

Министерство образования Республики Беларусь
Учреждение образования
Белорусский государственный университет
информатики и радиоэлектроники

УДК 004.855.5

Криводубский
Андрей Юрьевич

Модели и алгоритмы системы обучения и оценки знаний

АВТОРЕФЕРАТ

на соискание академической степени
магистра

по специальности 1-40 80 05 – Программная инженерия

Научный руководитель
Серебряная Л.В.
к.т.н., доцент

Минск 2022

КРАТКОЕ ВВЕДЕНИЕ

Процесс обучения в любом образовательном учреждении подразумевает проверку усвоения учащимися полученных знаний. Такая проверка важна для обратной связи и последующей корректировки процесса обучения.

Одним из эффективных видов проверки знаний студентов является тестирование, играющее очень важную роль при оценивании результатов обучения.

Наиболее эффективным методом тестирования является тестирование с помощью открытых вопросов – где отвечающий дает развернутый или короткий ответ на естественном языке. При таком виде тестирования практически полностью исключается возможность угадывания правильного ответа, а от студента требуется умение не просто вспомнить термин или определение, но и корректно сформулировать свои мысли.

Преимущества систем автоматического оценивания ответов на открытые вопросы по сравнению с классической проверкой преподавателем:

- уменьшение времени проверки – система позволяет сразу оценить результаты тестирования;

- увеличение объективности оценивания – автоматическое оценивание позволяет снизить ошибки из-за усталости или невнимательности и убрать возможные претензии студентов из-за предвзятого отношения.

Исходя из этого, было принято решение исследовать различные модели, используемые для задачи оценивания коротких ответов, и выполнить тонкую настройку наиболее подходящей из них с целью повышения качественных характеристик для использования в системе автоматического тестирования студентов.

ОБЩАЯ ХАРАКТЕРИСТИКА РАБОТЫ

Цель и задачи исследования

Целью диссертационной работы является разработка модели и алгоритма программного обеспечения для системы обучения и оценки знаний.

Для достижения поставленной цели необходимо решить следующие задачи:

1. Рассмотреть результаты последних исследований по теме автоматического оценивания ответов, выявить тенденции и актуальные вопросы;

2. Рассмотреть известные модели и алгоритмы автоматического оценивания ответов, выявить наиболее перспективные;

3. Выполнить тонкую настройку существующих нейросетевых моделей глубокого обучения для определения семантического сходства текстов, экспериментальным путем выявить оптимальную модель для применения в системе обучения и оценки знаний;

4. Осуществить проектирование системы автоматического оценивания знаний с использованием разработанной нейросетевой модели глубокого обучения.

Объектом исследования является текстовая информация учебных материалов и ответов обучающихся на естественном языке.

Предметом исследования являются модели и алгоритмы автоматической проверки знаний, представленных в виде ответов на естественном языке.

Основной *гипотезой*, положенной в основу диссертационной работы, является возможность применения современных моделей и алгоритмов обработки естественных языков на основе машинного обучения для автоматизации процесса обучения и оценки знаний. Внедрение систем с использованием данных моделей и алгоритмов в учебные заведения позволит повысить качество оценивания успеваемости учащихся, а также оптимизировать загрузку преподавательского состава.

Новизна полученных результатов

В результате проведенного исследования была разработана нейросетевая модель для решения задачи автоматического оценивания ответов на естественном языке, был разработан алгоритм автоматического оценивания ответов с использованием разработанной модели.

Положения, выносимые на защиту

1. Нейросетевая модель нахождения степени сходства ответа студента и ответа-эталона.
2. Алгоритм оценивания ответа студента по десятибалльной шкале.
3. Архитектура программного средства для реализации автоматического оценивания коротких ответов.

Апробация результатов диссертации

Основные положения диссертационной работы докладывались на II Международной научно-практической конференции «Актуальные вопросы развития современных технологий» (Петрозаводск, Россия, 2022); VIII Международной научно-практической конференции «Applied science of today: problems and new approaches» (Петрозаводск, Россия, 2022).

Опубликованность результатов диссертации

По теме диссертации опубликовано 2 работы в сборниках трудов и материалов международных конференций.

Структура и объем диссертации

Диссертация состоит из введения, общей характеристики работы, четырех глав, заключения, списка использованных источников, списка публикаций автора. В первой главе представлен анализ последних исследований в предметной области, выявлены основные тенденции и проблемы в рамках тематики исследования, показаны направления их решения. Вторая глава посвящена рассмотрению моделей и алгоритмов автоматического оценивания ответов и выбору моделей для дальнейшей тонкой настройки и исследования. В третьей главе проведен эксперимент по настройке различных вариаций нейросетевой модели глубокого обучения BERT, на основе наилучшей вариации создана модель для автоматического оценивания ответов, предложен алгоритм для оценивания ответов с использованием разработанной модели. В четвертой главе предложена архитектура программного средства для реализации автоматического оценивания коротких ответов с использованием разработанной модели глубокого обучения.

Общий объем работы составляет 67 страниц, из которых основного текста – 56 страниц, включая 35 рисунков и 6 таблиц, список использованных источников из 57 наименований на 5 страницах.

КРАТКОЕ СОДЕРЖАНИЕ РАБОТЫ

В первой главе были рассмотрены существующие системы обучения и оценки знаний, а также последние исследования и наработки в области решения задачи автоматического оценивания ответов. Выявлены актуальные направления исследований, существующие проблемы.

В рамках второй главы были более детально рассмотрены модели и алгоритмы, применяемые для задачи автоматического оценивания ответов. Начиная от простых моделей на основе мер сходства, и заканчивая современными моделями на основе нейронных сетей. Рассмотрены особенности строения наиболее популярной нейронной сети BERT. Был произведен выбор моделей глубокого обучения, которые могут быть настроены для решения задачи автоматического оценивания ответов.

В третьей главе были настроены выбранные модели глубокого обучения путем их дообучения на популярных наборах данных Mohler и SciEntsBank. Полученные результаты были сравнены между собой с помощью общепринятых для этой задачи метрик. На основании результатов сравнения была выбрана модель с наилучшей производительностью для применения в системе автоматической оценки знаний. С помощью процедуры дистилляции знаний поведение обученной нейросетевой модели было скопировано на мультязычную модель с целью добавления поддержки русского языка.

В четвертой главе была осуществлена проверка работоспособности разработанной модели с помощью приложения-прототипа, предложена потенциальная архитектура системы автоматического обучения и оценки знаний, которая может использовать обученную в ходе экспериментов модель для автоматического оценивания ответов.

ЗАКЛЮЧЕНИЕ

Были изучены и сравнены последние работы в области автоматического оценивания ответов. На основании данного исследования был сделан вывод, что существует два основных метода оценивания ответов: меры сходства и методы ML. Меры, основанные на сходстве, не требуют большого обучающего набора, однако эти методы не дают хороших результатов для открытых ответов. С другой стороны, методы ML расширяют возможности систем оценивания, давая хорошие результаты для полуоткрытых вопросов. Основным недостатком использования методов ML является то, что для каждого вопроса требуется большое размеченное обучающее множество, что может быть громоздким и непрактичным. Одной из последних тенденций является использование предварительно обученных нейросетевых моделей глубокого обучения и их тонкая настройка для применения в решении задачи ASAG. В частности, популярность получила модель BERT, в свою очередь основанная на модели трансформера.

Далее были рассмотрены различные модели и алгоритмы, использующиеся для автоматического оценивания ответов. Наиболее прогрессивными являются нейросетевые модели глубокого обучения, основанные на трансформерах. В связи с этим, именно ряд моделей на основе архитектуры BERT был выбран в качестве предмета исследования в дальнейшем.

В ходе тонкой настройки моделей глубокого обучения была выявлена и доработана модель с наилучшими характеристиками метрик корреляции и RMSE – *cross-encoder/stsb-roberta-large*, которая может быть как применена в системе автоматического оценивания ответов, так и дорабатываться и сравниваться с другими моделями в дальнейшем, благодаря использованию популярного набора данных для обучения.

Так как потенциальной конечной целью является практическое применение разработанной модели для проверки ответов студентов, с помощью процесса дистилляции знаний была обучена многоязычная модель на основе *sentence-transformers/stsb-xlm-r-multilingual*, которая сможет оценивать ответы на уровне сопоставимом с *cross-encoder/stsb-roberta-large*, однако при этом поддерживать входные данные на русском языке, без необходимости предварительного перевода сторонними сервисами.

Также был предложен алгоритм автоматического оценивания ответов с использованием полученной нейросетевой модели, который может быть положен в основу программного средства для решения задачи автоматического оценивания кратких ответов.

Разработанная модель была проверена на работоспособность в приложении-прототипе. Затем была предложена архитектура системы автоматической оценки знаний, которая позволяет системе функционировать при активной нагрузке, при этом минимизируя нагрузку на базу данных.

Дальнейшие исследования можно развивать в сторону использовать мультязычных моделей глубокого обучения, также можно будет изучить альтернативные модели глубокого обучения на основе трансформеров, например, XLNet, либо их комбинации, которые начинают составлять сильную конкуренцию сетям BERT.

СПИСОК ОПУБЛИКОВАННЫХ РАБОТ

1-А. Криводубский, А. Ю. Современные тенденции в автоматизированной оценке ответов на открытые вопросы / А.Ю. Криводубский // Актуальные вопросы развития современных технологий. – Петрозаводск, 2022. – № 2. – с. 101–107.

2-А. Криводубский, А. Ю. Доработка моделей машинного обучения для автоматического тестирования студентов / А.Ю. Криводубский // Applied science of today: problems and new approaches. – Петрозаводск, 2022. – № 8. – с. 144–149.