

Известия Алтайского государственного университета. 2025. № 4 (144). С. 61–66.
Izvestiya of Altai State University. 2025. No 4 (144). P. 61–66.

Научная статья

УДК 519.6:51-77:004.89 519.677

DOI: 10.14258/izvasu(2025)4-08

Программный комплекс для оценивания голосовых ответов обучающихся

Александр Александрович Дмитриев¹, Денис Александрович Дмитриев²

¹Алтайский государственный университет, Барнаул, Россия, dmitriev@asu.ru

²Алтайский государственный университет, Барнаул, Россия, dmitriev_d_a@vc.asu.ru

Original article

Software System for Assessing Students' Voice Responses

Alexander A. Dmitriev¹, Denis A. Dmitriev²

¹Altai State University, Barnaul, Russia, dmitriev@asu.ru

²Altai State University, Barnaul, Russia, dmitriev_d_a@vc.asu.ru

Аннотация. В работе рассмотрен подход к созданию программных средств для оценивания голосовых ответов обучающихся, полученных с помощью технологий распознавания речи. В связи с тем что интеллектуальные технологии на основе систем распознавания речи сегодня широко используются при дистанционном обучении, необходимо расширение применяемых инструментов для оценивания знаний в результате учебной деятельности. В качестве такого инструмента в работе предложен программный комплекс, позволяющий обрабатывать голосовой ответ обучающегося и с помощью алгоритмов анализа текста производить оценку распознанного ответа. Для распознавания голосового ответа в текстовое сообщение использовалась свободно распространяемая система распознавания речи Vosk. На этапе обработки текстового сообщения для получения оценки, отражающей правильность ответа обучающегося, применены алгоритмы символьной и статистической обработки естественного языка. При тестировании программного комплекса методами вычислительных экспериментов установлено соответствие между вычисляемыми баллами и шкалой оценивания, используемой преподавателем при приеме устного ответа. Результаты работы показывают, что программный комплекс может быть использован в электронном обучении для автоматического оценивания ответов при дистанционном приеме экзаменов с помощью голосового помощника.

Ключевые слова: дистанционное образование, система распознавания речи, вычислительные эксперименты

Abstract. The paper discusses an approach to creating software tools for assessing students' voice responses obtained using speech recognition technologies. Intelligent technologies based on speech recognition systems are widely used today in distance learning. Therefore, it is necessary to enhance and improve the tools used for assessing knowledge as a result of educational activities. The paper proposes a software system that allows one to process a student's voice response and evaluate the recognized response using text analysis algorithms. The freely distributed Vosk speech recognition system is used to convert the student's voice response into a text message. Symbolic and statistical natural language processing algorithms help process the text message and obtain a score that reflects the correctness of the student's answer. The developed software system has been tested with computational experiment methods. Tests prove that the calculated score points fully correspond to the assessment scale used by a teacher during oral exams. The work results demonstrate good prospects for the developed software system to assess exam answers automatically using a voice assistant in various e-learning environments.

Keywords: distance education, speech recognition system, computing experiment

Для цитирования: Дмитриев А.А., Дмитриев Д.А. Программный комплекс для оценивания голосовых ответов обучающихся // Известия Алтайского государственного университета. 2025. № 4 (144). С. 61–66. DOI: 10.14258/izvasu(2025)4-08.

Финансирование. Работа выполнена при реализации проекта «Речевые технологии искусственного интеллекта: цифровой двойник преподавателя» в рамках Программы «Приоритет-2030» стратегического развития ФГБОУ ВО «Алтайский государственный университет» на 2021–2030 гг.

Введение

Для расширения возможностей электронного обучения в последнее время набирает популярность использование интеллектуальных систем на основе распознавания речи и обработки текста на естественном языке [1]. Вместе подобные технологии при взаимодействии с обучающимся выступают в качестве «виртуального» преподавателя и обеспечивают выдачу учебного материала и контроль за выполнением заданий. Распространенным примером является задействование коммерческих и свободно распространяемых систем голосовых помощников в составе электронных курсов [2]. Достоинством такого подхода является то, что для изучения учебных материалов между системой электронного обучения и обучающимся поддерживается интерактивный диалог, в котором оба участника обмениваются голосовыми или текстовыми сообщениями, содержащими задания и ответы на них. Такой подход способствует углубленному изучению контента дистанционных курсов. Однако требует внедрения программных систем для автоматической проверки и оценки устных и текстовых ответов студентов. При этом решение такой задачи на практике сопряжено с рядом трудностей. Во-первых, связанных с работоспособным объединением технологий распознавания речи и обработки текста на естественном языке в составе единого программного комплекса. Во-вторых, с самой разработкой алгоритмов для оценивания ответа студента, учитывающих полноту, правильность и языковое оформление ответа, полученного средствами голосового помощника.

Существующие методы для автоматической проверки и оценки ответов основаны на применении алгоритмов для обработки текста и математического аппарата анализа данных. При простом сравнении отдельных знаков и слов ответа с заданным шаблоном используются критерии сходства на основе вычисления общих элементов. Для подсчета числа ключевых слов в тексте ответа применяется алгоритм «Мешок слов» [3]. В более сложных задачах, связанных с анализом структуры и грамматических особенностей текста ответа студента, исполь-

For citation: Dmitriev A.A., Dmitriev D.A. Software System for Assessing Students' Voice Responses. *Izvestiya of Altai State University*. 2025. No 4 (144). P. 61–66. (In Russ.). DOI: 10.14258/izvasu(2025)4-08.

Financing. The work was carried out during the implementation of the project "Speech technologies of artificial intelligence: a digital twin of a teacher" within the framework of the Priority-2030 Program of strategic development of the Altai State University for 2021–2030.

зуют вычисление информативных признаков. В качестве таких информативных параметров в работе [4] используются статистические характеристики текста. Для установления смысловой связи между ответом учащегося и текстовым шаблоном правильного ответа, составленным преподавателем, привлекаются методы семантического анализа тестов и нейронных сетей [5].

Целью настоящей работы является создание и апробация программного комплекса, основанного на цифровых технологиях распознавания речи и интеллектуальных алгоритмов обработки текста и позволяющего оценивать голосовые ответы учащихся.

Актуальность работы определяется важностью решения задачи автоматической проверки и оценки устных ответов учащихся, нерешенной в настоящее время в связи с ее сложностью и разнообразием конкретных ситуаций.

Описание программного комплекса

В работе использован программный комплекс, структурная схема которого представлена на рисунке 1.

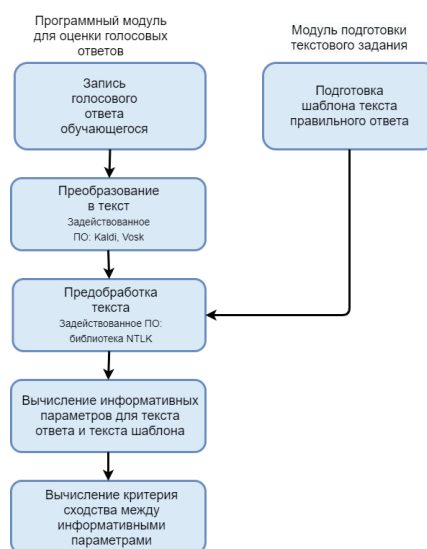


Рис. 1. Используемая обработка текста в программном комплексе

Представленный комплекс состоит из трех программных модулей: программного обеспечения для подготовки контрольных заданий, системы распознавания речи и программного модуля обработки текста на естественном языке для оценивания устного ответа обучающегося.

Подготовка контрольного задания осуществлялась путем выборки тестового задания из заранее составленной базы данных и его отображения на дисплее монитора через встроенный браузер операционной системы. База данных тестовых заданий состояла из набора текстовых вопросов, цифровых изображений к ним и текстового описания правильного ответа на контрольный вопрос. Текстовое описание правильного ответа заранее было подготовлено преподавателем и являлось шаблоном, с которым в дальнейшем сравнивался голосовой ответ обучающегося.

Устный ответ обучающегося на задание записывался в виде голосового сигнала и сохранялся в файл с расширением wav. С целью улучшения качества полученный голосовой сигнал обрабатывался алгоритмами частотной фильтрации для удаления шумовых составляющих, лежащих вне частотного диапазона звуков речи [6]. Для преобразования голосового сигнала в текст в работе использованы свободно распространяемое программное обеспечение для распознавания речи Kaldi и библиотека Vosk [7]. После распознавания речи записанного ответа обучающегося формировалось текстовое сообщение, которое в дальнейшем обрабатывалось программным модулем для оценки голосовых ответов.

Программное обеспечение для модуля оценки голосовых ответов было разработано с использованием библиотеки анализа и обработки текстов Natural Language Toolkit [8]. Алгоритм обработки текста состоял из следующих действий:

- удаление стоп-слов, небуквенных символов;
- приведение слов к начальной форме;
- выполнение алгоритма построения триграмм на основе текста;
- векторизация текста.

При обработке текста удалялись отдельные стоп-слова, не несущие смысловой информации для данного текста, и небуквенные символы, включающие знаки препинания и специальные знаки. В полученном тексте выполнялась лемматизация слов для приведения к их начальной форме [8]. Для последующего применения алгоритма векторизации обработанный текст разбивался на отдельные триграммы, которые состояли из тройки последовательно выбранных из текста слов.

Векторизация текста проводилась с помощью вычисления статистической меры TF-IDF [9]. На этом этапе для каждого слова из текста рассчитывались числовой параметр TF, соответствующий частоте употребления этого слова, и числовое значение обратной частоты IDF, с которой слово употребляется в составленных триграммах. Вычисление параметров TF и IDF позволило сопоставить тексту набор числовых коэффициентов, учитывающих особенности строения текста и частоты использования отдельных слов. Произведение этих двух параметров формировало вектор признаков, который в дальнейшем использовался для сравнения текстов.

Представленный алгоритм обработки текста и вычисления вектора признаков применялся для текста, записанного в шаблон правильного ответа, и для текстового сообщения, полученного при ответе обучающегося на поставленный в тестовом задании вопрос. Каждому тексту был сопоставлен свой рассчитанный вектор признаков. Полученные векторы признаков имели одинаковую длину. Для сравнения двух векторов в работе использовалась мера косинусного сходства [10].

Описанная обработка текстовых сообщений была реализована нами в виде программы на языке программирования Python. Для проверки корректности вычислений при работе программного обеспечения были обработаны голосовые записи студентов, которые отвечали на один тестовый вопрос преподавателя, показанный в таблице 1. Каждая голосовая запись была распознана в текст. После этого каждый текстовый ответ был сравнен с правильным ответом преподавателя. Для этого, согласно алгоритму, для текстов была проведена предобработка и представление слов в начальной форме в виде триграмм, вычислены свой вектор признаков, состоящий из числовых параметров TF-IDF. Длина каждого вектора признаков равна количеству слов из правильного ответа преподавателя. Данная обработка текста показана в таблице 1 для двух ответов учащихся.

Для каждого из ответов учащихся была вычислена мера косинусного подобия между информативными признаками ответа преподавателя и ответа студента. Для слов из ответа учащегося 1 и ответа преподавателя вычисленное значение косинусного сходства равно 0,87, а при сравнении ответа учащегося 2 с текстом правильного ответа — 0,61 соответственно. Уменьшение значения подобия двух текстов во втором случае связано с включением в текст ответа новых слов, не отражающих смысл правильного ответа.

Результаты промежуточной обработки текста для последующего оценивания

Вопрос: Дайте описание понятия цветовой модели RGB в компьютерной графике		
Ответ преподавателя	Цветовая модель представляется в виде декартовой системы координат, где каждый оттенок цвета кодируется тремя первичными цветами: красным, зеленым, синим	
Ответ учащегося, распознанный из голоса и обработанный алгоритмом	Учащийся 1: «Цветовая модель эр джи би является аппаратной и цветовой моделью используемой для цветных мониторов обычно данная цветовая система представляется в декартовой системе координат где каждый оттенок цвета кодируется тремя первичными цветами а именно красным и зеленым синим точками цвета»	Учащийся 2: «Цветовая модель р жи би является аппаратно ориентированной моделью которая используется в цветовых мониторах она задается в системе декарта в их координат где эр джи би это первичные цвета то есть красный и синий и зеленый для получения вторичных цветов осуществляется за счет смешения первичных например красный и зеленый дают в сумме желтый»
Пример лемматизации начала текста	Цветовой, модель, эр, джи, би, являться, аппаратный, цветовой, модель...	Цветовой, модель, р, жи, би, являться, аппаратный, цветовой, модель, использовать, цветной, монитор, обычно, данный, цветовой, система...

Применение комплекса для тестирования обучающихся

Разработанный программный комплекс был апробирован нами при приеме устных голосовых ответов обучающихся на дисциплинах «Обработка цифровых изображений» и «Сетевое администрирование». Для подготовки вопроса и правильного ответа преподавателя, записи голосовых ответов студентов использовался веб-интерфейс программного комплекса, представленный на рисунке 2.

Рис. 2. Графический интерфейс пользователя для приема голосовых ответов

По перечисленным дисциплинам на этапе внешнего контроля выполнения самостоятельной работы 16 студентам было предложено пройти дистанционное тестирование с применением представленного программного комплекса. Созданный тест содержал ряд вопросов, построенных по классической схеме «Вопрос с выбором ответа» и изображение с вопросом для последующего устного ответа. Шаблон правильного ответа для устного задания предварительно подготавливался преподавателем, состоял из машинописного текста на естественном языке и загружался в программный комплекс. Получение студентом такого тестового задания для дисциплины «Сетевое администрирование» выглядело следующим образом:

1. На экране браузера отображались в виде изображения формула из материалов курса или изображение объекта, связанного с курсом. Например, цифровое изображение сетевого маршрутизатора.

2. В текстовом виде выводилось задание для работы с изображением: «Дайте название и опишите область применения показанного на рисунке устройства в локальной сети».

3. Ответ на задание принимался в виде записанного голосового ответа студента, ограниченного по времени.

Голосовой ответ студента был прослушан преподавателем для его оценивания по 4-балльной шкале согласно используемым в рамках данного предмета критериям из фонда оценочных средств. На этапе проверки голосового ответа обучающегося программным комплексом была выставлена оценка от 0 до 100. Сопоставление результатов оценивания ответа студентов преподавателем и программным комплексом представлено в таблице 2.

Таблица 2

Сопоставление оценки ответа обучающегося с 4-балльной шкалой оценивания

4-балльная шкала оценивания ответа преподавателем	Оценка ответа программным комплексом	Число ответов с коллизиями, в % от общего числа ответов
Отлично	80–100	12.5
Хорошо	69–79	
Удовлетворительно	50–68	
Неудовлетворительно	0–49	18.8

Из таблицы 2 следует, что высокий балл за ответ студента, выставленный программным комплексом, соответствовал оценкам «отлично» и «хорошо», полученным от преподавателя. Распознанный текст такого ответа был структурирован, слова и предложения текста были семантически связаны между собой. В удовлетворительных ответах с диапазоном баллов [50–68] отдельные предложения были слабо связаны друг с другом. В тексте наблюдались многочисленные вставки паразитных слов, но в целом текст ответа по смыслу соответствовал правильному ответу на вопрос задания. При большем несопадении текста ответа шаблону, заданному преподавателем, оценка ответа программным комплексом уменьшалась до нулевых значений. Данные ответы характеризовались несвязным текстом, вставкой отдельных слов, отражающих тематику задания. Стоит отметить, что при полном несоответствии ответа студента текстовому шаблону отметка достигала близкого к нулевым значения.

При апробации программного комплекса был определен ряд факторов, влияющих на оценивание ответа студента. Проведенный анализ голосовых записей ответов студентов и результатов преобразования речи в текст показал, что уменьшение вычисленной оценки происходило суммарно в 31,3 % голосовых ответов обучающихся и зависело от качества распознавания речи. Снижение балла ответа наблюдалось при произношении отвечающим звуков латинских символов, которые он использовал при ответе. Распознавание таких звуков приводило к созданию отдельных букв или слогов на русском языке, которые в ряде случаев отличались от подготовленного преподавателем текста ответа. Следующей причиной для снижения оценки являлось одновременное про-

изношение ответа несколькими находящимися рядом студентами, что приводило к распознаванию программным комплексом совместной речи, приводящей к вставкам несвязных слов в текст ответа обучающегося. Другим обстоятельством, влияющим на оценку, являлся развернутый ответ отвечающего на поставленный вопрос, не совпадающий с ответом преподавателя. Отметим, что при должной организации проведения голосового экзамена, связанной с дистанцированной рассадкой учащихся и подготовкой продуманного ответа преподавателя, перечисленные коллизии могут быть сведены к минимуму.

Заключение

Проведено исследование действия разработанного программного комплекса для дистанционного приема устных экзаменов и оценки голосовых ответов учащихся. Представленный комплекс при оценивании голосовых ответов использует технологии распознавания речи и программное обеспечение для анализа и обработки текстов. Для выявления факторов, определяющих правильность выставяемой оценки, проведено тестирование программного комплекса методами вычислительных экспериментов. Показано влияние внешних условий при записи голосового сигнала и ошибок при распознавании слов на оценивание ответа обучающегося. Установлено соответствие между результатами оценивания голосового ответа обучающегося преподавателем и представленным программным комплексом.

Полученные результаты показали, что предложенный комплекс может использоваться в составе электронных курсов для автоматического оценивания голосовых ответов студентов при дистанционном обучении.

Библиографический список

1. Захаров А.А., Захарова И.Г., Шабалин А.М., Ханбеков Ш.И., Джалилзода Д.Б. Интеллектуальный голосовой помощник как пример реализации методологии инклюзивного дизайна // Образование и наука. 2024. Т. 26. № 3. С. 149–175.
2. Трегубов В.Н. Использование голосовых ассистентов для развития английской научной речи // International Journal of Open Information Technologies. 2020. № 6. С. 62–72.
3. Нугуманова А.Б., Бессмертный И.А., Пецина П., Байбуринов Е.М. Обогащение модели Bag-of-words семантическими связями для повышения качества классификации

текстов предметной области // Программные продукты и системы. 2016. № 2 (114). С. 89–97.

4. Талор М., Мане П. Голосовая система оценки ответов для учащихся с ограниченными физическими возможностями, использующих обработку естественного языка и машинное обучение // Научно-технический вестник информационных технологий, механики и оптики. 2023. Т. 23. № 2. С. 299–302.

5. Lakshmi V., Ramesh V. Evaluating Students' Descriptive Answers Using // Natural Language Processing and Artificial Neural Networks (IJCRT). 2017. Vol. 5. Iss. 4. P. 3168–3173.

6. Дмитриев А.А., Дмитриев Д.А. Применение цифровой обработки голосовых сигналов для улучшения распознавания речи // Проблемы правовой и технической защиты информации. 2023. № 10. С. 4–8.

7. Беленко М.В., Балакшин П.В. Сравнительный анализ систем распознавания речи с открытым кодом // Меж-

дународный научно-исследовательский журнал. 2017. № 4 (58). С. 13–18.

8. Patil S. M., Patil S. Evaluating Student Descriptive Answers Using Natural Language Processing // International Journal of Engineering Research & Technology (IJERT). 2014. Vol. 3. Iss. 3. P. 1716–1718.

9. Кравченко Ю.А., Мансур А.М., Мохаммад Ж.Х. Векторизация текста с использованием методов интеллектуального анализа данных // Известия ЮФУ. Технические науки. 2021. № 2 (219). С. 154–167.

10. Оськина К.А. Оптимизация метода классификации текстов, основанного на tf-idf, за счет введения дополнительных коэффициентов // Вестник Московского государственного лингвистического университета. Гуманитарные науки. 2016. № 15 (754). С. 175–187.

References

1. Zaharov A.A., Zaharova I.G., SHabalin A.M., Hanbekov SH.I., Dzhahalilzoda D.B. Intelligent Voice Assistant as an Example of Implementation of Inclusive Design Methodology. *The Education and Science Journal*. 2024. Vol. 26. No 3. P. 149–175. (In Russ.).

2. Tregubov V.N. Using Voice Assistants to Develop English Scientific Speech. *International Journal of Open Information Technologies*. 2020. No 6. P. 62–72. (In Russ.).

3. Nugumanova A.B., Bessmertnyj I.A., Pecina P., Bajburin E.M. Semantic Relations in Text Classification Based on Bag-of-words Model. *Software & Systems*. 2016. No 2 (114). P. 89–97. (In Russ.).

4. Talor M., Mane P. Voice Based Answer Evaluation System for Physically Disabled Students Using Natural Language Processing and Machine Learning. *Scientific and Technical Journal of Information Technologies, Mechanics and Optics*. 2023. Vol. 23. No 2. P. 299–302. (In Russ.).

5. Lakshmi V., Ramesh V. Evaluating Students' Descriptive Answers Using. *Natural Language Processing and Artificial Neural Networks (IJCRT)*. 2017. Vol. 5. No 4. P. 3168–3173.

6. Dmitriev A.A., Dmitriev D.A. Using of Digital Voice Signal Processing to Improve Speech Recognition. *Problems of Legal and Technical Protection of Information*. 2023. No 10. P. 4–8. (In Russ.).

7. Belenko M.V., Balakshin P.V. Comparative Analysis of Open Source Speech Recognition Systems. *International Research Journal*. 2017. No 4 (58). P. 13–18. (In Russ.).

8. Patil S. M., Patil S. Evaluating Student Descriptive Answers Using Natural Language Processing. *International Journal of Engineering Research & Technology (IJERT)*. 2014. Vol. 3. No 3. P. 1716–1718.

9. Kravchenko YU.A., Mansur A.M., Mohammad ZH.H. Text Vectorization Using Data Mining Methods. *Izvestiya of South Federal University. Engineering Sciences*. 2021. No 2 (219). P. 154–167. (In Russ.).

10. Oskina K.A. Optimization of the Tf-idf Based Text Classification Method by Introducing Additional Coefficients. *Vestnik of Moscow State Linguistic University*. 2016. No 15 (754). P. 175–187. (In Russ.).

Информация об авторах

А.А. Дмитриев, кандидат физико-математических наук, доцент кафедры информационной безопасности, Алтайский государственный университет, Барнаул, Россия;

Д.А. Дмитриев, аспирант Института цифровых технологий, электроники и физики, Алтайский государственный университет, Барнаул, Россия.

Information about the authors

A.A. Dmitriev, Candidate of Sciences in Physics and Mathematics, Associate Professor of the Department of Information Security, Altai State University, Barnaul, Russia;

D.A. Dmitriev, Postgraduate Student of the Institute of Digital Technology, Electronics and Physics, Altai State University, Barnaul, Russia.