

ПЛЕНАРНЫЕ ДОКЛАДЫ

АВТОМАТИЗАЦИЯ РЕЦЕНЗИРОВАНИЯ ОТЧЕТОВ ПО ЛАБОРАТОРНЫМ РАБОТАМ С ИСПОЛЬЗОВАНИЕМ БОЛЬШИХ ЯЗЫКОВЫХ МОДЕЛЕЙ

О. Д. Асенчик, В. А. Воробьев

Учреждение образования «Гомельский государственный технический университет имени П.О. Сухого», Республика Беларусь

В условиях цифровой трансформации высшего технического образования происходит активное внедрение инновационных технологий в учебный процесс с целью повышения его эффективности. Одним из трудоемких аспектов в деятельности преподавателя IT-дисциплин остается проверка отчетов по лабораторным работам. Особенность таких работ заключается в сложности структуры и большом количестве уникальных вариантов заданий. Отчеты студентов по программированию, базам данных или сетевым технологиям содержат не только текстовые описания, но и фрагменты программного кода, таблицы, диаграммы, SQL-запросы и результаты их выполнения. Проверка таких работ требует от преподавателя оценки соответствия формальным требованиям, анализа логики представленных решений, а также корректности синтаксиса и семантики кода, что при большом количестве студентов и вариантов заданий приводит к значительным временным затратам и риску субъективной оценки.

Появление и развитие больших языковых моделей (LLM) открывает новые возможности для автоматизации анализа и оценки текстов и программного кода при его наличии. Современные LLM демонстрируют высокую эффективность анализа контента благодаря их способности к глубокому семантическому разбору сложных технических текстов, обработке структурированных данных (кода, формул) и генерации содержательных оценочных суждений, основанных на выявлении логических взаимосвязей в представленном материале. Важным является возможность LLM выявлять не только очевидные ошибки, но и скрытые проблемы, такие как нарушение принципов проектирования в коде или логические несоответствия в объяснениях студентов, что значительно превышает возможности традиционных автоматизированных систем проверки. LLM демонстрируют высокую способность к пониманию контекста и выявлению логических связей в текстах [1], что делает их особенно подходящими для оценки работ, содержащих текст и код.

Прямое применение общедоступных моделей через окно чата является достаточно трудозатратным и требует активного участия пользователя и совершения множества повторяющихся действий. В связи с этим создание апробированного интерактивного инструмента, который бы автоматизировал процесс рецензирования отчетов по лабораторным работам, предоставляя необходимые удобные сервисные возможности для управления проверками, хранения их истории и адаптации под конкретные учебные курсы представляется нам актуальным.

Целью исследования является разработка и апробация программного комплекса для автоматизированного рецензирования отчетов по лабораторным работам студентов с использованием больших языковых моделей, направленного на снижение нагрузки на преподавательский состав и повышение качества, скорости и объективности обратной связи.

Разработанный программный комплекс представляет собой Web-приложение. Бизнес-логика реализована с использованием фреймворка NestJS на платформе Node.js, а графический интерфейс в виде одностраничного приложения – с использованием библиотеки React.js. Для хранения информации (курсы, группы, студенты, лабораторные работы, промпты, модели, провайдеры, API-ключи, результаты проверок) используется встраиваемая СУБД SQLite. Такой стек обеспечивает кроссплатформенность и удобство развертывания как в компьютерной сети, так и на отдельном компьютере.

Программный комплекс не привязан к конкретной LLM, а предоставляет интерфейс для подключения различных LLM через систему провайдеров и API-ключей. Поддерживается работа как с облачными моделями через OpenAI API, так и с моделями, развернутыми локально с помощью платформы Ollama. Для обеспечения надежности взаимодействия с внешними сервисами реализован механизм повторных запросов с настраиваемой задержкой при ошибках.

Процесс проверки включает следующие шаги:

Предварительная настройка: создание дисциплины (дисциплин) и промптов для проверки, создание лабораторных работ и загрузка текстов заданий по ним; выбор и настройка параметров подключения к сервисам, предоставляющим доступ к LLM, задание параметров их работы (размер контекста, температура и др.).

Система будет генерировать запрос к LLM на основе задания к лабораторной работе, текста отчета студента и специального промпта курса, который задает общие критерии оценки. Промпт инструктирует модель вернуть ответ, содержащий оценку, а также текстовый отзыв.

Пример использованного промпта: «Вы преподаватель в университете и эксперт по базам данных и системам управления базами данных, информационным системам и разработке программных приложений для них. Вы выдали студенту задание для выполнения лабораторной работы по учебной дисциплине «Базы данных». Оцени предоставленный студентом отчет по лабораторной работе: на соответствие выполненного задания исходному заданию с учетом варианта студента; правильность и полноту выполнения задания. Выставь оценку отчету студента по шкале от 1 до 10. Оценка ниже 4 выставляется в случае несоответствия варианту или выполнения не всех пунктов задания. Напиши отзыв на выполненную работу и укажи недостатки, если они есть».

Особенностью системы является возможность настройки текстов промптов под конкретные требования дисциплин и заданий, что позволяет адаптировать анализ к различным типам лабораторных работ.

Предобработка данных. Архив с работами студентов скачивается из системы Moodle, в которой они публиковали свои отчеты и который имеет определенную структуру. Загруженный архив обрабатывается: отчеты в форматах .docx и .pdf конвертируются в текст с помощью специализированных библиотек; из имен файлов и папок извлекается информация о студентах (Ф.И.О.), которая сопоставляется с записями в базе данных.

Выполнение проверки. Реализовано два режима проверки, выбор между которыми осуществляется автоматически на основе паттерна проектирования «Стратегия». Первый – базовый, с использованием одной выбранной LLM. Второй – комплексный, при котором один и тот же отчет анализируется несколькими моделями параллельно. Этот подход позволяет нивелировать возможные слабости отдельных моделей.

При повторной проверке система сравнивает новую версию отчета со старой, определяя, были ли исправлены ранее выявленные недостатки, и формирует новую рецензию и оценку.

Обработка и сохранение результата. Ответ модели анализируется на соответствие ожидаемому формату, валидируется и сохраняется в базе данных. Преподавателю через веб-интерфейс становятся доступны все результаты проверки, сгруппированные по группам и студентам.

Основным результатом исследования является разработанный программный комплекс. Его код и подробное описание приведено в [2]. В качестве используемых для проведения экспериментов LLM были выбраны DeepSeek V3, DeepSeek R1, Llama 3.3 70B, Gemini 2.5 Flash, Qwen QwQ [3]. Все они доступны через API различных провайдеров. Политика ценообразования для запросов с использованием API достаточно гибкая. Ко всем из этих моделей у определенных провайдеров можно получить бесплатный или относительно недорогой доступ, с качеством, приемлемым для решения задач настоящей работы [4].

В ходе апробации проводилась верификация работы программного комплекса на реальных отчетах студентов по дисциплине «Базы данных». Отчеты содержали поясняющие тексты, таблицы, SQL-скрипты и текстовые результаты их выполнения, фрагменты кода на C#. В ходе экспериментов наибольшую эффективность продемонстрировал Gemini 2.5, особенно для анализа объемных отчетов. DeepSeek-R1 также продемонстрировал хорошие результаты при работе с кодом и при анализе текста.

В процессе проверок система успешно выявляла факты несоответствия заданию, синтаксические ошибки в коде и логические недочеты, такие как, например, некорректная нормализация базы данных, неоптимальные SQL-запросы или способы работы с базой данных с использованием C#. Сформированные моделью списки замечаний и сильных сторон в значительной степени совпадали с теми, которые давал бы преподаватель при внимательном рецензировании. Рецензии Gemini 2.5 отличались исключительной подробностью и глубиной.

Следует отметить особую важность структурирования выданных студентам заданий и содержащиеся в них требования к структуре и форме представления результатов в отчете. В случае недостаточной структурированности, ясности и полноты задания и требований к отчетам LLM могут упускать специфический контекст, заложенный преподавателем в задание. Также LLM не способны напрямую анализировать графический материал, что затрудняет проверку работ с диаграммами и схемами.

Разработанный программный комплекс является мощным и удобным инструментом, который может использоваться для рецензирования работ по различным дисциплинам. При этом решение об окончательной оценке должно оставаться за преподавателем, который использует систему для первичного анализа и формирования развернутой обратной связи. Внедрение данного комплекса в учебный процесс позволит, на наш взгляд, достичь значимых эффектов. Снижается нагрузка на преподавателя за счет автоматизации рутинной части проверки, что высвобождает время для более качественной подготовки к занятиям и индивидуальной работы со студентами. Повышается качество и скорость обратной связи: студенты получают детальный и структурированный разбор своих работ в короткие сроки, что способствует более эффективному освоению материала и своевременному исправлению ошибок. Обеспечивается объективность и стандартизация оценки на этапе первичного разбора работ.

Литература

1. The emergence of large language models as tools in literature reviews: a large language model-assisted systematic review / D. Scherbakov, N. Hubig, V. Jansari [et al.] // Journal of the American Medical Informatics Association. – 2025. – Vol. 32 (6). – P. 1071–1086.
2. Reports-check / Software Suite for Reviewing Student Laboratory Reports. – URL: <https://github.com/Olgasn/reports-check> (дата обращения: 10.09.2025).
3. Models / Artificial Analysis. – URL: <https://artificialanalysis.ai/-leaderboards/models> (дата обращения: 10.09.2025).
4. OpenRouter / OpenRouter. – URL: <https://openrouter.ai/> (дата обращения: 10.09.2025).

РЕАЛИЗАЦИЯ СЕТЕВОГО СОТРУДНИЧЕСТВА ПОЛИТЕХНИЧЕСКОГО УНИВЕРСИТЕТА С РЕГИОНАЛЬНЫМИ ШКОЛАМИ

Н. А. Леонова

Федеральное государственное автономное образовательное учреждение высшего образования «Санкт-Петербургский политехнический университет Петра Великого», Российская Федерация

Санкт-Петербургский политехнический университет Петра Великого действует как «драйвер технологического суверенитета», сочетая фундаментальные исследования, образовательные инновации и промышленную кооперацию. Его роль определена в стратегиях национального развития до 2036 г. и подтверждается конкретными проектами: от цифровых двойников до подготовки инженеров новой формации.

В эпоху технологического суверенитета политехнический университет реализует стратегию сетевого взаимодействия со школами, центрами дополнительного образования для ранней инженерной подготовки. Особое внимание уделено взаимодействию с «региональными школами» – образовательными организациями, находящимися в малых населенных пунктах и удаленные от Федеральных образовательных учреждений. Это взаимодействие не только формирует мотивированный поток абитуриентов, готовых к современным вызовам науки, но и повышает престиж физико-математического образования среди школьников.

В основе модели сетевого сотрудничества находится триада «школа – вуз – промышленность». В рамках данной модели университет ставит перед собой и успешно решает три ключевых задачи:

- повышение качества физико-математического образования школьников через интеграцию ресурсов вуза, школ и предприятий;
- ранняя профориентация с акцентом на инженерные специальности (ИТ, робототехника, биотехнологии);
- ликвидация разрыва между школьной программой и требованиями вуза/работодателей.

Ключевые форматы сотрудничества с региональными школами включают:

а) практико-ориентированные проекты для школьников.

Лабораторные практикумы: Использование высокотехнологичного оборудования вуза (лазерные установки, вакуумные камеры, стенды для изучения сверхпроводимости).

Инженерные проекты:

- расчет эффективности ветрогенераторов для Ленинградской области (физика + экология);
- разработка ИТ-решений для производственных задач предприятий-партнеров;
- б) программы для педагогов.