MOOK [Bakharia, 2017; Wong, Zhang, 2018].

Среди научных публикаций в области учебной аналитики MOOK автору не удалось найти исследований состояния тревожности у слушателей MOOK по математике на основе анализа отзывов и комментариев. Целью статьи является разработка методологии для выявления математической тревожности на основе анализа отзывов слушателей массовых открытых онлайн-курсов по математике с помощью методов машинного обучения.

2. Подходы к анализу тревожностей

2.1. Математическая тревожность

Математическая тревожность — это серьезная проблема, связанная с фрустрирующим опытом изучения математики. Усиление от урока к уроку переживания трудности, неуспешности, разочарования способно вызвать чувство неприязни, страх, беспокойство, постоянный стресс и нежелание в будущем выбирать профессию, требующую математических знаний и навыков. Под математической тревожностью понимают особое эмоциональное состояние обучающегося, вызывающее такие сильные эмоции, как ненависть, отвращение, и приводящее к избеганию обучающимся любых вопросов, связанных с математикой [Ashcraft, Moore, 2009]. Математическая тревожность широко распространена и возникает даже у обучающихся на инженерных специальностях, если они испытывают трудности при освоении базовых концепций [Ma, 1999]. Некоторые исследователи даже квалифицируют математическую тревожность как клиническую патологию, которая приводит к ухудшению когни-

тивных навыков, замкнутости, неблагоприятным эмоциональным состояниям даже при высокой успеваемости обучающегося в целом [Stella, 2021].

Эмпирически показана зависимость между низкой успеваемостью по математике и наличием математической тревожности [Ashcraft, Moore, 2009; Ma, 1999]. Дополнительным фактором, усиливающим математическую тревожность, может быть переход от начального образования к среднему, а также переход из школы в колледж или университет, связанный с изменениями условий обучения [Field et al., 2021].

Для измерения математической тревожности и оценки дискомфорта, который испытывают люди при изучении математики, разработаны опросники, например шкала AMAS (*Abbreviated Math Anxiety Scale*) для измерения тревожности у старших подростков и взрослых и ее модификации — шкалы mAMAS и EES-AMAS для оценки математической тревожности у детей. Опросники основаны на оценивании эмоциональной реакции на различные задания, связанные с математикой [Carey et al., 2017; Primi et al., 2020].

Методы предотвращения математической тревожности основаны на подмене отрицательных эмоций положительными, что достигается через опыт и целенаправленную работу по адаптации, через применение педагогических инструментов и инновационных методов с привлечением дополненной и виртуальной реальности, с использованием системы подсказок для поддержки обучающегося в процессе решения задач по математике [Дюличева, 2020].

Учитывать природу математической тревожности и ее причины особенно важно при разработке MOOK по математике, поскольку в них отсутствует непосредственный контакт инструкторов с обучающимися. О распространенности проблемы математической тревожности можно косвенно судить по названиям некоторых MOOK, например на *Udemy* один из англоязычных онлайн-курсов по математике для начального уровня называется «Исчисление для того, кто ненавидит исчисление», для среднего уровня — «Свободная от стресса статистика для IBDP/AP студентов: Мини-курс 2».

При этом исследования, нацеленные на подбор методов обучения, дизайна и стиля MOOK с таким расчетом, чтобы предотвратить возникновение математической тревожности в результате онлайн-взаимодействия, практически отсутствуют, а также открытым остается вопрос об оценке математической тревожности, возникающей при изучении MOOK, поскольку слушатели онлайн-курсов неохотно участвуют в дополнительных опросах.

2.2. Выявление тревожностей на основе анализа текстовых данных

Анализ тональности текстовых данных (комментариев, отзывов, контента персональных страниц в социальных сетях) эффективно используется не только для оценки распространенности негативных настроений в обществе, но и для выявления депрессивных состояний, тревожности и других психических расстройств.

Депрессивные состояния выявляются на основании текстовых данных, представленных в социальных медиа, с помощью модели BERT для анализа тональности текстов и колл-грамм, для лексического анализа текстов с вычислением таких показателей, как коэффициент взаимной информации, Т-оценка и количество идиосинкразических единиц, характеризующих болезненную реакцию в сообщениях на некоторые раздражения [Wolk, Chlasta, Holas, 2021]. Тревожность, возникшую в обществе в связи с пандемией COVID-19, оценивали на основе анализа комментариев к видео на *Youtube* с помощью различных подходов к векторизации текста (TF-IDF, «мешок слов») и методов машинного обучения (SVM, «случайный лес», бустинг и др.) [Saifullah, Fauziah, Aribowo, 2021]. Комментарии в социальных медиа анализировали с помощью модели представления языка RoBERTa с архитектурой «Трансформер», нейронных сетей LSTM и модели BERT для решения задачи распределения представленных в комментариях расстройств на пять классов: депрессия, тревога, биполярное расстройство, синдром дефицита внимания и гиперактивности, посттравматическое стрессовое расстройство [Murarka, Radhakrishnan, Ravichandran, 2020]. Депрессивные и тревожные состояния выявлялись на основании 4500 твитов при помощи SVM с различными способами векторизации текстов и моделей представления языка BERT, ALBERT [Owen, Camacho-Collados, Espinosa-Anke, 2020].

Опираясь на опыт указанных выше исследований [Wolk, Chlasta, Holas, 2021; Murarka, Radhakrishnan, Ravichandran, 2020; Stella, 2021], мы предлагаем методологию для анализа состояний математической тревожности на основе выявления отзывов на MOOK по математике с негативной тональностью, кластеризации этих отзывов и визуализации выявленных состояний с помощью графа знаний.

3. Датасет и методология исследования

В настоящее время на платформах *Coursera*, *Udemy*, *EdX* представлено множество массовых открытых онлайн-курсов по математике с простейшими критериями сортировки: по рейтингу, цене, продолжительности видео, уровню, наличию субтитров и т.п. Количество англоязычных MOOK по математике на этих платформах, найденных по запросу с ключевым словом *math*, представлено на рис. 1. Наибольшее количество MOOK по ма-

тематике для начального уровня и наименьшее — для продвинутого размещено на *Udemy*.

Для каждого уровня мы проанализировали названия онлайн-курсов и построили облака наиболее часто встречающихся слов, представленные в табл. 1. Этапу построения облаков наиболее часто встречающихся слов с помощью *Python*-библиотеки *Wordcloud* предшествовал этап предобработки названий онлайн-курсов, включающий удаление пунктуации и стоп-слов.

Мы исследовали две категории англоязычных онлайн-курсов начального уровня на *Udemy*: 27 курсов по математике, в названиях которых встречались ключевые слова «фундаментальный», «основы», «стать мастером», и 11 курсов, в названиях которых содержались ключевые слова «ментальный» и «ведический». На следующем этапе был выполнен скрапинг отзывов на курсы рассматриваемых категорий: собраны 1326 отзывов от уникальных пользователей на курсы категории «Основы математики» и 572 отзыва от уникальных пользователей на курсы категории «Ментальная/ведическая математика».

Как видно из табл. 1, построенной на основе частотного анализа, на онлайн-курсах по математике начального уровня на *Udemy* уделяется большое внимание изучению основ алгебры, статистики, исчислению, тригонометрии, теории вероятностей и теории графов, а также ментальной математике и основам математики для освоения машинного обучения. *Coursera* предоставляет онлайн-курсы по математике, направленные на определенную специализацию обучающегося, а также на изучение языков программирования, машинного обучения, в частности глубокого обучения, и анализ данных. На EdX онлайн-курсы по математике начального уровня больше ориентированы на введение в машинное обучение, квантовую и классическую механику, среднего уровня — на фундаментальные курсы по математике (матричная алгебра, линейная алгебра, дифференциальное исчисление и т.п.), механике и электронике, продвинутого уровня — на квантовое исчисление и решение прикладных задач.

Основные этапы анализа отзывов на онлайн-курсы *Udemy* по математике и выявления математической тревожности из негативных отзывов показаны на рис. 2.

На первом этапе выполняется предобработка отзывов с удалением знаков пунктуации, стоп-слов, приведением к нижнему регистру и токенизацией средствами *Python*-библиотеки *NLTK*.

На втором этапе проводится анализ тональности отзывов с помощью алгоритма *VADER*. Тональность измеряется с помощью методов, основанных на правилах и словарях, или на основе методов машинного обучения. С помощью алгоритма *VADER* на основе правил и тональных словарей отзыв получает четы-

Рис. 1. Распределение математических МОК по уровням знаний, %

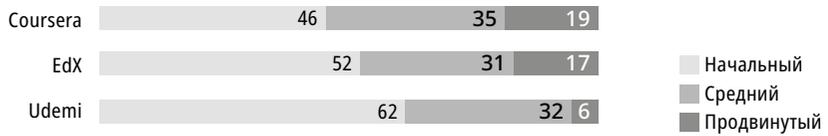


Таблица 1. Облака слов, характеризующие MOOK разных уровней по математике

MOOK	Уровень		
	Начальный	Средний	Продвинутый
Udemy			
Coursera			
EdX			

Рис. 2. Основные этапы анализа отзывов на MOOK по математике



ре оценки: позитивную, нейтральную, негативную и составную. Преимуществом применения алгоритма VADER является то, что в нем нет необходимости использовать датасет для обучения алгоритма, а недостатком — отсутствие учета слов, которых нет в словаре тональности, при вычислении итоговой оценки тональности. Для итогового оценивания тональности отзыва используется составная оценка, которая получается как нормализованная сумма баллов валентностей, вычисленная на основе некоторой эвристики и лексикона настроений. Нормализация суммы валентностей в диапазоне от -1 (крайне отрицательная тональность) до $+1$ (крайне положительная тональность) выполняется по формуле [Hutto, Gilbert, 2014; Adarsh et al., 2019]:

$$\text{compoundScore} = \frac{x}{\sqrt{x^2 + \alpha}},$$

где α — параметр, по умолчанию равный 15; x — сумма оценок тональностей всех слов фразы (отзыва).

На третьем этапе выделяются предложения с негативной тональностью, которые не содержат ключевых слов, описывающих текущий онлайн-курс и/или инструктора. Для этого используются кастомные словари, содержащие слова «курс», «уроки» и т. п. и их синонимы, слова «инструктор», «тьютор» и т. п., а также их синонимы и имена инструкторов (например, Криста Кинг).

На четвертом этапе близкие по семантике предложения с описанием математической тревожности группируются с помощью модели представления языка BERT, алгоритма кластеризации *kMeans* и метода главных компонент.

Преобразование текста отзывов в плотные числовые векторы выполняется с помощью предварительно обученной модели представления языка BERT, которая основана на использовании двунаправленной нейронной сети-кодировщика с архитектурой «Трансформер» [Devlin et al., 2019]. BERT демонстрирует высокую точность и производительность на небольших датасетах. После получения векторных представлений предложений с негативной тональностью определяется оптимальное количество кластеров на основе метода локтя и применяется метод *kMeans*, идея которого заключается в следующем.

1. На вход подается оптимальное число кластеров k и случайно выбирается k векторных представлений с описанием математической тревожности в качестве первоначальных центров кластеров.
2. Каждому векторному представлению предложения с описанием математической тревожности присваивается номер кластера, определяемый на основании наименьшего евкли-

Рис. 3. Исходное распределение отзывов на онлайн-курсы по математике с учетом их тональности (негативная, нейтральная, позитивная) для двух категорий онлайн-курсов — «Основы математики» и «Ментальная/ведическая математика»



дова расстояния от векторного представления предложения до центра кластера.

3. С учетом распределения по кластерам пересчитываются координаты центроидов как среднее значение всех векторных представлений, принадлежащих каждому кластеру.
4. Шаги 2 и 3 повторяются до тех пор, пока изменяются значения координат центроидов.

Применение BERT совместно с *kMeans* позволяет выделить группы предложений (кластеры) с учетом структурных особенностей языка.

На пятом этапе с помощью определения частей речи и регулярных выражений (шаблонов) выделяются сущности и отношения между ними для построения графа знаний, описывающего состояние, извлеченные из отзывов.

4. Аналитика отзывов МООК по математике и выявление математической тревожности

4.1. Анализ тональности отзывов МООК по математике

Приведем результаты оценки тональности 1326 отзывов на курсы категории «Основы математики» и 572 отзывов на курсы категории «Ментальная/ведическая математика».

Как видно из гистограммы на рис. 3, отзывы на онлайн-курсы по математике в основном имеют позитивную тональность. Оценка тональности вычисляется с помощью алгоритма VADER [Hutto, Gilbert, 2014], реализованного в *Python*-библиотеке *VADER Sentiment*.

Таблица 2. Пример анализа тональности англоязычного отзыва на один из онлайн-курсов по математике на *Udemy* с помощью VADER

Отзыв/его часть	Оценка тональности			
	Позитивная	Нейтральная	Негативная	Составная
Я всегда презирал математику в школе, потому что учителя никогда не делали ее интересной. Я чуть не бросил старшую школу из-за того, что не посещал занятия по математике, потому что я ее ненавижу. После завершения этого курса у меня определенно есть хороший фундамент по основам, так что это был не просто отличный экскурс, но мне действительно пришлось заново выучить все концепции. Криста — потрясающий учитель, и я мог только представить, насколько она была бы замечательной, если бы вела очный класс. Этот курс наполнен замечаниями и практическими тестами, которые можно использовать в вашем обучении, и если вы пройдете этот курс, то выйдете хорошо подготовленными для того, чтобы изучить следующий уровень математики. Я планирую перейти к изучению курсов алгебры и линейной алгебры от Кристи, и, кто знает, если все пойдет хорошо, я смогу перейти к курсам по исчислению, геометрии, теории вероятностей и статистике. У Кристи действительно приятный голос, и она настолько упрощает все понятия, что даже ребенок сможет их понять. Настоятельно рекомендую и думаю, что есть причина, по которой ее курсы получают самые лучшие отзывы. Спасибо, Криста, ты классная!	0,212	0,74	0,048	0,9913
Я всегда презирал математику в школе, потому что учителя никогда не делали ее интересной. Я чуть не бросил старшую школу из-за того, что не посещал занятия по математике, потому что я ее ненавижу.	0	0,677	0,323	-0,917

Сложность задачи оценки тональности отзывов на онлайн-курсы по математике заключается в том, что отзыв в целом может получить позитивную или нейтральную оценку тональности, но при этом содержать одно или несколько предложений, описывающих фрустрирующий опыт, связанный с изучением математики. В табл. 2 приведен типичный пример такого отзыва. В первой строке таблицы представлен исходный отзыв на один из онлайн-курсов по математике на *Udemy*, во второй строке выделена часть этого отзыва, в которой нет слов, описывающих курс или инструктора. В столбцах приведены оценки тональностей для отзыва (позитивный сентимент) и его части (негативный сентимент), полученные на основе алгоритма VADER.

В дальнейшем мы сосредоточимся на анализе негативных отзывов. Из исходных датасетов с отзывами выделим предложения с негативной тональностью, которые не содержат слов

Рис. 4. Получение векторных представлений на основе BERT

```

['я всю жизнь боролся с математикой из-за отсутствия базовых знаний',
'я всегда презирал математику в школе потому что учителя никогда не делали её интересной',
'я плохо разбираюсь в математике',
'математика всегда была моей самой большой слабостью',
'получил прочную основу по математике которую в детстве / юношестве полностью пропустил',
'математика стала менее пугающей для меня',
'это еще больше усугубляет мою тревожность по поводу математики',
'математика никогда не была моим любимым предметом',
'я избегал использовать математику в магистратуре',
'...']

```

↓

```

[[-0.7948846  -0.2880064  -0.35478127  ...  -0.68020815  -0.03355978
  0.7341734  ]
 [ 0.33664528  0.4001624  -0.4852074  ...  -0.46567246  0.13488093
  0.43704456 ]
 [ 0.17897569  -0.1904699  -0.07447997  ...  -0.2082912  0.09511402
  0.33010995 ]
 ...
 [-1.105307   -0.04232654  -0.13710353  ...  0.63763756  -0.5516867
  0.7343703  ]
 [-0.83153987  0.4879239   -0.03810127  ...  -0.6294967   -0.15833776
  1.2576165  ]
 [-1.2190384   -0.8249754   -0.5068008   ...  0.21856284  0.2639225
  0.6180943  ] ]

```

из словарей, описывающих онлайн-курс и его тьютора. Получим датасет, содержащий 231 предложение с негативной тональностью в категории «Основы математики» и 93 предложения с негативной тональностью в категории «Ментальная/ведическая математика».

4.2. Кластеризация негативных отзывов на MOOK по математике

Среди предложений с негативной тональностью на основе регулярных выражений выделим предложения, которые содержат ключевые слова *math*, *mathematic* и т. п. Это 52 предложения в категории «Основы математики» и 10 предложений в категории «Ментальная/ведическая математика». Выполним кластеризацию выделенных предложений, используя векторное представление с помощью модели BERT и алгоритма *kMeans*. BERT позволяет получить плотные векторные представления предложений с описанием математической тревожности. Результат представлен на рис. 4.

Применение модели представления языка BERT совместно с алгоритмом *kMeans* позволяет выделить кластеры с учетом семантического сходства предложений [Li et al., 2020]. Для применения алгоритма кластеризации *kMeans* необходимо знать число кластеров. Оптимальное число кластеров было определено на основе метода локтя и равно 5. Распределение предложений с описанием математической тревожности по кластерам с применением метода главных компонент и метода *kMeans*

Рис. 5. Распределение предложений с описанием математической тревожности по кластерам с применением метода главных компонент, метода *kMeans* и векторного представления на основе BERT

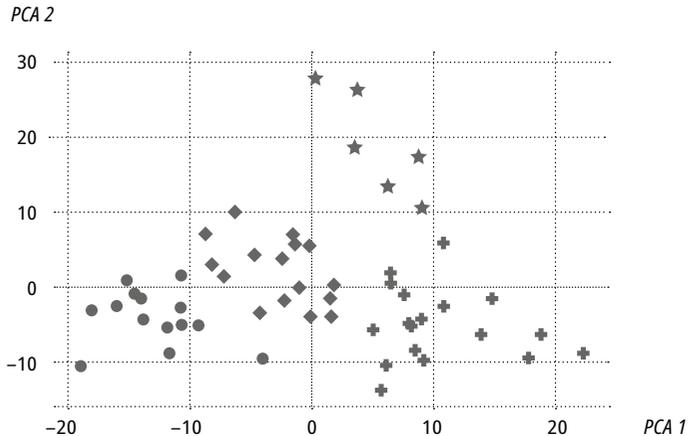


Таблица 3. Результаты кластеризации англоязычных предложений с негативной тональностью с учетом их семантической близости

Номер кластера	Мощность кластера	Пример случайного предложения из кластера	Ключевые слова кластера	Средняя составная оценка тональности кластера
1	7	Я обнаружил, что узнал довольно много новых трюков, которыми меня не научили в школе	Ментальные блоки, пробелы, школа	-0,286
2	6	Теперь математика стала менее пугающей	Менее, пугающий, запутанный, странный	-0,383
3	16	Математика всегда была моей самой большой слабостью	Слабость, враг, фобия, ненавидеть	-0,405
4	16	Каждый урок бросал мне вызов и приводил к росту математических навыков	Вызов, навыки, помогать, шаг за шагом, опыт, отрицательный, постепенно	-0,326
5	17	Очень жаль, что в мои школьные годы не было такого отличного учителя	Жаль, не было, ненавидеть, тест, школа, негативный	-0,498

представлено на рис. 5. Результаты кластеризации представлены в табл. 3.

Ключевые слова в каждом кластере выделены с помощью частотного анализа. Кластер 1 содержит предложения, описывающие успехи при изучении онлайн-курса по математике на фоне прошлого фрустрирующего опыта освоения этой дис-

циплины, в них используются такие выражения, как «ментальные блоки», «пробелы в школьных знаниях».

Кластер 2 включает описания снижения математической тревожности, например с помощью выражений «менее пугающая», «менее запутанная» по отношению к математике.

Кластер 3 объединил предложения, в которых выражены сильные эмоции, связанные с проблемами в математике, например «математика всегда была моей самой большой слабостью», «математика всегда была моим врагом», «я всегда страстно ненавижу математику».

Предложения в кластере 4 описывают постепенное преодоление проблем, связанных с изучением математики, на онлайн-курсах через погружение в математику шаг за шагом, путем постепенного приобретения навыков. Например, в них используются выражения «рост моих математических навыков», «учитель помогает мне понять, как решать задачу», «...шаг за шагом показывая, как решить задачу».

Кластер 5 содержит предложения, описывающие сожаления, связанные с прошлым неудачным опытом при изучении математики и упущенными возможностями, например «очень жаль, что в мои школьные годы не было такого отличного учителя», «...из-за незнания математики я провалил тест», «я чуть не бросил старшую школу из-за того, что не посещал занятия по математике, потому что я так сильно ее ненавижу».

4.3. Построение графа знаний на основе анализа негативных отзывов МООК по математике

Графы знаний позволяют наглядно представлять и структурировать отношения между сущностями и описывать их атрибуты. Вершины графа представляют сущности (документы, навыки, вакансии, музыкальные мелодии и т. п.), а ребра — отношения (расстояние Жаккара, события и т. п.).

На основе графа знаний и методов машинного обучения анализируют научные публикации [Chi et al., 2018], граф знаний можно использовать для наглядного представления профессиональных навыков и их сопоставления с вакансиями для упрощения анализа рынка труда [Groot de, Schutte, Graus, 2021].

Для исследования математической тревожности построим граф знаний с помощью *Python*-библиотек *SpaCy* и *NX* соответственно на основе определения частей речи и извлечения сущностей. Из предложений с негативной тональностью для построения графа знаний извлекались местоимение «я» и слова «математика», «школа» и т. п. и шаблон из слова *math* и следующего за ним или предшествующего ему существительного или местоимения в качестве субъектов и объектов, определяющих вершины графа. Отношение между субъектом и объектом приписывалось ребрам графа и определялось глаголом и предше-

Рис. 6. Граф знаний, построенный на основе выборки предложений с негативной тональностью



ствующим ему или следующим за ним наречием или прилагательным (при наличии).

Граф знаний, построенный на основе выборки предложений с негативной тональностью, приведен на рис. 6. Он позволяет визуализировать сущности и взаимосвязи между ними, а также упрощает интерпретацию результатов. В частности, он демонстрирует негативные эмоции, которые возникали при изучении математики, такие как фобия, страх, неприятности, связанные с изучением этого предмета, а также выявляет отношения, например «всегда вызвала», «является скучной» и т. п.

Наглядное представление сущностей, описывающих математическую тревожность, и отношений между ними может быть использовано учителями, тьюторами, разработчиками MOOK, а также психологами для анализа причин математической тревожности, поиска способов их устранения и оказания своевременной помощи обучающимся.

5. Выводы Математическая тревожность остается серьезной проблемой, препятствующей приобретению математических знаний. В статье предложена методология выявления математической тревожности на основе методов интеллектуального анализа данных. В частности, алгоритм анализа тональности VADER использовался для выявления предложений с негативной тональ-

ностью, описывающих отношение к математике, к предыдущему опыту ее изучения в школе; модель представления языка BERT применялась для векторизации предложений, содержащих математическую тревожность; метод локтя, метод *kMeans* и метод главных компонент использовались для определения оптимального числа кластеров, выделения кластеров семантически близких предложений, содержащих описание математической тревожности, и их визуализации; с помощью метода распознавания частей речи и графа знаний визуализировались отношения между обучающимся и его переживаниями, связанными с прошлым фрустрирующим опытом изучения математики. Полученные результаты могут быть использованы преподавателями для совершенствования содержания MOOK по математике и психологами для разработки рекомендаций по предотвращению и терапии тревожных состояний, связанных с изучением математики.

Таким образом, количество отзывов с подробным описанием отношения слушателей курсов к математике, текущему курсу и тьютору, ограничено, но результаты, полученные в этой статье, могут быть использованы для исследования математической тревожности и математической фобии при анализе датасетов больших объемов с отзывами на разных языках и применены как дополнительный инструмент учебной аналитики в анализе тревожных состояний при изучении математики.

Работа опубликована при поддержке Программы университетского партнерства НИУ ВШЭ.

Литература

1. Быстрова Т. Ю., Ларионова В. А., Синицын А. В., Толмачев А. В. (2018) Учебная аналитика MOOK как инструмент прогнозирования успешности обучающихся. № 4. С. 139–166. doi:10.17323/1814-9545-2018-4-139-166
2. Дюlicheva Ю. Ю. (2020) О применении технологии дополненной и виртуальной реальности в процессе обучения математике и физике // Открытое образование. Т. 24. No 3. С. 44–55. doi:10.21686/1818-4243-2020-3-44-55
3. Косова Е. А., Изетова М. Ю. (2020) Доступность массовых открытых онлайн-курсов по математике для обучающихся с ограниченными возможностями здоровья // Вопросы образования / Educational Studies Moscow. № 1. С. 205–229. doi:10.17323/1814-9545-2020-1-205-229
4. Adarsh R., Ashwin Patil, Shubham Rayar, Veena K. M. (2019) Comparison of VADER and LSTM for Sentiment Analysis // International Journal of Recent Technology and Engineering. Vol. 7. Iss. 6S. P. 540–543.
5. An Ya-H., Pan L., Kan M.-Ye., Dong Q., Fu Y. (2019) Resource Mention Extraction for MOOC Discussion Forums // IEEE Access. Vol. 7. P. 87887–87900. doi:10.1109/ACCESS.2019.2924250
6. Ashcraft M. H., Moore A. M. (2009) Mathematics Anxiety and the Affective Drop in Performance // Journal of Psychoeducational Assessment. Vol. 27. No 3. P. 197–205. doi:10.1177/0734282908330580

7. Bakharia A. (2017) PerspectivesX: A Proposed Tool for Scaffold Collaborative Learning Activities within MOOCs. arXiv:1704.04846
8. Brinton Ch.G., Buccapatnam S., Chiang M., Poor H.V. (2015) Mining MOOC Clickstreams: On the Relationship between Learner Behavior and Performance. arXiv:1503.06489
9. Capuano N., Caballe S., Conesa J., Greco A. (2020) Attention-Based Hierarchical Recurrent Neural Networks for MOOC Forum Posts Analysis // *Journal of Ambient Intelligence and Humanized Computing*. Vol. 12. No 5. P. 1–13. doi:10.1007/s12652-020-02747-9
10. Carey E., Hill F., Devine A., Szucs D. (2017) The Modified Abbreviated Math Anxiety Scale: A Valid and Reliable Instrument for Use with Children // *Frontiers in Psychology*. Vol. 8. Art. No 11. doi:10.3389/fpsyg.2017.00011
11. Chi Ya., Qin Y., Song R., Xu H. (2018) Knowledge Graph in Smart Education: A Case Study of Entrepreneurship Scientific Publication Management // *Sustainability*. Vol. 10. Art. No 995. doi:10.3390/su10040995
12. Dernoncourt F., Taylor C., O'Reilly U.-M., Veeramachaneni K., Wu Sh., Hallowa Sh. (2013) MocoViz: A Large Scale, Open Access, Collaborative, Data Analytics Platform for MOOCs // *Proceedings of the NIPS Workshop on Data-Driven Education (Lake Tahoe, Nevada, USA, December, 9–10, 2013)* doi:10.13140/2.1.3749.1201
13. Devlin J., Chang M.-W., Lee K., Toutanova K. (2019) BERT: Pre-Training of Deep Bidirectional Transformers for Language Understanding. arXiv:1810.04805v2
14. Field A. P., Evans D., Bloniewski T., Kovas Yu. (2021) Predicting Maths Anxiety from Mathematical Achievement across the Transition from Primary to Secondary Education // *Royal Society Open Science*. Vol. 6. Art. No 191459. doi:10.1098/rsos.191459
15. Griffiths L., Pratt D., Jennings D., Schmoller S. (2019) A MOOC for Adult Learners of Mathematics and Statistics: Tensions and Compromises in Design // G. Burrill, D. Ben-Zvi (eds) *Topics and Trends in Current Statistics Education Research*. Cham, Switzerland: Springer International. P. 351–371.
16. Groot de M., Schutte J., Graus D. (2021) Job Posting-Enriched Knowledge Graph for Skills-Based Matching. arXiv:2109.02554v1
17. Hou Yi., Zhou P., Wang T., Yu L., Hu Y., Wu D. (2016) Context-Aware Online Learning for Course Recommendation of MOOC Big Data. arXiv:1610.03147
18. Hutto C.J., Gilbert E. (2014) VADER: A Parsimonious Rule-based Model for Sentiment Analysis of Social Media Text // *Proceedings of the Eight International AAAI Conference on Weblogs and Social Media (Ann Arbor, Michigan, USA, June 1–4, 2014)*. P. 216–225.
19. Jeon B., Park N. (2020) Dropout Prediction over Weeks in MOOCs by Learning Representations of Clicks and Videos. arXiv:2002.01955
20. Lambert S. (2015) Reluctant Mathematician: Skill-Based MOOC Scaffolds Wide Range of Learners // *Journal of Interactive Media in Education*. No 1. Art. No 21. doi:10.5334/jime.bb
21. Li B., Zhou H., He J., Mingxuan M., Yang Y., Li L. (2020) On the Sentence Embeddings from Pre-trained Language Models. arXiv:2011.05864v1
22. Liu Z., Xiong F., Zou K., Wang H. (2018) Predicting Learning Status in MOOCs Using LSTM. arXiv:1808.01616v1
23. Lundqvist K. O., Liyanagunawardena Th., Starkey L. (2021) Evaluation of Student Feedback within a MOOC Using Sentiment Analysis and Target Groups // *International Review of Research in Open and Distributed Learning*. Vol. 21. No 3. P. 140–156. doi:10.19173/irrodl.v21i3.4783
24. Luo Y., Li J., Xie Z., Zhou G., Xiao X. (2018) MOOC Course Evaluation Based on Big Data Analysis. *Advances in Computer Science Research // Proceedings of the 2018 International Conference on Computer Science, Electronics and Communication Engineering (CSECE2018) (Wuhan, China, February 7–8, 2018)*. Vol. 80. P. 349–352. doi:10.2991/csece-18.2018.75

25. Ma X. (1999) A Meta-Analysis of the Relationship between Anxiety toward Mathematics and Achievement in Mathematics // *Journal for Research in Mathematics Education*. Vol. 30. No 5. P. 520–540. doi:10.2307/749772
26. Murarka A., Radhakrishnan B., Ravichandran S. (2020) Detection and Classification of Mental Illnesses on Social Media using RoBERTa. arXiv:2011.11226v1
27. Owen D., Camacho-Collados J., Espinosa-Anke L. (2020) Towards Preemptive Detection of Depression and Anxiety in Twitter // *Proceedings of the Social Media Mining for Health Applications (Barcelona, Spain, Online, December 12, 2020)*. arXiv:2011.05249
28. Primi C., Donati M. A., Izzo V. A. et al. (2020) The Early Elementary School Abbreviated Math Anxiety Scale (the EES-AMAS): A New Adapted Version of the AMAS to Measure Math Anxiety in Young Children // *Frontiers in Psychology*. Vol. 11. Art. No 1014. doi:10.3389/fpsyg.2020.01014
29. Saifullah S., Fauziah Yu., Aribowo A. S. (2021) Comparison of Machine Learning for Sentiment Analysis in Detecting Anxiety based on Social Media Data. <https://arxiv.org/ftp/arxiv/papers/2101/2101.06353.pdf>
30. Sinha T. (2014) Supporting MOOC Instruction with Social Network Analysis. arXiv:1401.5175
31. Soares F., Lopes A. P. (2016) Teaching Mathematics using Massive Open Online Courses // *Proceedings of 10th International Technology, Education and Development Conference (Valencia, Spain, March, 7–9, 2016)*. P. 2635–2641.
32. Stella M. (2021) Network Psychometrics and Cognitive Network Science Open New Ways for Detecting, Understanding and Tackling the Complexity of Math Anxiety: A Review. arXiv:2108.13800v1
33. Tang S., Peterson J. C., Pardos Z. A. (2016) Modeling Student Behavior using Granular Large Scale Action Data a MOOC. arXiv:1604.04789
34. Taranto E., Robutti O., Arzarello F. (2020) Learning within MOOCs for Mathematics Teacher Education // *ZDM: The International Journal on Mathematics Education*. Vol. 52. No 2. P. 1–15. doi:10.1007/s11858-020-01178-2
35. Wang Ch., Huang S., Zhou Ya. (2021) Sentiment Analysis of MOOC Reviews via ALBERT-BiLSTM Model // *MATEC Web of Conferences*, 336, 05008. doi:10.1051/mateconf/202133605008
36. Wen M., Yang D., Rose C. P. (2014) Sentiment Analysis in MOOC Discussion Forums: What Does It Tell Us? // *Proceedings of the 7th International Conference on Educational Data Mining, EDM 2014 (London, UK, July 4–7, 2014)*. P. 130–137.
37. Wolk A., Chlasta K., Holas P. (2021) Hybrid Approach to Detecting Symptoms of Depression in Social Media Entries // *Proceedings of the Twenty-Fifth Pacific Asia Conference on Information System (Dubai, UAE, June 20–24, 2021)*. <https://arxiv.org/ftp/arxiv/papers/2106/2106.10485.pdf>
38. Wong A. W., Wong K., Hindle A. (2019) Tracing Forum Posts to MOOC Content using Topic Analysis. arXiv:1904.07307
39. Wong J.-S., Zhang X. L. (2018) MessageLens: A Visual Analytics System to Support Multifaceted Exploration of MOOC Forum Discussions // *Visual Informatics*. Vol. 2. Iss. 1. P. 37–49. doi:10.1016/j.visinf.2018.04.005
40. Zhang F., Liu D., Liu C. (2020) MOOC Video Personalized Classification Based on Cluster Analysis and Process Mining // *Sustainability*. Vol. 12. Art. No 3066. doi:10.3390/su12073066

References

- Adarsh R., Ashwin Patil, Shubham Rayar, Veena K. M. (2019) Comparison of VADER and LSTM for Sentiment Analysis. *International Journal of Recent Technology and Engineering*, vol. 7, iss. 6S, pp. 540–543.
- An Ya-H., Pan L., Kan M.-Ye., Dong Q., Fu Y. (2019) Resource Mention Extraction for MOOC Discussion Forums. *IEEE Access*, vol. 7, pp. 87887–87900, doi:10.1109/ACCESS.2019.2924250

- Ashcraft M. H., Moore A. M. (2009) Mathematics Anxiety and the Affective Drop in Performance. *Journal of Psychoeducational Assessment*, vol. 27, no 3, pp. 197–205. doi:10.1177/0734282908330580
- Bakharia A. (2017) *PerspectivesX: A Proposed Tool for Scaffold Collaborative Learning Activities within MOOCs*. arXiv:1704.04846
- Brinton Ch.G., Buccapatnam S., Chiang M., Poor H.V. (2015) *Mining MOOC Clickstreams: On the Relationship Between Learner Behavior and Performance*. arXiv:1503.06489
- Bystrova T., Larionova V., Sinitsyn E., Tolmachev A. (2018) Uchebnaya analitika MOOK kak instrument prognozirovaniya uspehnosti obuchayushchikhsya [Learning Analytics in Massive Open Online Courses as a Tool for Predicting Learner Performance]. *Voprosy obrazovaniya / Educational Studies Moscow*, no 4, pp. 139–166. doi:10.17323/1814-9545-2018-4-139-166
- Capuano N., Caballe S., Conesa J., Greco A. (2020) Attention-Based Hierarchical Recurrent Neural Networks for MOOC Forum Posts Analysis. *Journal of Ambient Intelligence and Humanized Computing*, vol. 12, no 5, pp. 1–13. doi:10.1007/s12652-020-02747-9
- Carey E., Hill F., Devine A., Szucs D. (2017) The Modified Abbreviated Math Anxiety Scale: A Valid and Reliable Instrument for Use with Children. *Frontiers in Psychology*, vol. 8, art. no 11. doi:10.3389/fpsyg.2017.00011
- Chi Ya., Qin Y., Song R., Xu H. (2018) Knowledge Graph in Smart Education: A Case Study of Entrepreneurship Scientific Publication Management. *Sustainability*, vol. 10, art. no 995. doi:10.3390/su10040995
- Dernoncourt F., Taylor C., O'Reilly U.-M., Veeramachaneni K., Wu Sh., Halawa Sh. (2013) MoocViz: A Large Scale, Open Access, Collaborative, Data Analytics Platform for MOOCs. *Proceedings of the NIPS Workshop on Data-Driven Education (Lake Tahoe, Nevada, USA, December, 9–10, 2013)*. doi:10.13140/2.1.3749.1201
- Devlin J., Chang M.-W., Lee K., Toutanova K. (2019) *BERT: Pre-Training of Deep Bidirectional Transformers for Language Understanding*. arXiv:1810.04805v2
- Dyulichева Yu. Yu. (2020) O primeneniі tekhnologii dopolnennoy i virtual'noy real'nosti v protsesse obucheniya matematike i fizike [About the Usage of the Augmented Reality Technology in Mathematics and Physics Learning]. *Open Education*, vol. 24, no 3, pp. 44–55. doi:10.21686/1818-4243-2020-3-44-55
- Field A. P., Evans D., Bloniewski T., Kovas Yu. (2021) Predicting Maths Anxiety from Mathematical Achievement across the Transition from Primary to Secondary Education. *Royal Society Open Science*, vol. 6, art. no 191459. doi:10.1098/rsos.191459
- Griffiths L., Pratt D., Jennings D., Schmoller S. (2019) A MOOC for Adult Learners of Mathematics and Statistics: Tensions and Compromises in Design. *Topics and Trends in Current Statistics Education Research* (eds G. Burrill, D. Ben-Zvi), Cham, Switzerland: Springer International, pp. 351–371.
- Groot de M., Schutte J., Graus D. (2021) *Job Posting-Enriched Knowledge Graph for Skills-Based Matching*. arXiv:2109.02554v1
- Hou Yi., Zhou P., Wang T., Yu L., Hu Y., Wu D. (2016) *Context-Aware Online Learning for Course Recommendation of MOOC Big Data*. arXiv:1610.03147
- Hutto C.J., Gilbert E. (2014) VADER: A Parsimonious Rule-based Model for Sentiment Analysis of Social Media Text. *Proceedings of the Eight International AAAI Conference on Weblogs and Social Media (Ann Arbor, Michigan, USA, June 1–4, 2014)*, pp. 216–225.
- Jeon B., Park N. (2020) *Dropout Prediction over Weeks in MOOCs by Learning Representations of Clicks and Videos*. arXiv:2002.01955
- Kosova Y., Izetova M. (2020) Dostupnost' massovykh otkrytykh onlayn-kursov po matematike dlya obuchayushchikhsya s ogranichennymi vozmozhnostyami zdorov'ya [Accessibility of Massive Open Online Courses on Mathematics for Students with Disabilities]. *Voprosy obrazovaniya / Educational Studies Moscow*, no 1, pp. 205–229. doi:10.17323/1814-9545-2020-1-205-229

- Lambert S. (2015) Reluctant Mathematician: Skill-Based MOOC Scaffolds Wide Range of Learners. *Journal of Interactive Media in Education*, no 1, art. no 21. doi:10.5334/jime.bb
- Li B., Zhou H., He J., Mingxuan M., Yang Y., Li L. (2020) *On the Sentence Embeddings from Pre-trained Language Models*. arXiv:2011.05864v1
- Liu Z., Xiong F., Zou K., Wang H. (2018) *Predicting Learning Status in MOOCs Using LSTM*. arXiv:1808.01616v1
- Lundqvist K. O., Liyanagunawardena Th., Starkey L. (2021) Evaluation of Student Feedback within a MOOC Using Sentiment Analysis and Target Groups. *International Review of Research in Open and Distributed Learning*, vol. 21, no 3, pp. 140–156. doi:10.19173/irrodl.v21i3.4783
- Luo Y., Li J., Xie Z., Zhou G., Xiao X. (2018) MOOC Course Evaluation Based on Big Data Analysis. *Advances in Computer Science Research. Proceedings of the 2018 International Conference on Computer Science, Electronics and Communication Engineering (CSECE2018) (Wuhan, China, February 7–8, 2018)*, vol. 80, pp. 349–352. doi:10.2991/csece-18.2018.75
- Ma X. (1999) A Meta-Analysis of the Relationship between Anxiety toward Mathematics and Achievement in Mathematics. *Journal for Research in Mathematics Education*, vol. 30, no 5, pp. 520–540. doi:10.2307/749772
- Murarka A., Radhakrishnan B., Ravichandran S. (2020) *Detection and Classification of Mental Illnesses on Social Media using RoBERTa*. arXiv:2011.11226v1
- Owen D., Camacho-Collados J., Espinosa-Anke L. (2020) Towards Preemptive Detection of Depression and Anxiety in Twitter. *Proceedings of the Social Media Mining for Health Applications (Barcelona, Spain, Online, December 12, 2020)*. arXiv:2011.05249
- Primi C., Donati M. A., Izzo V. A. et al. (2020) The Early Elementary School Abbreviated Math Anxiety Scale (the EES-AMAS): A New Adapted Version of the AMAS to Measure Math Anxiety in Young Children. *Frontiers in Psychology*, vol. 11, art. no 1014. doi:10.3389/fpsyg.2020.01014
- Saifullah S., Fauziah Yu., Aribowo A. S. (2021) *Comparison of Machine Learning for Sentiment Analysis in Detecting Anxiety based on Social Media Data*. Available at: <https://arxiv.org/ftp/arxiv/papers/2101/2101.06353.pdf> (accessed 20 October 2021).
- Sinha T. (2014) *Supporting MOOC Instruction with Social Network Analysis*. arXiv:1401.5175
- Soares F., Lopes A. P. (2016) Teaching Mathematics using Massive Open Online Courses. *Proceedings of 10th International Technology, Education and Development Conference (Valencia, Spain, March, 7–9, 2016)*, pp. 2635–2641.
- Stella M. (2021) *Network Psychometrics and Cognitive Network Science Open New Ways for Detecting, Understanding and Tackling the Complexity of Math Anxiety: A Review*. arXiv:2108.13800v1
- Tang S., Peterson J. C., Pardos Z. A. (2016) *Modeling Student Behavior using Granular Large Scale Action Data a MOOC*. arXiv:1604.04789
- Taranto E., Robutti O., Arzarello F. (2020) Learning within MOOCs for Mathematics Teacher Education. *ZDM: The International Journal on Mathematics Education*, vol. 52, no 2, pp. 1–15. doi:10.1007/s11858-020-01178-2
- Wang Ch., Huang S., Zhou Ya. (2021) Sentiment Analysis of MOOC Reviews via ALBERT-BiLSTM Model. *MATEC Web of Conferences*, 336, 05008. doi:10.1051/mateconf/202133605008
- Wen M., Yang D., Rose C. P. (2014) Sentiment Analysis in MOOC Discussion Forums: What does it tell us? *Proceedings of the 7th International Conference on Educational Data Mining, EDM 2014 (London, UK, July 4–7, 2014)*, pp. 130–137.
- Wolk A., Chlasta K., Holas P. (2021) Hybrid Approach to Detecting Symptoms of Depression in Social Media Entries. *Proceedings of the Twenty-Fifth Pacific Asia Conference on Information System (Dubai, UAE, June 20–24, 2021)*. Available at:

<https://arxiv.org/ftp/arxiv/papers/2106/2106.10485.pdf> (accessed 20 October 2021).

Wong A. W., Wong K., Hindle A. (2019) *Tracing Forum Posts to MOOC Content using Topic Analysis*. arXiv:1904.07307

Wong J.-S., Zhang X. L. (2018) MessageLens: A Visual Analytics System to Support Multifaceted Exploration of MOOC Forum Discussions. *Visual Informatics*, vol. 2, iss. 1, pp. 37–49. doi:10.1016/j.visinf.2018.04.005

Zhang F., Liu D., Liu C. (2020) MOOC Video Personalized Classification Based on Cluster Analysis and Process Mining. *Sustainability*, vol. 12, art. no 3066. doi:10.3390/su12073066