

**Кузовкин Владимир Валерьевич**

аспирант

ФГАОУ ВО «Томский государственный университет систем управления и  
радиоэлектроники»

г. Томск, Томская область

## **ИНТЕЛЛЕКТУАЛЬНАЯ СИСТЕМА АНАЛИЗА РУКОПИСНЫХ РЕШЕНИЙ ОТКРЫТЫХ ЗАДАЧ КАК ИНСТРУМЕНТ ФОРМИРУЮЩЕГО ОЦЕНИВАНИЯ**

***Аннотация:** в статье представлена веб-ориентированная система автоматизированной проверки математических задач открытого типа, использующая мультимодальные большие языковые модели для анализа фотографий рукописных решений. В отличие от классических систем компьютерного тестирования, которые ограничиваются сравнением конечного ответа, предлагаемая платформа выполняет пошаговый разбор хода решения, классифицирует ошибки (арифметические, логические, неполнота решения) и формирует персонализированную обратную связь в формате «тип ошибки – объяснение – верное решение». Приводятся архитектура системы, принципы промпт-инжиниринга для получения структурированного ответа от ИИ, результаты пилотной апробации на задачах Единого государственного экзамена по математике. Показано, что применение мультимодального ИИ позволяет сократить время проверки преподавателем в несколько раз при сохранении качества обратной связи.*

***Ключевые слова:** искусственный интеллект в образовании, формирующее оценивание, мультимодальные модели, обработка рукописных решений, открытые задачи, ЕГЭ по математике.*

### *Введение*

Современные тенденции цифровизации образования и развития технологий искусственного интеллекта (ИИ) открывают новые возможности для автоматизации рутинных процессов обучения и контроля знаний. Одним из наиболее трудоёмких этапов работы педагога является проверка письменных решений задач,

особенно по математике, где важна не только правильность конечного ответа, но и корректность хода решения, обоснованность логических переходов и отсутствие арифметических ошибок. Традиционные системы компьютерного тестирования (в том числе системы дистанционного обучения) преимущественно ориентированы на задания закрытого типа – с выбором ответа или вводом числа, – что не позволяет оценить глубину понимания предмета и процессуальные аспекты решения [1].

Активное развитие мультимодальных больших языковых моделей (LLM), способных обрабатывать одновременно текст и изображения, создаёт предпосылки для создания интеллектуальных систем нового поколения, способных анализировать рукописные математические решения, выявлять ошибки и предоставлять развёрнутую обратную связь учащемуся. Подобные системы реализуют концепцию формирующего оценивания (formative assessment), которое рассматривается как один из наиболее эффективных инструментов повышения качества образования [2].

Целью данной работы является разработка и экспериментальная апробация веб-системы, использующей мультимодальные LLM для автоматического анализа фотографий ученических решений открытых математических задач, классификации ошибок и генерации персонализированной обратной связи. В статье представлены архитектура платформы, принципы взаимодействия с ИИ-моделями, а также результаты пилотного тестирования на материале задач Единого государственного экзамена по математике.

### *Обзор аналогов*

Существующие системы автоматической проверки математических заданий можно условно разделить на две группы. К первой относятся системы символьной проверки, такие как STACK, WeBWorK, Maple T.A., которые анализируют ответ, представленный в виде формулы или числа, и способны распознавать эквивалентные алгебраические выражения. Однако они не работают с рукописным вводом и не анализируют ход решения [3]. Вторую группу составляют системы на основе нейронных сетей и компьютерного зрения, ориентированные на

распознавание отдельных символов или формул. К этой группе можно отнести приложения Photomath и Microsoft Math Solver, которые фокусируются на выдаче готового ответа и пошагового алгоритма, но не выполняют проверку ученического решения с диагностикой конкретных ошибок [4].

Среди российских образовательных платформ, активно применяемых при подготовке к Единому государственному экзамену, следует выделить проект «РешуЕГЭ» Д.Д. Гущина, открытый банк заданий ФИПИ (в том числе сервис «КЕГЭ» для подготовки к ЕГЭ по информатике и ИКТ), а также «Яндекс.Репетитор». «РешуЕГЭ» предоставляет обширный каталог задач с автоматической проверкой ответов тестовой части и эталонными решениями для заданий с развёрнутым ответом, однако проверка второй части по-прежнему остаётся ручной [5]. Открытый банк заданий ФИПИ, включая интерактивный тренажёр «КЕГЭ», обеспечивает доступ к актуальным экзаменационным прототипам, но также не содержит инструментов автоматического анализа рукописных решений [6]. «Яндекс.Репетитор» предлагает персонализированные тренировочные варианты с мгновенной проверкой тестовой части, тогда как задания с развёрнутым ответом проверяются экспертами, а встроенные механизмы самопроверки основаны исключительно на сравнении конечного ответа с эталоном [7]. Таким образом, ни одна из широко распространённых российских платформ не предоставляет функциональности автоматического анализа фотографий рукописных черновиков и формирования индивидуального диагностического заключения по ходу решения.

В последние годы появились научные работы, исследующие применение больших языковых моделей для оценки математических рассуждений, однако использование мультимодальных моделей для анализа именно рукописных работ в контексте отечественной системы образования остаётся малоизученной областью [8]. Данное исследование восполняет указанный пробел, предлагая законченное веб-решение, интегрированное с облачным хранилищем изображений и мультимодальными ИИ-моделями, и апробированное на задачах ЕГЭ по математике.

### *Постановка задачи*

При разработке системы мы опирались на формальную модель учебной задачи для электронного обучения, предложенную в [9]. Согласно этой модели, задача представляется в виде кортежа, включающего идентификатор, текст условия, правильный ответ, классификацию по темам и экзаменам, характеристики типичных ошибок и другие метаданные. Использование такой модели позволило унифицировать описание каждой задачи, обеспечить её корректную передачу в промпт мультимодальной языковой модели и организовать хранение результатов анализа.

С учётом указанной модели задача автоматического анализа рукописного решения формулируется следующим образом. Пусть имеется множество математических задач открытого типа  $D$ , для каждой задачи  $d \in D$  известен текст условия  $T(d)$  и правильный ответ  $A_{\text{true}}(d)$ . Ученик предоставляет рукописное решение в виде одного или нескольких изображений  $I = \{i_1, i_2, \dots, i_n\}$ . Требуется на основе изображений  $I$  и информации о задаче  $d$  с помощью мультимодальной LLM сгенерировать структурированный отклик, содержащий:

- 1) тип ошибки (`error_type`) – классификацию ошибки (арифметическая, логическая, неполное решение и т. п.) или вердикт «Решение верно»;
- 2) пошаговый анализ (`reasoning`) – детальное объяснение, почему решение является неверным или верным, с возможным использованием LaTeX-нотации для формул;
- 3) правильное решение (`correct_solution`) – корректный ход решения, приводящий к верному ответу.

Дополнительным требованием является сохранение всех изображений и результатов анализа в облачном хранилище для возможности последующего просмотра учителем и формирования отчётов.

### *Архитектура системы*

Предлагаемая платформа продолжает развитие подхода, заложенного в модели учебной задачи [9–10], и реализует концепцию Data-Driven-образования, предполагающую сбор, накопление и анализ данных о каждом действии ученика. Архитектура строилась таким образом, чтобы объединить классическую базу

знаний (задачи, ответы, рубрикаторы) с современными мультимодальными ИИ-сервисами, обеспечив при этом масштабируемость и отказоустойчивость. Система состоит из пяти основных слоёв.

1. *Клиентский слой (frontend)*. Адаптивный веб-интерфейс реализован на HTML5/CSS3/JavaScript с использованием библиотеки Vulma CSS. Он обеспечивает:

- постраничный просмотр заданий теста с динамической пагинацией (AJAX-подгрузка без перезагрузки страницы);
- ввод числового ответа в текстовое поле;
- загрузку до четырёх фотографий рукописного решения на каждую задачу;
- отображение таймера, прошедшего с начала тестирования;
- предпросмотр загруженных изображений с возможностью удаления.

Клиентская часть активно использует JavaScript для асинхронного сохранения ответов и файлов, что позволяет фиксировать состояние ученика в реальном времени.

2. *Серверный слой (backend)*. Ядро системы написано на Python с использованием микрофреймворка Flask и ORM SQLAlchemy. Сервер отвечает за:

- управление жизненным циклом теста (старт, пагинация, завершение);
- сохранение текстовых ответов и метаданных в серверной сессии (test\_answers, test\_ai\_results, test\_photos);
- маршрутизацию запросов к ИИ-моделям и обработку их ответов;
- формирование итоговой страницы результатов, где сводятся ответы, фотографии и анализ ИИ.

Взаимодействие с базой данных организовано через модели SQLAlchemy, что упрощает расширение схемы хранения.

3. *Слой облачного хранения (S3-совместимый)*. Все загруженные учеником изображения конвертируются в формат WebP (для экономии трафика) и сохраняются в облачное объектное хранилище с публичным доступом. Ключ формируется по шаблону test\_uploads/<session\_id>/<task\_id>/<uuid>.webp, что гарантирует уникальность и изоляцию данных разных сессий. Публичные URL

изображений затем передаются мультимодальным моделям в качестве входных данных, а также отображаются на странице результатов в виде миниатюр.

4. *Интеграционный слой с ИИ.* Для анализа рукописных решений используются мультимодальные большие языковые модели, доступные через API (OpenRouter, NVIDIA API). В текущей версии задействованы microsoft/phi-4-multimodal-instruct (быстрый первичный анализ) и модели семейства NVIDIA Nemotron (глубокий разбор с повторными попытками). Запросы отправляются с экспоненциальной задержкой (до 5 попыток) для компенсации временной недоступности сервисов. Промпт инженерия описана в следующем подразделе.

5. *Слой постобработки.* Ответ ИИ, полученный в виде текста, проходит через парсер, выделяющий JSON-структуру с полями `error_type`, `reasoning` и `correct_solution`. Если модель возвращает некорректный JSON, производится повторный запрос; на последней попытке ответ используется «как есть» с соответствующим комментарием. Результат сохраняется в сессионной переменной и затем встраивается в страницу результатов без дополнительных AJAX-вызовов, что снижает нагрузку на сервер и улучшает пользовательский опыт.

Таким образом, архитектура платформы обеспечивает сквозную автоматизацию процесса от предъявления открытой задачи до предоставления персонализированной обратной связи, опираясь при этом на единую модель учебной задачи [9] и накапливая данные для дальнейшего совершенствования алгоритмов. На рис. 1. Показан дизайн взаимодействия с платформой.

КUZOVKIN

Обучение ▾ Новости ▾ Онлайн доска Веб зазеркалье Партнёрам

Прошло времени: 00 : 00 : 23

Закончить тестирование

Показывать заданий: 1 3 5 10 15

Предыдущая 1 Следующая

**1. № 16651**  
Решите уравнение:  $\frac{2x+1}{5} = 1$

Вести число Загрузить фото

Ваш ответ: 2

**2. № 6**  
Велосипедист отъехал от станции в тот момент, когда пешеход отошел от нее на 1,6 км, и через 15 мин догнал пешехода. С какой скоростью шел пешеход, если велосипедист ехал в  $2\frac{1}{3}$  раза быстрее?

Вести число Загрузить фото

Ваш ответ: Введите число

**3. № 14**  
Расстояние между двумя пристанями равно 12,3 км. За сколько времени моторная лодка проплывет путь от одной пристани до другой и обратно, если собственная скорость лодки  $7,2 \frac{км}{ч}$ , а скорость течения реки составляет  $\frac{1}{6}$  скорости лодки?

Рис. 1. Вид редактора

КUZOVKIN

Обучение ▾ Новости ▾ Онлайн доска Веб зазеркалье Партнёрам

Затраченное время: 00:00:25

| №  | Идентификатор | Правильный ответ | Полученный ответ | Статус    |
|--|---------------|------------------|------------------|-----------|
| 1  | 16651         | 2                | Неверно          | Неверно   |
| <p><b>№ 16651 – УСЛОВИЕ ЗАДАЧИ</b><br/>Решите уравнение: <math>\frac{2x+1}{5} = 1</math></p> <p><b>ЧЕРНОВИК УЧЕНИКА</b></p> <p>Автоматический разбор ошибок ИИ</p> <p><b>Тип проблемы:</b> Арифметическая ошибка при делении (вместо деления на 2 выполнено умножение на 2)</p> <p><b>Анализ шагов:</b> Решение уравнения начинается верно: <math>\frac{2x+1}{5} = 1 \rightarrow</math> умножаем обе части на 5 <math>\rightarrow 2x + 1 = 5</math>. Затем вычитаем 1: <math>2x = 4</math>. На этом этапе всё правильно. Однако далее в решении написано <math>x = 4 \cdot 2</math>, что неверно – должно быть <math>x = \frac{4}{2} = 2</math>. Вместо деления на 2, ученик умножил 4 на 2, получив 8. Это арифметическая ошибка при делении.</p> <p><b>ПОЛНОЕ РЕШЕНИЕ ИЗ БАЗЫ ЗНАНИЙ</b></p> <p><math>\frac{2x+1}{5} = 1 \rightarrow 2x + 1 = 5 \rightarrow 2x = 4 \rightarrow x = \frac{4}{2} = 2</math> Ответ: <math>x = 2</math>.</p> |               |                  |                  |           |
| 2  | 6             | 4                | 4,8              | Неверно   |
| 3  | 14            | —                | 3:514            | Пропущено |

← В избранное Работа над ошибками Похожий тест? Сгенерировать аналоги (AI)

Рис. 2. Вид результатов теста

### Экспериментальная апробация

Пилотное тестирование системы было проведено на базе авторской платформы *kuzovkin.info*. В эксперименте приняли участие учащиеся 10–11 классов (15 учеников), выполнявшие подборку из  $M$  задач открытого типа,

соответствующих кодификатору ЕГЭ по математике профильного уровня. Каждое решение фотографировалось и загружалось в систему. После завершения тестирования для всех задач с загруженными фото автоматически вызывался ИИ-анализ.

Оценивались следующие метрики:

- точность классификации ошибок (совпадение с экспертной оценкой преподавателя);
- полнота разбора (наличие всех обязательных полей);
- время обработки одного решения (от момента отправки фото до получения ответа ИИ);
- удовлетворённость учащихся полученной обратной связью (по 5-балльной шкале).

Предварительные результаты показали, что система корректно определяет наличие ошибки и её тип в 95% случаев. Среднее время анализа одного решения составило 10 секунд. Большинство учащихся отметили полезность автоматических комментариев и выразили желание использовать систему при подготовке к экзаменам. Учитель, участвовавший в экспертной оценке, подтвердил, что автоматический анализ позволяет сократить время проверки работ ориентировочно в 3–5 раз.

#### *Обсуждение и выводы*

Предложенная система демонстрирует принципиальную возможность применения мультимодальных больших языковых моделей для формирующего оценивания математических решений в условиях реального учебного процесса. В отличие от традиционных автоматизированных систем, она не ограничивается констатацией «верно/неверно», а предоставляет развёрнутую диагностическую информацию, способствующую осознанной коррекции ошибок учащимся.

К ограничениям текущей реализации можно отнести зависимость от качества и доступности сторонних ИИ-API, а также необходимость дальнейшего повышения точности классификации специфических ошибок (например, связанных с неверной интерпретацией условия). Перспективы развития включают

интеграцию с популярными LMS (Moodle, Сферум), добавление голосового ввода и расширение списка поддерживаемых предметов.

Практическая значимость работы заключается в создании действующего прототипа, который может быть внедрён в образовательные организации в качестве инструмента поддержки учителя и самостоятельной подготовки учащихся. Результаты исследования могут быть использованы при проектировании интеллектуальных обучающих систем нового поколения.

### *Список литературы*

1. Black P. Assessment and Classroom Learning / P. Black, D. Wiliam // Assessment in Education: Principles, Policy & Practice. – 1998. – Vol. 5, № 1. – P. 7–74.
2. Bennett R.E. Formative assessment: a critical review / R.E. Bennett // Assessment in Education: Principles, Policy & Practice. – 2011. – Vol. 18, № 1. – P. 5–25.
3. Sangwin C.J. Computer Aided Assessment of Mathematics / C.J. Sangwin. – Oxford : Oxford University Press, 2013. – 272 p.
4. Photomath – URL: <https://photomath.com> (дата обращения: 06.06.2025).
5. Гущин Д.Д. РешуЕГЭ : образовательный портал для подготовки к экзаменам / Д.Д. Гущин. – URL: <https://ege.sdangia.ru> (дата обращения: 06.06.2025).
6. Федеральный институт педагогических измерений : открытый банк заданий ЕГЭ. – URL: <https://fipi.ru/ege/otkrytyy-bank-zadaniy-ege> (дата обращения: 06.06.2025).
7. Яндекс.Репетитор : сервис для подготовки к ЕГЭ и ОГЭ. – URL: <https://yandex.ru/tutor> (дата обращения: 06.06.2025).
8. Mathematical Capabilities of ChatGPT / S. Frieder, L. Pinchetti, R.-R. Griffiths [и др.] // Advances in Neural Information Processing Systems. – 2024. – Vol. 36. – P. 1–10.
9. Кручинин В.В. Модель учебной задачи в электронной системе обучения и ее наполнение / В.В. Кручинин, В.В. Кузовкин // Доклады Томского государственного университета систем управления и радиоэлектроники. – 2023. – Т. 26. – № 2. – С. 93–100. – DOI 10.21293/1818-0442-2023-26-2-93-100. – EDN AHSUFC

10. Кузовкин В.В. Программное обеспечение обучающего веб-портала для учителей, репетиторов и учеников / В.В. Кузовкин // Информационные технологии. – 2024. – Т. 30. – № 4. – С. 214. – DOI 10.17587/it.30.214-223. – EDN VSJTIX