DOI: 10.58168/MIST2026_502-510
УДК 004

ДАЛЬНЕЙШЕЕ РАЗВИТИЕ ОНЛАЙН-ПЛАТФОРМЫ ДЛЯ ПОДГОТОВКИ К ВСЕРОССИЙСКОЙ ПРОВЕРОЧНОЙ РАБОТЕ (ВПР)

Т.В. Скворцова[1], П.А. Пискурский[1], О.А. Авилова[1]

[1]ФГБОУ ВО «Воронежский государственный лесотехнический университет имени Г.Ф. Морозова»

Аннотация. В статье проводится анализ по улучшению и оптимизации онлайн-платформы для поддержания её востребованности на рынке. Были изучены современные семантические эмбеддинги (ИИ-модели), направленные на решение основных недостатков алгоритмов TF-IDF и косинусного сходства. Также была рассмотрена перспективная возможность, направленная на использование метода трансферного обучения (дообучения) ИИмоделей для снижения вычислительных затрат, достижения оптимальной производительности и увеличения точности обработки текстов.

Ключевые слова: ИИ, нейронные сети, машинное обучение, латентно-семантический анализ, WEB-разработка.

FURTHER DEVELOPMENT OF THE ONLINE PLATFORM FOR PREPARATION FOR THE ALL-RUSSIAN TEST ASSIGNMENT (VPR)

T.V. Skvortsova[1], P.A. Piskursky[1], O.A. Avilova[1]

[1]Voronezh State University of Forestry and Technologies named after G.F. Morozov

Abstract. The article analyzes how to improve and optimize the online platform to maintain its market demand. Modern semantic embeddings (AI models) aimed at solving the main disadvantages of TF-IDF and cosine similarity algorithms were studied. A promising opportunity aimed at using the transfer learning method of AI models to reduce computational costs, achieve optimal performance and increase the accuracy of word processing was also considered.

Keywords: AI, neural networks, machine learning, latent semantic analysis, WEB development.

Введение

В 2016 году в систему образования была введена Всероссийская проверочная работа (ВПР) для отслеживания состояния системы образования путём проверки знаний обучающихся в образовательных организация от Федерального института оценки качества образования (далее – ФИОКО) [1]. Конечные цели ВПР – дать не только независимую оценку просвещения знаний, но и выявлять пробелы в знаниях обучающихся для проведения последующей методической работы.

В первое время, итоги ВПР показывали значительно хорошие результаты, однако позже выяснилось, что учителя и педагоги помогали обучающимся выполнять задания, из-за этого снижалась объективность результатов для независимой оценки [2]. Следовательно, для повышения объективности были введены методической работы по комплексному повторению материалов, привлечены независимые наблюдатели по контролю проведения ВПР и тренировочные варианты заданий для подготовки.

Согласно результатам информационного обзора оценки качества образования от ФИОКО за период 2021-2024 гг., количество «двоек» сократилось, а количество высоких оценок наоборот увеличились [3]. Несмотря на положительную динамику, обучающиеся столкнулись психологическим давлением, выпадением из привычного процесса, а на учителей увеличилась нагрузка, что приводит образовательной организации к дополнительным расходам и времени на подготовку ВПР [4].

Для решения данных проблем подойдет современная автоматизированная онлайн-платформа подготовки к ВПР с применением искусственного интеллекта (ИИ), машинного обучения для оценки развернутых ответов.

Архитектура онлайн-платформы

Онлайн-платформа по подготовке к ВПР представляет из себя многофункциональный веб-сайт, состоящий из бэкенда и фронтенда. Бэкенд же, в свою очередь, состоит из веб-сервера, который отвечает за обработку авторизации, ответы на задания, передачу информации между обучающимися и преподавателями, сбор и отображение статистики. Фронтенд с помощью языков разметки и скриптового языка обеспечивает пользователя удобным интерфейсом веб-сайта, включающим интерактивные элементы и современное оформление.

В платформе типы задания будут разными, в соответствии с ВПР.

Большая часть типов заданий подлежат к простой обработке ответов:

– тест с одинарным ответом;

– тест с множественными ответами;

– тест с проверкой точного строкового совпадения (побуквенной идентичности ответа);

– задание с установками соответствий; – задание с заполнением таблиц.

Также, будет тип задания, который позволяет оценивать развернутый ответ. Однако, оценить такой ответ простыми алгоритмами будет крайне сложно или вообще невозможно.

Векторная модель поиска

Для оценки развернутых ответов, машине нужно понимать, с чем имеет дело. Компьютер не понимает строковые символы, изображения или ещё что-то абстрактное, лишь только нули и единицы. Для этого используется эмбеддинг – процесс преобразования абстрактных данных в набор чисел [5]. Этот набор чисел является вектором, которое располагается в векторном пространстве (например, двумерная координатная плоскость).

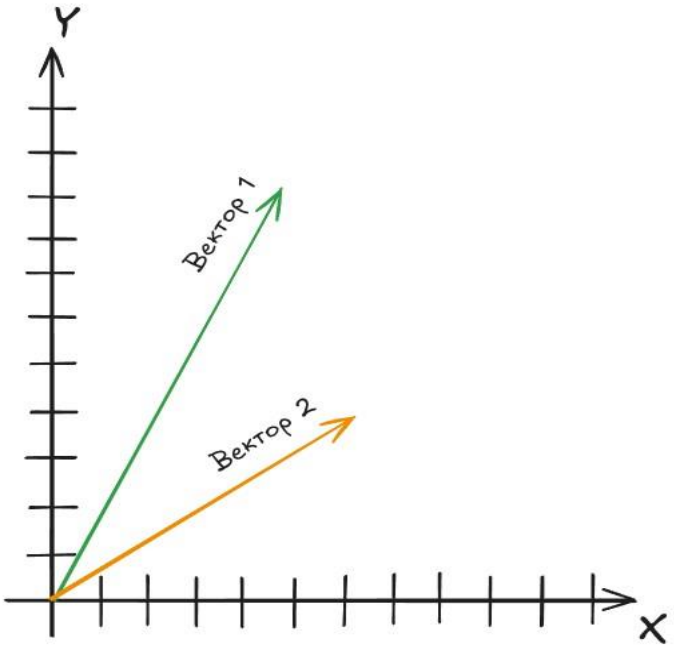

Рисунок 1 – Вектора на координатной плоскости

В нашем случае, для преобразования текста в эмбеддинга, воспользуемся алгоритмом TF-IDF – статистической мерой, которая используется для оценки важности термина в документе, являющемся частью коллекции документов или

корпуса [6]. Важность термина возрастает пропорционально частоте его появления в документе, но компенсируется частотой термина в корпусе. Его преимущества – это: учитывать не только частоту термина, но и его редкость; снижать влияние общих терминов, не несущие различительной информации; увеличивает вес специфичных терминов; создает более информативные векторные представления.

Расчёт веса термина производится как произведение локальной и глобальной компонент:

$$w_{t,d} = tf(t,d) \cdot idf(t) \quad (1)$$

Формула TF-IDF состоит из двух компонентов:

– TF (Term Frequency, tf(t ,d)) – насколько часто встречается термин в документе (2);

– IDF (Inverse Document Frequency, idf(t)) – насколько уникален этот термин во всем корпусе. Чем в большем количестве документов встречается термин, тем меньше его вес (3).

$$tf(t,d) = \frac{f_{t,d}}{\sum_k f_{k,d}} \quad (2)$$

$$w_{t,d} = tf(t,d) \cdot idf(t) \quad (3)$$

На выходе получается матрица весов терминов (рисунок 2).

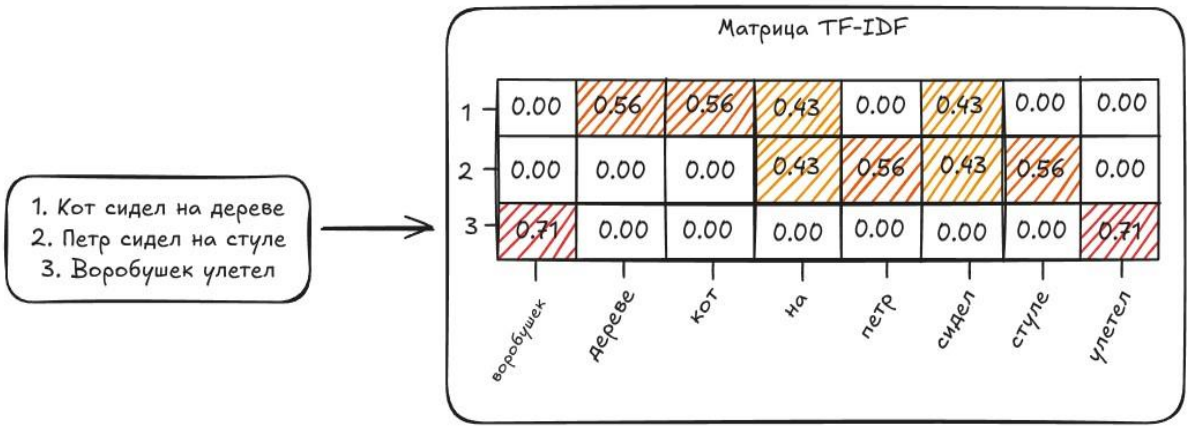

Рисунок 2 – Матрица TF-IDF

Дабы понять, являются ли два текста (документа), представленные в виде векторов, близкие по смыслу, воспользуемся метрикой, измеряющая угол между двумя векторами – косинусное сходство [7]. Чем меньше угол, тем больше сходство. Для текстовых данных используется положительное пространство, поэтому диапазон является [0, 1]. Математически это выражается через скалярное произведение и норму веткоров:

$$cos(\theta) = \frac{a \cdot b}{|a| * |b|} \quad (4)$$

где: $a \cdot b$ – скалярное произведение векторов;

$|a|$и$|b|$ – нормы (длины) векторов;

θ – угол между векторами.

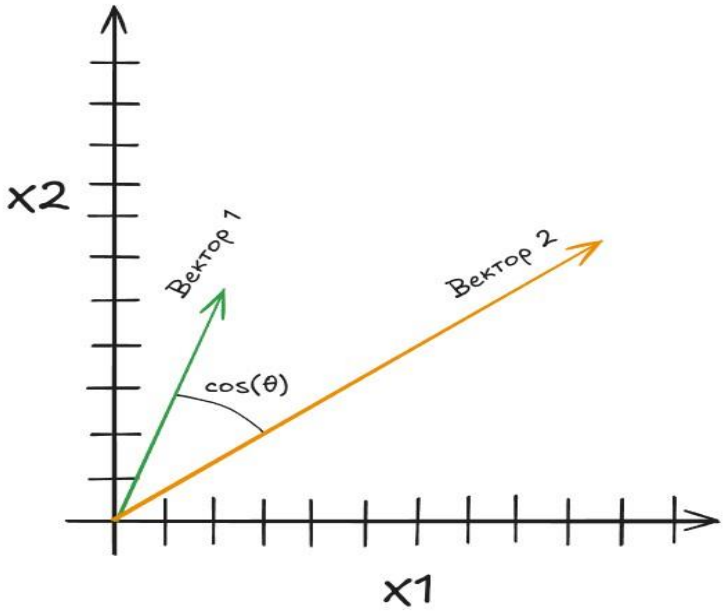

Рисунок 3 – Угол между двумя векторами

Ключевое преимущество косинусного сходства

Главный плюс косинусного сходства – оно не зависит от масштаба. Если векторы смотрят в одну сторону, но сильно отличаются по длине, сходство всё равно будет высоким. Ещё эта метрика устойчива к шуму и словам, которые не несут смысла. Она хорошо ловит фундаментальные семантические закономерности в текстовых эмбеддингах и не ломается от лексического мусора и стоп-слов. Поэтому её можно применять к неструктурированным текстам, где полно нерелевантных выражений и фоновых помех [8].

Ограничения векторных моделей и что с ними делать

Алгоритмы вроде TF-IDF простые и популярные, но у них есть серьёзные минусы. Они сильно портят качество, когда нужно автоматически оценивать развёрнутые ответы. TF-IDF работает только с поверхностными лексическими совпадениями. Если общих слов нет, она не поймёт, что тексты близки по смыслу. Один и тот же смысл, выраженный разными словами и конструкциями, приводит к резкому падению косинусной меры. Многозначные термины TF-IDF тоже не

различает – может завысить совпадение там, где это не нужно. И главное: векторные модели оценивают общее сходство признаков, но не смотрят на порядок изложения, причинно-следственные связи и качество аргументации. А это ключевые вещи, когда мы оцениваем развёрнутые ответы в гуманитарных и социальных дисциплинах [9].

Для решения этой проблемы, можно было бы попробовать создать нейронную сеть с «нуля» и обучать её, но для таких задач это крайне невыгодно как по времени, так и экономически: необходимо иметь мощное оборудование и огромный датасет, чтобы эффективно обучать модель. Использования готовой большой модели с огромными миллиардными параметрами также потребует мощное оборудование и много времени на обработку развернутых ответов.

Готовые семантические эмбеддинги (модели) больше подходят для решения таких задач. Они:

- не требуют обучения;
- понимают семантику, перефразирование, синонимы;
- точность на 30–50% выше TF-IDF на реальных данных;

Замена TF-IDF на предобученные семантические эмбеддинги даст максимальный прирост качества с минимальными усилиями. Однако, недостаточно просто внедрить модель и сразу её использовать.

Трансферное обучение

Всегда есть вероятность, что модель выдаст неверный результат при оценки развернутого ответа, который всё же является правильным. Чтобы предотвращать такие ситуации, модель нужно дообучать – трансферное обучение. Это метод, который просто дообучает существующую ML-модель для решения других задач [10]. Разница классического и трансферного обучения представлена в таблице 1.

Более подходящая модель под нашу задачу подойдет «paraphrasemultilingual-MiniLM-L12-v2». Благодаря таким преимуществам, как предобучение на задачах семантического сходства и перефразирования, компактная архитектура и поддержка более 50 языков позволяет обеспечивать устойчивость к вариативности формулировок определений и выполнять как инференсию, так и периодическое дообучение исключительно на CPU без потребления видеопамяти. Модель может работать на чистом CPU благодаря её компактности и небольшого числа параметров, что делает её применимой на стандартном учебном сервере без GPU [11, 12].

Существующие модели для трансферного обучения представлены в таблице 2.

Таблица 1 – Разница классического и трансферного обучения

	Классическое обучение	**Трансферное обучение**
Обучение	Нужно с «нуля»	На основе модели, обученной на датасете
Вычислительные затраты	Высокие	Низкие
Необходимый объем данных	Большой	Малый
Использование знаний	Каждая модель обучается независимо	Использует знания из предварительно обученной модели
Время достижения оптимальной производительности	Длительное	Быстрое
Эффективность	Не эффективна в условиях ограниченных ресурсов	Эффективна в условиях ограниченных ресурсов

Таблица 2 – Мультиязычные модели семантических эмбеддингов

Модель	**Языки**	**Параметры**	**Особенности**
paraphrase-multilingualMiniLM-L12-v2	50+	33M	Оптимизирована под семантическое сходство и перефразирование; минимальные требования к CPU
multilingual-e5-small	100+	110M	Поддержка query/document-режима; быстрая инференсия на CPU
paraphrase-multilingualmpnet-base-v2	50+	278M	Высокая точность на семантических задачах; устойчивость к перефразированию; требуется ≥ 4 ядер CPU для дообучения
multilingual-e5-base	100+	278M	Универсальность; поддержка длинных текстов до 512 токенов

На пилотном этапе модель будет ограничена в рамках ответов, которые связаны с дисциплиной «Информатика», охватывающие определения терминов и описание архитектурных концепций. На этих данных модель будет тестироваться и адаптироваться методом трансферного обучения.

Для более эффективного обучения модели будет использоваться система подачи апелляции. Если пользователь после прохождения теста увидит, что его развернутый ответ оценен как «неверный», хотя является верным, то он подаёт

апелляцию. Администратор или преподаватель увидит данную апелляцию и ему необходимо принять решение, является ли ответ верным или нет. После принятия решения, оно вместе с развернутым ответом отправляется на дообучение модели.

Будущая актуальность

При успешной валидации модели на дисциплине «Информатика» планируется расширение системы на другие предметные области ВПР, такие как естественно-научный цикл, математические и гуманитарные науки.

Для повышения доступности платформы будет реализована полноценная поддержка английского языка, что обеспечит работу с иностранными обучающимися и образовательными организациями стран СНГ. Функционал отчётности будет дополнен расширенной аналитикой (динамика освоения тем, типовые ошибки по классам) и возможностью экспорта отчётов в форматы DOCX, PDF и ODT для интеграции в документооборот учебного заведения.

Ключевым элементом масштабируемости станет внедрение внешнего REST API, позволяющего обеспечивать обмен данными с корпоративными системами управления обучением образовательной организации (напр. система «Гиперион» от ВГЛТУ), а также интеграцию с национальным мессенджером MAX через образовательного бота с поддержкой WebApp-платформы. Это обеспечит доступ к заданиям и результатам в привычной для обучающихся среде мгновенных сообщений, повысит вовлечённость и снизит барьер входа в образовательную платформу.

Заключение

В статье мы показали, почему классические алгоритмы вроде TF-IDF и косинусного сходства лучше заменить на предобученные семантические эмбеддинги, когда речь идёт об автоматической оценке развёрнутых ответов. Модель «paraphrase-multilingual-MiniLML12-v2» даёт оптимальный баланс – точность выросла на 30–50% по сравнению с TF-IDF, она устойчива к пересказу одних и тех же мыслей разными словами, и при этом не требует мощного железа (нормально работает на CPU, GPU не нужен). Если добавить дообучение на реальных апелляциях преподавателей, модель легко адаптируется под особенности конкретного предмета. Пилотное внедрение на «Информатике» станет первым шагом, а дальше можно будет масштабировать решение на все предметы ВПР и создать единую цифровую среду, где знания будут оценивать объективно.

Список литературы

1. Мероприятия по оценке качества образования – Портал. URL: https://fioco.ru/ru/osoko (дата обращения: 08.02.2026).

2. Эмбеддинги для начинающих. URL: https://habr.com/ru/companies/otus/articles/787116/ (дата обращения (08.02.2026).

3. Косинусное сходство – Метрика сходства текста – Изучите машинное обучение. URL: https://studymachinelearning.com/cosine-similarity-text-similarity-metric/ (дата обращения 08.02.2026).

4. Косинусное сходство: преимущества для встраивания текста. URL: https://www.csharpcorner.com/article/cosine-similarity-benefits-for-text-embeddings/ (дата обращения 08.02.2026).

5. Косинусное сходство – GeeksforGeeks. URL: https://www.geeksforgeeks.org/dbms/cosine-similarity/ (дата обращения 08.02.2026).

6. Трансферное обучение: подробный гайд для начинающих. URL: https://habr.com/ru/companies/skillfactory/articles/835020/ (дата обращения 08.02.2026).

References

1. Education Quality Assessment Activities – Portal. URL: https://fioco.ru/ru/osoko (accessed: 08.02.2026).

2. Embeddings for Beginners. URL: https://habr.com/ru/companies/otus/articles/787116/ (date of access (08.02.2026).

3. Cosine Similarity – Text Similarity Metric – Study Machine Learning. URL: https://studymachinelearning.com/cosine-similarity-text-similarity-metric/ (date of access 08.02.2026).

4. Cosine Similarity: Benefits for Text Embeddings. URL: https://www.csharpcorner.com/article/cosine-similarity-benefits-for-text-embeddings/ (date of access 08.02.2026).

5. Cosine Similarity – GeeksforGeeks. URL: https://www.geeksforgeeks.org/dbms/cosine-similarity/ (date of access February 8, 2026).

6. Transfer learning: a detailed guide for beginners. URL: https://habr.com/ru/companies/skillfactory/articles/835020/ (accessed February 8, 2026).