

АВТОМАТИЗИРОВАННОЕ ОЦЕНИВАНИЕ КОРОТКИХ ОТВЕТОВ ОБУЧАЮЩИХСЯ С ИСПОЛЬЗОВАНИЕМ ЯЗЫКОВЫХ МОДЕЛЕЙ

Ч. Б. Миннегалиева¹ [0000-0003-4648-1623], И. И. Кашапов² [0009-0004-6341-8380],

О. Д. Морозова³ [0009-0009-5670-561X]

¹⁻³Институт вычислительной математики и информационных технологий,
Казанский федеральный университет, Казань

¹mchulpan@gmail.com, ²k.i.i.08@yandex.ru, ³olka30.olka30@mail.ru

Аннотация

Методы проверки ответов обучающихся с использованием языковых моделей в настоящее время исследуются разными специалистами. Результаты автоматизированного оценивания зависят от предметной области и особенностей учебной дисциплины. В работе проанализированы ответы студентов, полученные в ходе изучения курса «Компьютерная графика и дизайн». При помощи языковых моделей определены векторы документов. Предложен метод оценивания ответов через нахождение косинусного сходства полученных векторов и уточнение оценок проверкой ключевых слов. Результаты могут использоваться при предварительной проверке ответов студентов и являются базой для дальнейших исследований.

Ключевые слова: языковая модель, контроль знаний, обработка текста, ключевое слово ответа, автоматизированная оценка ответов обучающихся, косинусное сходство, векторное представление документа, BERT, word2vec, открытый вопрос

ВВЕДЕНИЕ

Проверка и оценка знаний, умений, навыков студентов являются обязательной составляющей обучения, процессом обратной связи [1]. От результатов и условий проведения контроля знаний зависят качество образовательного процесса и степень удовлетворенности обучающихся. Составление заданий для проверочных работ, тестирований и экзаменов всегда являлось достаточно трудоём-

кой работой для учителей и преподавателей [2]. Вопросы объективного оценивания заданий, выполненных обучающимися, и процедура проведения контрольных мероприятий также продолжают обсуждаться участниками образовательного процесса. С развитием онлайн-образования контроль знаний остаётся одной из наиболее острых и дискуссионных проблем организации обучения [3], так как в данном случае проверка знаний, умений и навыков в большинстве случаев проводится дистанционно. Увеличивается доля самостоятельной работы обучающихся, и, как правило, возрастает количество студентов, обучающихся у одного преподавателя. В настоящее время специалистами изучаются различные подходы к автоматизации составления заданий и оцениванию выполненных работ. Разработан инструмент анализа проектной работы, позволяющий преподавателю выставлять оценки, анализируя собранные данные о работе студентов [4]. Программный комплекс (система контроля знаний), использующий конструктивно-выборочный метод, позволяет на практике применять новый тип заданий для автоматизированного контроля знаний и навыков по программированию [5]. Предлагаются новые подходы, основанные на формализации знаний. Например, система, базой которой является онтология *OntoMath^{Edu}*, предназначена для решения ряда задач, в том числе для автоматического тестирования знаний [6].

Во время тестирования знаний применяются различные формы заданий. На данном этапе развития технологий платформы онлайн-обучения предоставляют возможность автоматической проверки ответов на такие вопросы, которые требуют выбора одного или нескольких верных из множества имеющихся, сопоставления, ввода числового ответа, заполнения пропусков и другие. При выполнении заданий открытого типа обучающиеся самостоятельно формулируют ответ. Включение в тесты таких заданий повышает объективность контроля знаний, так как в этом случае не предъявляются варианты верных и неверных ответов и снижается вероятность угадывания или подбора верного варианта.

В настоящей работе рассмотрены подходы к автоматизированному оцениванию коротких ответов. Такие ответы формулируются на естественном языке, их длина может изменяться от одного предложения до одного абзаца.

Специалисты начали исследовать проблему автоматизированного оценива-

ния коротких ответов более 50 лет назад [7]. Сегодня изучается возможность получения такой оценки при помощи языковых моделей, которые в последние годы используются для решения разных задач, в том числе в образовании [8, 9]. Рассмотрены подходы к автоматизации решения текстовых математических задач с использованием моделей [10]. Предложена методика оценки и анализа формулировок результатов обучения. В качестве исходных данных выступают формулировки результатов обучения, между которыми вычисляется семантическое расстояние. Трансформация исходного текста в вектор происходит с помощью языковой модели [11]. Представлено исследование задачи автоматической классификации коротких связных текстов (эссе) на английском языке [12]. Авторы этой работы подробно рассмотрели оценочные метрики для задач автоматического оценивания коротких ответов, учитывая специфику наборов данных. Так как база ответов обучающихся, как правило, имеет меньший объём, чем, например, база отзывов или данных с веб-страниц, интерпретация результатов должна идти с учётом этого факта [13]. По этой причине исследователи уделяют большое внимание подготовке набора данных для изучения проблемы автоматизированного оценивания коротких ответов, полученных на разных языках [14]. В работе [15] авторы оценили при помощи предварительно обученных моделей 840 ответов на 21 вопрос. Улучшить точность оценки можно при расширении базы ответов. Также значение имеет сбалансированность набора, например, сопоставимое количество верных и неверных ответов. В этой работе получены результаты применения семи моделей, но отмечено, что они требуют перепроверки с другими наборами данных. После такой проверки планируется перейти к следующей цели — расширить проект, включив в него работу над ответами в виде эссе [16].

Нами были собраны 1319 коротких ответов студентов на 41 вопрос, полученных в ходе изучения дисциплины по выбору «Компьютерная графика и дизайн». Ответам была дана экспертная оценка. Большинство ответов было оценено экспертом как верные (1) или неверные (0). В некоторых случаях, например, если ответ предполагал перечисление, экспертная оценка могла выражаться дробным числом от 0 до 1. В настоящей работе представлены результаты оценивания ответов при помощи языковых моделей и уточнения полученных оценок с использованием проверки ключевых слов.

ОЦЕНИВАНИЕ ОТВЕТОВ С ИСПОЛЬЗОВАНИЕМ ЯЗЫКОВЫХ МОДЕЛЕЙ

Языковые модели word2vec и BERT применяются для решения различных задач, связанных с обработкой текста. Например, в [17] с использованием этих моделей проведен sentiment-анализ данных социальных сетей. В [18] определены наиболее эффективные модели семейства BERT для выявления деструктивного контента в социальных медиа. В настоящее время исследователи изучают не только процесс обучения и полученные результаты, но и объяснимость и прозрачность алгоритмов моделей, что в перспективе должно привести к их более оптимальному использованию (см., например, [19]).

Для оценивания ответов ранее мы применили модель word2vec, обученную на текстах по компьютерной графике. Результаты показали, что при помощи этой модели можно выявить ответы, явно неверные и близкие к верным по формулировке. В случаях, когда значение косинусной меры было близко к 0.5, предлагался дополнительный этап проверки [20].

Ответы студентов, полученные во время опросов, зачетов и экзаменов, дополнительно были проверены при помощи готовой языковой модели на основе BERT [21]. Если значение косинусной меры для векторов корректного ответа и ответа студента не превосходило 0.5, мы считали, что модель верно оценила ответ студента. Например, вариантом корректного (шаблонного) ответа на вопрос «Что называется воксельной моделью?» является ответ «Трехмерная модель, состоящая из вокселей (элементов объема, кубиков)». Значения косинусной меры между векторами некоторых ответов студентов и шаблонного ответа приведены в таблице 1. Ответы здесь даны в том виде, в котором получены от студентов, орфография и пунктуация сохранены.

Таблица 1. Результаты оценивания ответов

Ответы студентов	Экспертная оценка	Косинусное сходство
«Воксельная модель – состоящая из вокселей (кубов)»	1	0.716
«адаптация цветовой модели»	0	0.237
«Воксельная модель – трехмерный растр, состоит	1	0.452

из вокселей (как пиксели в 2D, только в 3D). Из них строится изображение»		
---	--	--

Как видно из таблицы 1, в первом и втором случаях модель оценила ответы верно, в третьем случае правильный ответ студента был оценен как неверный. Всего в 76.5 % случаев модель на основе трансформеров верно оценила ответы студентов. Но точность по отдельным вопросам варьируется от 46 до 97 процентов.

Была осуществлена проверка, как повторение теории влияет на результаты теста. Для этого студентам непосредственно перед тестированием была предоставлена возможность прочитать необходимый теоретический материал, содержащий ответы на предлагаемые вопросы. Количество правильных ответов ожидаемо увеличилось. Также точность результата проверки ответов, которые получены после прочтения, выше на 11.1%, что связано в первую очередь с тем, что, формулируя ответы, студенты использовали чаще те слова и предложения, которые только что повторили. Данные по 10 вопросам представлены на рис. 1 (точность оценки в %).

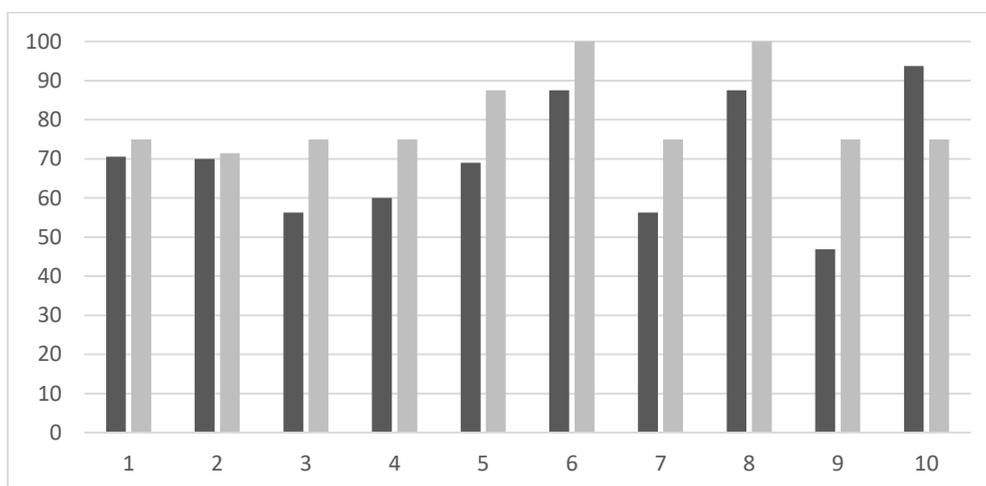


Рис. 1. Доля верно оцененных ответов (левые столбцы – ответы получены без повторения учебного материала, правые столбцы – ответы получены после повторения учебного материала)

В тех случаях, когда опрос проходит непосредственно после изучения темы, использование в формулировках корректного (шаблонного) ответа выражений из теоретического материала (лекций, рекомендованной литературы) позволит более точно оценить ответы. Если во время тестирования проверяются остаточные знания, при достаточно заполненной базе ответов возможно оценивание на основе сравнения верных ответов, полученных от студентов. Например, обучающиеся получили задание «Графический редактор – это ... (дописать определение, примеры приводить не нужно)». Как правило, это задание не вызывает затруднений, с графическим редактором обучающиеся знакомы со школы. Верные ответы студентов были сформулированы по-разному, например, «Приложение для редактирования изображений», «инструмент (приложение), которое позволяет редактировать и создавать изображения», «Приложение, с помощью которого можно создавать и изменять изображения» и т. д. Было найдено косинусное сходство между ответами студентов и корректным (шаблонным) ответом. После этого были выбраны 10 верных ответов, полученных от обучающихся. Затем сформирован вектор, являющийся средним для векторов верных ответов. Далее было вычислено косинусное сходство между средним вектором и векторами 10 случайных ответов студентов. В таблице 2 приведены оценки, выставленные экспертом, значения косинусной меры между векторами и абсолютные значения разности между оценками, выставленными экспертом и значениями косинусной меры.

Таблица 2. Результаты оценивания ответов различными способами

Номера ответов студентов	Экспертная оценка	Значение косинусной меры		Абсолютное значение разности между экспертной оценкой и значением косинусной меры	
		1-й способ	2-й способ	1-й способ	2-й способ
1	0.5	0.396	0.534	0.104	0.034
2	1	0.397	0.581	0.603	0.419
3	0	0.176	0.346	0.176	0.346
4	1	0.394	0.569	0.606	0.431
5	0.5	0.342	0.535	0.158	0.035

6	0	0.372	0.452	0.372	0.452
7	1	0.627	0.749	0.373	0.251
8	1	0.574	0.763	0.426	0.237
9	1	0.339	0.614	0.661	0.386
10	0	0.612	0.755	0.612	0.755

При 1-м способе рассматривались вектор ответа студента и вектор корректного (шаблонного) ответа. При 2-м способе – вектор ответа студента и вектор, являющийся средним для векторов верных ответов студентов. В большинстве случаев результаты оценивания стали лучше. Абсолютное значение разности между экспертной оценкой и значениями косинусной меры между векторами увеличилось для ответов, которые содержат слова верных ответов, но сформулированы ошибочно. Это ответ 3: «инструмент для работы с мультимедиями», ответ 6: «Приложение /облако, где можно редактировать или вносить какие-либо правки», ответ 10: «Графический редактор – приложение для редактирования» (орфография и пунктуация сохранены). Уточнить оценки ответов возможно проверкой ключевых слов.

ПРОВЕРКА КЛЮЧЕВЫХ СЛОВ

Под ключевыми словами корректного (шаблонного) ответа на вопрос будем понимать значимые, важные слова ответа. Специалистами проверка ключевых слов предлагается как одна из возможностей совершенствования процесса оценивания коротких ответов [22].

Для проверки ключевых слов сначала была проведена предобработка ответов, включая токенизацию, удаление стоп-слов и знаков препинания, а также лемматизацию. Чтобы избежать дублирования исходной информации в ответах студентов, были удалены слова, которые содержались в формулировке самого вопроса. Такие слова не несут дополнительного смысла и могут исказить результаты анализа. Например, если ответ на вопрос «Как получается цвет в модели RGB?» начинается с повторения слов вопроса: «Цвет в модели RGB получается ...», данные слова были удалены. Этап предобработки помог стандартизировать и очистить текстовые данные, делая их готовыми для дальнейшего анализа и обеспечивая более точные и непротиворечивые результаты.

На следующем шаге были определены ключевые слова, для каждого из них назначен вес, который отражает значимость ключевого слова в ответе. Совпадения ключевых слов между корректным ответом и ответом студента оцениваются с учетом их весов. На приведенный выше вопрос про модель RGB было получено 64 ответа. Вариант корректного ответа: «смешиванием основных компонент цветовой модели – Red, Green, Blue (красный, зеленый, синий)». В 51 случае (79.7 %) оценки, выставленные экспертом и определенные через косинусное сходство с использованием языковой модели, совпали. В 7 случаях из оставшихся неверно оцененных 13 ответов проверка ключевых слов помогла уточнить оценки. Ключевыми были определены слова: «красный» (или «Red»), «зеленый» («Green»), «синий» («Blue»). В итоге в 6 случаях из 64 оценка, определенная через косинусное сходство, не совпала с экспертной. Некоторые из неверно оцененных ответов приведены в таблице 3 (орфография и пунктуация сохранены).

Таблица 3. Результаты оценивания ответов

Ответы студентов	Экспертная оценка	Косинусное сходство	Ключевые слова в ответе
«Цвет получается путем смешивания 3 основных цветов в одном пикселе»	1	0.267	отсутствуют
«Путем смешивания 3 цветов RGB в разных пропорциях»	1	0.241	приведены сокращенно
«Путем слияния в разных количествах параметров R G и B, варьирующих в значении от 0 до 255, где все по 0 – черный цвет, все по 255 – белый»	1	0.123	приведены сокращенно
«RGB – Red Green Blue?»	0	0.535	присутствуют

Вопрос про модель RGB не вызывал затруднений у студентов. В первых трех случаях из приведенных в таблице 3 эксперт не снизил оценку за то, что цвета не

были названы полностью. Отметим, что экспертные оценки не снижались за грамматические, орфографические, речевые, стилистические и другие ошибки, если они не искажали смысл предложения, так как проверялись знания по компьютерной графике. В четвертом ответе обучающийся не написал про смешение (сложение) компонентов, поэтому эксперт оценил ответ как неверный. Косинусное сходство близко к 0.5, ключевые слова присутствуют. После вычисления косинусного сходства и проверки ключевых слов в 90.7 % случаев ответы студентов на данный вопрос были оценены верно.

Таким образом, векторизация ответов при помощи языковых моделей предоставляет возможность предварительного оценивания ответов обучающихся. В работе были проанализированы ответы, которые определяются, в основном, как верные или неверные. Разработка критериев позволит получить более объективные оценки эксперта и повысит точность результатов применения моделей, но увеличит время работы преподавателей. Полученные результаты могут помочь при анализе развернутых ответов на вопросы, требующих пояснения нескольких понятий.

ЗАКЛЮЧЕНИЕ

Представлены результаты анализа ответов студентов, сформулированных в свободной форме. Векторы предложений получены при помощи языковой модели на основе BERT, далее определялось косинусное сходство ответа обучаемого и корректного (шаблонного) ответа. В 76.5 % всех случаев языковая модель позволила оценить верность ответа обучающегося. Точность полученной оценки зависит от формулировки корректного ответа, времени и места тестирования в графике учебного процесса. При достаточном заполнении базы ответов возможно вместо шаблонного ответа взять вектор, являющийся средним вектором для векторов верных ответов обучающихся. Описаны подходы к уточнению оценки при помощи проверки ключевых слов. Результаты могут использоваться при предварительной проверке ответов студентов, в составе цифровых образовательных ресурсов при выполнении тестов на самопроверку с последующим обсуждением оценок.

Благодарности

Работа выполнена за счет средств Программы стратегического академического лидерства Казанского (Приволжского) федерального университета («ПРИОРИТЕТ-2030»)

СПИСОК ЛИТЕРАТУРЫ

1. Шепелюк О.Л. Методы контроля знаний студентов в условиях реализации образовательных стандартов // Глобальный научный потенциал. 2021. № 8 (125). С. 112–114.
2. Бухман Л.М. Проблемы тестового контроля знаний и их решение // Известия Самарского научного центра Российской академии наук. 2010. Т. 12, № 5. С. 21–24.
3. Тихонова Н.В. Организация контроля знаний студентов в условиях удаленного обучения // Казанский лингвистический журнал. 2021. Т. 4, № 1. С. 111–125.
4. Атнагулов А.А., Абрамский М.М. О подходе к автоматизации оценки знаний в области разработки программного обеспечения на основе анализа данных проектной работы // Электронные библиотеки. 2023. Т. 26, № 5. С. 589–599.
5. Жуков И.А. Система контроля знаний и практических навыков по программированию // Информатика и образование. 2023. Т. 38, № 2. С. 66–74.
6. Nevzorova O.A., Falileeva M.V., Kirillovich A.V., Lipachev E.K., Shakirova L.R., Dyupina A.E. OntoMathEdu Educational Ontology: Problems of Ontological Engineering // Pattern Recognition and Image Analysis. 2023. V. 33 (3). P. 460–466.
7. Burrows S., Gurevych I. and Stein B. The Eras and Trends of Automatic Short Answer Grading // International Journal of Artificial Intelligence in Education. 2015. V. 25. P. 60–117.
8. Yan L., Sha L., Zhao L., Li Y., Martinez-Maldonado R., Chen G., Li X., Jin Y., Gašević D. Practical and ethical challenges of large language models in education: A systematic scoping review // British Journal of Educational Technology. 2023. V. 55, Iss. 1. P. 90–112.
9. Zirar A. Exploring the impact of language models, such as ChatGPT, on student learning and assessment // Review of Education. 2023. V. 11, e3433.
10. Sundaram S.S., Gurajada S., Padmanabhan D., Sam Abraham S., Fisichella M.

Does a language model “understand” high school math? A survey of deep learning based word problem solvers // WIREs Data Mining and Knowledge Discovery. 2024. 1534.

11. *Гиниятуллин В.М., Ермолаев Е.В., Салихова М.А., Хлыбов А.В., Чурилов Д.А., Чурилова Е.А.* Исследование структуры и содержания компетенций с помощью языковой модели ELMO // *Современные наукоемкие технологии.* 2021. № 8. С. 58–65.

12. *Лагутина Н.С., Лагутина К.В., Бредерман А.М., Касаткина Н.Н.* Классификация текстов по уровням CEFR с использованием методов машинного обучения и языковой модели BERT // *Моделирование и анализ информационных систем.* 2023. Т. 30, № 3. С. 202–213.

13. *Ahmed A., Joorabchi A., Hayes M.* On Deep Learning Approaches to Automated Assessment: Strategies for Short Answer Grading // *In Proceedings of the 14th International Conference on Computer Supported Education (CSEDU 2022).* 2022. V. 2. P. 85–94.

14. *Agarwal D., Gupta S., Baghel N.* ScAA: A Dataset for Automated Short Answer Grading of Children’s Free-text Answers in Hindi and Marathi // *Proceedings of the 17th International Conference on Natural Language Processing, Patna, India, December 18 – 21, 2020.* P. 430–436.

15. *Wilianto D., Girsang A.S.* Automatic Short Answer Grading on High School’s E-Learning Using Semantic Similarity Methods // *TEM Journal.* 2023. V. 12, Is. 1. P. 297–302.

16. *Divya A., Haridas V. and Narayanan J.* Automation of Short Answer Grading Techniques: Comparative Study using Deep Learning Techniques // *2023 Fifth International Conference on Electrical, Computer and Communication Technologies (ICECCT), Erode, India.* 2023. P. 1–7.

17. *Ярушкина Н.Г., Мошкин В.С., Константинов А.А.* Применение языковых моделей word2vec и bert в задаче сентимент-анализа текстовых сообщений социальных сетей // *Автоматизация процессов управления.* 2020. № 3 (61). С. 60–69.

18. *Минаев В.А., Симонов А.В.* Сравнение моделей-трансформеров BERT при выявлении деструктивного контента в социальных медиа // *Информация и безопасность.* 2022. Т. 25, № 3. С. 341–348.

19. *Surov I.* Opening the Black Box: Finding Osgood's Semantic Factors in Word2vec Space // Informatics and Automation. 2022. V. 21, No. 5. P. 916–936.

20. *Миннегалиева Ч.Б., Сабитова Г.А., Гаялиев А.М.* Метод предварительной оценки ответов обучающихся на основе векторной модели документов // Электронные библиотеки. 2023. Т. 26, № 3. С. 324–339.

21. *Reimers N., Gurevych I.* Sentence-BERT: Sentence Embeddings using Siamese BERT-Networks // In Proceedings of the 2019 Conference on Empirical Methods in Natural Language Processing and the 9th International Joint Conference on Natural Language Processing (EMNLP-IJCNLP) Hong Kong, China Association for Computational Linguistics. 2019. P. 3982–3992.

22. *Кожевников В.А., Сабинин О.Ю.* Система автоматической проверки ответов на открытые вопросы на русском языке // Научно-технические возможности СПбГПУ. 2018. Т. 11, № 3. С. 57–72.

AUTOMATED STUDENTS' SHORT ANSWERS GRADING USING LANGUAGE MODELS

Chulpan Minnegalieva¹ [0000-0003-4648-1623], **Ilnur Kashapov**² [0009-0004-6341-8380],
Olga Morozova³ [0009-0009-5670-561X]

^{1–3}*Institute of Computer Mathematics and Information Technologies, Kazan Federal University*

¹mchulpan@gmail.com, ²k.i.i.08@yandex.ru, ³olka30.olka30@mail.ru

Abstract

Methods for assessing student answers using language models are currently being studied by various specialists. The results of automated assessment depend on the subject area and characteristics of the academic discipline. This paper analyzes the students' answers received during the course «Computer Graphics and Design». It is proposed to determine the cosine similarity of document vectors obtained using language models and refine the estimates by checking keywords. The results obtained can be used for preliminary assessment of students' answers and are the basis for further research.

Keywords: *language model, knowledge control, natural text processing, student answer keyword, automated short answer grading, cosine similarity, document vector representation, BERT, word2vec, open-ended question*

REFERENCES

1. *Shepelyuk O.L.* Metody kontrolya znanij studentov v usloviyah realizacii obrazovatel'nyh standartov // *Global'nyj nauchnyj potencial*. 2021. № 8 (125). S. 112–114.
2. *Buhman L.M.* Problemy testovogo kontrolya znanij i ih reshenie // *Izvestiya Samarskogo nauchnogo centra Rossijskoj akademii nauk*. 2010. T. 12, № 5. S. 21–24.
3. *Tihonova N.V.* Organizaciya kontrolya znanij studentov v usloviyah udalennogo obucheniya // *Kazanskij lingvisticheskiy zhurnal*. 2021. T. 4, № 1. S. 111–125.
4. *Atnagulov A.A., Abramskij M.M.* Student teams project work analysis tool development // *Russian Digital Libraries Journal*. 2023. T. 26, No 5. P. 589–599.
5. *Zhukov I.A.* Sistema kontrolya znanij i prakticheskikh navykov po programmirovaniyu // *Informatika i obrazovanie*. 2023. T. 38, № 2. S. 66–74.
6. *Nevzorova O.A., Falileeva M.V., Kirillovich A.V., Lipachev E.K., Shakirova L.R., Dyupina A.E.* OntoMathEdu Educational Ontology: Problems of Ontological Engineering // *Pattern Recognition and Image Analysis*. 2023. V. 33 (3). P. 460–466.
7. *Burrows S., Gurevych I. and Stein B.* The Eras and Trends of Automatic Short Answer Grading // *International Journal of Artificial Intelligence in Education*. 2015. V. 25. P. 60–117.
8. *Yan L., Sha L., Zhao L., Li Y., Martinez-Maldonado R., Chen G., Li X., Jin Y., Gašević D.* Practical and ethical challenges of large language models in education: A systematic scoping review // *British Journal of Educational Technology*. 2023. V. 55, Is. 1. P. 90–112.
9. *Zirar A.* Exploring the impact of language models, such as ChatGPT, on student learning and assessment // *Review of Education*. 2023. V. 11, e3433.
10. *Sundaram S.S., Gurajada S., Padmanabhan D., Sam Abraham S., Fisichella M.* Does a language model “understand” high school math? A survey of deep learning based word problem solvers // *WIREs Data Mining and Knowledge Discovery*. 2024. 1534.

11. *Giniyatullin V.M., Ermolaev E.V., Salihova M.A. Hlybov A.V., Churilov D.A., Churilova E.A.* Issledovanie struktury i sodержaniya kompetencij s pomoshch'yu yazykovoju modeli ELMO // *Sovremennye naukoemkie tekhnologii*. 2021. № 8. S. 58–65.

12. *Lagutina N.S., Lagutina K.V., Brederman A.M., Kasatkina N.N.* Klassifikaciya tekstov po urovnjam CEFR s ispol'zovaniem metodov mashinnogo obucheniya i yazykovoju modeli BERT // *Modelirovanie i analiz informacionnyh sistem*. 2023. T. 30, № 3. S. 202–213.

13. *Ahmed A., Joorabchi A., Hayes M.* On Deep Learning Approaches to Automated Assessment: Strategies for Short Answer Grading // *In Proceedings of the 14th International Conference on Computer Supported Education (CSEDU 2022)*. 2022. V. 2. P. 85–94.

14. *Agarwal D., Gupta S., Baghel N.* ScAA: A Dataset for Automated Short Answer Grading of Children's Free-text Answers in Hindi and Marathi // *Proceedings of the 17th International Conference on Natural Language Processing, Patna, India, December 18 – 21, 2020*. P.430–436.

15. *Wilianto D., Girsang A.S.* Automatic Short Answer Grading on High School's E-Learning Using Semantic Similarity Methods // *TEM Journal*. 2023. V. 12, Is. 1. P. 297–302.

16. *Divya A., Haridas V. and Narayanan J.* Automation of Short Answer Grading Techniques: Comparative Study using Deep Learning Techniques // *2023 Fifth International Conference on Electrical, Computer and Communication Technologies (ICECCT), Erode, India*. 2023. P. 1–7.

17. *Yarushkina N.G., Moshkin V.S., Konstantinov A.A.* Primenenie yazykovykh modelej word2vec i bert v zadache sentiment-analiza tekstovykh soobshchenij social'nyh setej // *Avtomatizaciya processov upravleniya*. 2020. № 3 (61). S. 60–69.

18. *Minaev V.A., Simonov A.V.* Sravnenie modelej-transformerov BERT pri vyavlenii destruktivnogo kontenta v social'nyh media // *Informaciya i bezopasnost'*. 2022. T. 25, № 3. S. 341–348.

19. *Surov I.* Opening the Black Box: Finding Osgood's Semantic Factors in Word2vec Space // *Informatics and Automation*. 2022. V. 21, No. 5. P. 916–936.

20. *Minnegalieva Ch.B., Sabitova G.A., Gayaliev A.M.* Method of pre-assessment

of students' answers based on the vector model of documents// Russian Digital Libraries Journal. 2023. T. 26, No 3. P. 324–339.

21. *Reimers N., Gurevych I.* Sentence-BERT: Sentence Embeddings using Siamese BERT-Networks // In Proceedings of the 2019 Conference on Empirical Methods in Natural Language Processing and the 9th International Joint Conference on Natural Language Processing (EMNLP-IJCNLP) Hong Kong, China Association for Computational Linguistics. 2019. P. 3982–3992.

22. *Kozhevnikov V.A., Sabinin O.Yu.* Sistema avtomaticheskoy proverki otvetov na otkrytye voprosy na russkom yazyke // Nauchno-tehnicheskie vozmozhnosti SPbGPU. 2018. T. 11. № 3. S. 57–72.

СВЕДЕНИЯ ОБ АВТОРАХ



МИННЕГАЛИЕВА Чулпан Бакиевна – к. п. н., доцент кафедры информационных систем Института вычислительной математики и информационных технологий, Казанский федеральный университет, г. Казань.

Chulpan MINNEGALIEVA – Associate Professor, Institute of Computer Mathematics and Information Technologies, Kazan Federal University.

email: mchulpan@gmail.com

ORCID: 0000-0003-4648-1623.



КАШАПОВ Ильнур Илхамович – магистр 1 г.о., Институт вычислительной математики и информационных технологий, Казанский федеральный университет, г. Казань.

Ilnur KASHAPOV – master's student, Institute of Computer Mathematics and Information Technologies, Kazan Federal University.

email: k.i.i.08@yandex.ru

ORCID: 0009-0004-6341-8380



МОРОЗОВА Ольга Дмитриевна – магистр 1 г.о., Институт вычислительной математики и информационных технологий, Казанский федеральный университет, г. Казань.

Olga MOROZOVA – master's student, Institute of Computer Mathematics and Information Technologies, Kazan Federal University.

email: olka30.olka30@mail.ru

ORCID: 0009-0009-5670-561X

Материал поступил в редакцию 2 мая 2024 года